UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

TECHNOLOGY FORECASTING FOR NATIONAL RESEARCH AND EDUCATION NETWORKS USING STRUCTURAL EQUATION MODELLING-BASED CONTEXT-SENSITIVE DATA FUSION

Leonard Staphorst

Submitted in Fulfilment of the Requirements for the Degree
Doctor of Philosophy
through the
Department of Engineering and Technology Management
in the
Graduate School of Technology Management
of the
Faculty of Engineering, Built Environment and Information Technology
at the
University of Pretoria

Under the Supervision of
Professor Leon Pretorius
and
Professor Marthinus Pretorius

July 2024

ABSTRACT

Technology Intelligence entails an intricate process of gathering and transforming data related to technological advancements into refined actionable intelligence by identifying and analysing emergent characteristics and interrelational linkages. Specialists in this field process this refined information to develop knowledge critical for directing strategic technology management decisions. As primary sources of technology-related data, technology indicators facilitate an expansive characterisation and evaluation of various technologies throughout their entire lifecycle. Engaging in Future-oriented Technology Analysis necessitates a rigorous examination of information from these indicators, equipping decision-makers with sophisticated insights for Technology Forecasting, a vital tool in anticipating and preparing for future technology trends and developments.

This study posits that one can conceptualise Technology Forecasting as a context-sensitive Data Fusion process using Structural Equation Modelling. To this end, this study undertook inductive reasoning to develop generic frameworks for Structural equation Modelling, context-sensitive Data Fusion and the relational mapping of technology indicators for Technology Forecasting, employing analytic literature assessment, conceptual framework development and grounded theory as supportive research methodologies. These generic frameworks were then integrated through inductive reasoning, incorporating comparative and cross-disciplinary analyses and framework unification to develop the study's proposed framework for Technology Forecasting using Structural Equation Modelling-based, context-sensitive Data Fusion.

The proposed Technology Forecasting framework includes methodologies and processes for integrating data from varied sources while emphasising the importance of complex hierarchal relational interconnections and context-related information to augment technology indicator relevance. Data Fusion methods attuned to context refine the output knowledge produced by accounting for the influence of external, context-related variables. Structural Equation Modelling is a robust statistical methodology that can discern and assess the complex hierarchical relationships between latent and observable variables within a given problem and its context, demonstrating efficacy in executing context-sensitive Data Fusion.

The study developed an autoregression model instantiation of the framework, tailored explicitly for longitudinal forecasting in the National Research and Education Network technology domain. This model instantiation, formulated through deductive reasoning, incorporates insights from action research within the South African National Research Network. It is supplemented by analysis of secondary data from the Trans-European Research and Education Network Association's Compendiums, which record infrastructure, services, and ecosystem-related trends for National Research and Education Networks in Europe.

This autoregressive model instantiation, although found suboptimal, innovatively delineated various technology-related indicators from the National Research and Education Network technology domain as distinct model technology-related constructs, for example, the measurements of core network traffic, while also integrating indicators for context-related constructs, such as the spectrum of institutions typically serviced by National Research and Education Networks. Employing secondary data from the Trans-European Research and Education Network Association's annual Compendiums, the study undertook a Partial Least Squares regression analysis to empirically evaluate this autoregressive model instantiation to ascertain key model parameters, such as indicator loadings and path coefficients. The study also engaged in an extensive reliability and validity analysis of this model instantiation, affirming the empirical analysis's repeatability and the model instantiation's internal consistency in providing technology forecasting outputs in the National Research and Education Network technology domain.

Next, the study developed a cross-sectional model instantiation for the National Research and Education Network technology domain. Although incapable of longitudinal technology forecasting, this model instantiation marked a considerable improvement in performance over the autoregressive model instantiation. Its development not only amalgamated knowledge from action research conducted within the South African National Research Network and insights from the annual Compendiums of the Trans-European Research and Education Networking Association but also hypotheses from scholarly literature. Partial Least Squares regression analysis, employing data from the Trans-European Research and Education Network Association Compendiums, confirmed various hypothesised relationships, except the anticipated positive correlation between a National Research and Education Network's infrastructure and advanced services capabilities. This exception underscores the influence of technology leapfrogging within the National Research and Education Network community, potentially disrupting established technology development and adoption patterns.

The study concluded with an extensive analysis of the proposed framework's strengths and weaknesses. Informed by a broad scholarly discourse on the strengths and weaknesses of Data Fusion and Structural Equation Modelling, this evaluation scrutinised the framework's capability to incorporate context-related information in forecasting computations and its potential susceptibility to inaccuracies arising from structural model misspecification. This investigation employed various model instances specific to the National Research and Education Network technology domain for this assessment, including a structurally disarranged model instantiation.

This study heralds a transformative advancement in Technology Forecasting, particularly within the National Research and Education Networks technology domain, by introducing a novel amalgamation of Structural Equation Modelling and context-sensitive Data Fusion to perform transversal and longitudinal technology prediction using technology and context-related indicators. This innovative approach contributes to Engineering and Technology Management by offering an advanced tool for strategic planning and technology trend analysis. The core publications from the study demonstrated the development, practical application and assessment of this integrated framework. In contrast, the study's supplementary publications enhanced the understanding and application of Partial Least Squares analysis tools to perform the regression analysis required to create various model instantiations for the National Research and Education Networks technology domain.

All research data for this study, including the associated National Research and Education Network Compendiums, Structural Equation Modelling path diagrams, processed indicator data, and generated core and supplementary publications, are organised and stored within dedicated private collections on Figshare.

Keywords: Context-Sensitive Data Fusion, National Research and Education Networks, Partial Least Squares Regression Analysis, Structural Equation Modelling, Technology Forecasting, Technology Indicators, Trans-European Research and Education Network Association

ACKNOWLEDGEMENTS

My heartfelt thanks go to the Council for Scientific and Industrial Research and the University of Pretoria for their financial support, which has been crucial in advancing this research.

I want to acknowledge the contribution of Dr Pieter de Villiers at the University of Pretoria, who assisted in conceptualising Data Fusion through Structural Equation Modelling. His expertise and guidance in implementing Structural Equation Modelling as a Data Fusion analysis tool have significantly impacted this study.

I am deeply grateful to my parents, Johannes and Marietjie Staphorst, whose unwavering support and encouragement have been the foundation of my journey. Their belief in me has been essential in keeping me going and achieving success.

I am indebted to my sister, Ronell Staphorst, and my aunt, Joey Rabie, whose endless encouragement has inspired and strengthened me. Their unwavering support has guided me towards excellence.

Furthermore, I am immensely thankful for my daughter, Tessa Staphorst. Your presence in my life brings joy and serves as a constant motivation. Tessa, you remind me daily of the importance of pursuing dreams with determination and passion. I sincerely appreciate your support throughout my thesis journey. Your unique perspective and caring contributions have played a role in its completion.

I want to thank Prof Leon Pretorius and Prof Marthinus Pretorius, whose invaluable guidance and expertise have been crucial in my research journey. They have provided insightful feedback and significantly shaped my work and understanding of the subject. I am deeply grateful for their mentorship and steadfast commitment to excellence, which have profoundly influenced my growth as a researcher.

I sincerely thank each of you for making this journey successful and fulfilling.

This dissertation is dedicated to:

My loving family, whose continuous support and encouragement carried me through the arduous times.

"The next best thing to a good education is a pushy mother."
(Source: M.P. Staphorst)

ETHICS AND ORIGINALITY DECLARATION

I, **Leonard Staphorst**, declare that the dissertation/thesis, **Technology Forecasting for National Research and Education Networks using Structural Equation Modelling-Based Context-Sensitive Data Fusion**, which has been submitted in partial fulfilment of the requirements for the degree of **Philosophiae Doctor (Technology Management)**, at the University of Pretoria, is my work and has not previously been submitted by me for any degree at the University of Pretoria or any other tertiary institution.

I declare that I obtained the applicable research ethics approval to conduct the research described in this dissertation/thesis.

I declare that I have observed the ethical standards required by the University of Pretoria's ethics code for researchers and have followed the policy guidelines for responsible research.

Signature: Staphorst

Date: 3 July 2024

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND ABBREVIATIONS

3G	Third Generation
5G	Fifth Generation
AAI	Authentication and Authorization Infrastructure
AARNET	Australian Academic and Research Network
AI	Artificial Intelligence
ARG	Alternate Reality Games
ASEM	American Society for Engineering Management
ASREN	Arab States Research and Education Networks
AVE	Average Variance Extracted
CAnet	Canadian Academic and Research Networks
CB-SEM	Covariance-Based Structural Equation Modelling
CLPM	Cross-Lagged Panel Model
CMM	Capability Maturity Model
CSNET	Computer Science Network
DEA	Data Envelopment Analysis
DANTE	Delivery of Advanced Network Technology to Europe
DF	Data Fusion
DFG	Data Fusion Group
DMU	Decision-Making Unit
DSEM	Dynamic Structural Equation Modelling
EMBOK	Engineering Management Body of Knowledge
EMJ	Engineering Management Journal
ESANET	East and Southern Africa Network
ESnet	Energy Sciences Network
FAIR	Findable, Accessible, Interoperable, and Reusable
FTA	Future-oriented Technology Analysis
HCI	Human-Computer Interaction
HEI	Higher Education Institution
ICT	Information and Communication Technology
IoT	Internet of Things
IPv4	Internet Protocol version 4

IPv6	Internet Protocol version 6
IRU	Indefeasible Right of Use
ISP	Internet Service Provider
IXPs	Internet Exchange Points
JANET	Joint Academic Network
JDL	Joint Directors of Laboratories
KanREN	Kansas Research and Education Network
KENET	Kenya Education Network
LISREL	Linear Structural Relations
LTE	Long-Term Evolution
ML	Machine Learning
MPLS	Multiprotocol Label Switching
NIPALS	Non-Iterative Partial Least Squares
NORDUnet	Nordic University and Research Network
NREN	National Research and Education Network
NSFNET	National Science Foundation Network
OSI	Open Systems Interconnection
PCA	Principal Component Analysis
PLS	Partial Least Squares
PLS-SEM	Partial Least Squares Structural Equation Modelling
R&D	Research and Development
RDA	Research Data Alliance
RedIRIS	Red de Interconexión de Recursos Informáticos
REN	Research and Education Network
RENU	Research and Education Network of Uganda
ROI	Return of Investment
RREN	Regional Research and Education Network
SANReN	South African National Research Network
SDN	Software Defined Networking
SEM	Structural Equation Modelling
SOA	State-of-the-Art
SURFnet	Stichting Universitaire Rekencentrum Groningen Foundation Network
SVD	Singular Value Decomposition
SWITCH	Swiss Academic and Research Network

TB	Terabytes
TCP/IP	Transmission Control Protocol/Internet Protocol
TCSI	Technology Commercialization Success Index
TENET	Tertiary Education and Research Network of South Africa
TERENA	Trans-European Research and Education Network Association
TF	Technology Forecasting
TFDEA	Technology Forecasting using Data Envelopment Analysis
TFSEMDF	Technology Forecasting using Structural Equation Modelling-Based Data Fusion
TI	Technology Intelligence
TIO	Technology Indicator Ontology
TRIZ	Teoriya Resheniya Izobretatelskikh Zadach
TVET	Technical and Vocational Education and Training
UA	UbuntuNet Alliance
UNINET	University Network
WACREN	West and Central African Research and Education Network
ZAMNET	Zambia Network

CHAPTER 1 – INTRODUCTION

1.1 INTRODUCTION TO THE RESEARCH PROBLEM

The trajectory of technological advancement is accelerating, adhering to established growth paradigms such as Moore's Law (Mack, 2011), Nielsen's Law (Nielsen, 2013), and Metcalfe's Law (Metcalfe, 1995). The introduction, refinement, and global distribution of innovative technologies, along with a movement towards international collaboration and open innovation, are catalysing this rapid development (Nyberg & Palmgren, 2011). Consequently, this evolution has given rise to global markets for technology-centric products and services (Porter, 2007), rendering it imperative for corporations to monitor and forecast technological trends to ensure their viability, profitability, and growth in an increasingly competitive landscape. Possessing this foresight enables firms to construct a resilient foundation capable of withstanding or adapting to the rapidly evolving market demands (Porter, 2007). Moreover, the adept management of internal and external shifts is paramount in sustaining a competitive edge for enterprises operating within dynamic and technologically intensive markets (Lichtenthaler, 2004).

The evolution from Third Generation (3G) and Long-Term Evolution (LTE) networks to Fifth Generation (5G) telecommunications infrastructure underscores the criticality of precise Technology Forecasting (TF). The burgeoning adoption of mobile devices and the escalating demand for mobile internet services have catalysed a swift progression in mobile technology paradigms. This advancement necessitates strategic foresight in planning and predicting the deployment of emergent technologies such as 5G (Kalem et al., 2020).

This shift accentuates the need for robust forecasting capabilities and the strategic determination of optimal launch windows for new services. For instance, a detailed analysis of the telecommunications sector revealed a pivotal timeline for implementing 5G services, identified through a confluence of forecast parameters that significantly impact network investment strategies (Kalem et al., 2020). This scenario highlights the pivotal role of technology prediction in navigating companies through the rapidly evolving technological milieu, ensuring their competitive stance and agility in responding to market dynamics and introducing novel technologies.

Technology Intelligence (TI) is pivotal within the ambit of technology management. It encompasses systematically collecting and analysing technology-related data, transforming it into actionable intelligence by discerning interconnections among various elements. This process elevates raw data to strategic knowledge, aiding decision-makers in strategic planning (Chang et al., 2008; Lichtenthaler, 2004).

Technology indicators, such as technology maturity and innovation levels, are vital to this process. They serve as critical data sources for the ongoing assessment and evaluation of technologies across their lifecycle (Chang et al., 2008). By employing a forward-looking methodology such as FTA, decision-makers can analyse the insights derived from these technology indicators. This analysis facilitates the acquisition of TF knowledge, enabling a deeper understanding of future technology trajectories and trends (Porter, 2005).

Haenlein and Kaplan (2004) postulate that regression analysis comprises a spectrum of statistical methodologies aimed at modelling and scrutinising the dynamics between dependent and independent variables through empirical data. These methods endeavour to explain the variance in dependent variables as functions of alterations in independent variables, thus facilitating the prediction and forecasting of outcomes based on specific independent variable inputs. Conventional regression approaches like multiple and logistic regression fall into the category of first-generation techniques. These methods generally presuppose the independence of multiple dependent variables, constraining their efficacy in modelling intricate interdependencies, such as those potentially present in output variables within a TF framework (Haenlein & Kaplan, 2004).

In contrast, Jöreskog (1973) introduced Covariance-Based Structural Equation Modelling (CB-SEM) as an advanced, second-generation technique to surmount these constraints. Structural Equation Modelling (SEM) enables concurrent modelling of intricate relationships among multiple dependent and independent constructs. A critical limitation of first-generation regression methods is their assumption of direct observability of all variables, thereby excluding unobservable or latent constructs from their scope (Haenlein & Kaplan, 2004). SEM, however, incorporates these latent constructs, enhancing the model's comprehensiveness. Steinberg (2009) and Steinberg and Rogova (2008) advocate for the suitability of SEM in context-sensitive Data Fusion (DF) applications. SEM is adept at supporting complex structural models essential for situation state estimation in TF. It also accommodates non-linear, non-Gaussian elements and cyclical dependencies among model variables, encompassing latent and directly observable factors.

Sohn and Moon (2003) underscore the imperative for TF methodologies to rigorously account for the structural interconnections between technology indicators and TF output metrics. In this context, SEM is a notably advantageous approach, adept at capturing and modelling the intricate hierarchical interrelations between these indicators and TF outputs. By offering a more robust framework, they illustrate that SEM transcends conventional factor and path analysis methods, such as Bayesian Networks (Steinberg, 2009). This enhanced capability of SEM is exemplified in its application to forecast TCSI, serving as a pivotal TF output metric. This approach underscores the efficacy of SEM in navigating the complex landscape of technology indicators and their predictive relationships with commercialisation success, offering a more comprehensive tool for TF endeavours.

Buchroithner (1998) and Wald (1997) define DF as an intricate framework that amalgamates disparate data streams to synthesise information of augmented quality, where the criterion for increased quality is contingent upon the specific application context. Originating within the military sphere for the synthesis of superior tactical intelligence via the processing of multi-faceted sensor data (Wald, 1999), DF in this sphere equates context with situational constructs, conceptualised as a nexus of relational dynamics or instantiated interrelations (Steinberg, 2009). This contextual element is pivotal at every juncture of the DF methodology, bolstering data congruence and interconnection and prognosticating situational states (Steinberg, 2009). Contemporary explorations in this domain have ventured into context-sensitive DF modalities, adeptly honing the intelligence extracted at each processing tier, informed by the influence of exogenous context-specific variables (Steinberg, 2009).

Expanding upon the foundational work of Staphorst et al. (2013, 2014), Dash and Paul (2021) have advanced the application of SEM in TF and management research. Staphorst et al. (2016a) initially advocated for integrating context-sensitive DF in the TI process, highlighting SEM's utility in elucidating the complex interconnections among various technology indicators, thus enhancing the calibre of technological insights for forecasting applications. Dash and Paul's (2021) research resonates with and extends these initial assertions, delving deeper into SEM's development from conceptual frameworks to empirical model construction using technology user data. They emphasise the effectiveness of Partial Least Squares Structural Equation

Modelling (PLS-SEM) over CB-SEM for estimating intricate cause-and-effect relationships, aligning with Staphorst et al.'s emphasis on SEM's predictive capabilities in assessing emerging technologies. Moreover, Dash and Paul (2021) suggest future research directions that build on Staphorst et al.'s groundwork, proposing the exploration of both composite-based and factor-based models in SEM, the investigation of complex consumer behaviour relationships, and the integration of advanced model structures with moderation effects, thereby reinforcing the significance of SEM in TF and its applicability in understanding consumer behaviour and technology adoption trends.

A National Research and Education Network (NREN) equips research and education communities within a country with broadband network connections and services specifically designed for their needs. GÉANT (2022) notes that NRENs often extend their services to public sector organisations, including hospitals, municipalities, and libraries. Typically, each country operates a single NREN, such as the South African National Research Network (SANReN) (SANReN, n.d.) and the Joint Academic Network (JANET) in the United Kingdom (Cooper et al., 1991). However, in some countries, such as the United States, multiple entities, like the Energy Sciences Network (ESnet) and Internet2, serve different research and education sectors or geographical areas. NRENs predominantly utilise fibre optic cables to offer high-speed connectivity and advanced services to researchers, educators, and students at lower costs than standard network providers. These networks are transforming due to technological advancements, leading to the emergence of new business models, innovative infrastructure solutions, and service offerings. Increased collaboration among NRENs also characterises this evolution (GÉANT, 2022).

This study centred on applying SEM to implement context-sensitive DF to augment the precision and effectiveness of TF. This endeavour culminated in developing, applying and assessing a framework for Technology Forecasting using Structural Equation Modelling-Based Data Fusion (TFSEMDF) within the NREN technology domain. The study synthesises an array of theoretical postulations, analyses, and empirical findings derived from an extensive series of peer-reviewed scholarly contributions by Staphorst et al. (2013, 2014, 2016a, 2016b).

Staphorst et al. (2013) postulated that TF can be improved by integrating data from various technological and contextual sources through context-sensitive DF techniques, leading to more reliable and insightful forecasting results. This study highlighted the essential role of SEM in analysing the relationships among variables relevant to specific technological and context-related factors, culminating in the TFSEMDF framework’s introduction. TFSEMDF is an SEM-based DF framework that utilises technology and context-related indicators to perform robust TF. Staphorst et al. (2013) concluded by proposing a comprehensive methodology focussed on creating an autoregressive model instantiation of TFSEMDF for assessing the framework's longitudinal forecasting effectiveness within the NREN technology domain.

Staphorst et al. (2014) showcased the practical application of the proposed SEM-based DF framework in TF for the NREN technology domain. This effort involved the creation of the autoregressive NREN model instantiation, presented by Staphorst et al. (2013), informed by action research within the SANReN (Gustavsen, 2008; SANReN, n.d.) and data from the Trans-European Research and Education Network Association’s (TERENA's) NREN Compendiums (TERENA, 20211, 2012). Although shown to be non-optimal, this autoregressive NREN model instantiation could track and predict longitudinal technology and contextual trends in the NREN ecosystem, such as the reach of the NREN.

Furthering this research, Staphorst et al. (2016a) refined the TFSEMDF framework while applying it to a cross-sectional NREN model instantiation (Staphorst et al., 2014), building upon earlier work of Staphorst et al. (2013). This cross-sectional NREN model instantiation integrated findings from action research in SANReN (Gustavsen, 2008; SANReN, n.d.), the 2011 TERENA NREN Compendium (TERENA, 2011), and theoretical propositions from academic research.

This study culminated in a detailed examination of the TFSEMDF framework's strengths and weaknesses, expanding on the initial efforts of Staphorst et al. (2016b). This pivotal work identified key strengths, notably the framework's adeptness in incorporating contextual information for enhanced forecasting accuracy. It also acknowledged various weaknesses, such as the framework's susceptibility to inaccuracies due to errors in structural model specification. The original analysis conducted by Staphorst et al. (2016b) encompassed various model instantiations from Staphorst et al. (2014, 2016a) and introduced a structurally disarranged model instantiation for the NREN technology domain.

1.2 BACKGROUND

1.2.1 STRUCTURAL EQUATION MODELLING

Regression analysis encompasses a range of statistical techniques specifically designed for modelling and analysing relationships between dependent and independent variables in empirical data (Haenlein & Kaplan, 2004). These techniques primarily reveal how variations in dependent variables function as outcomes of changes in independent variables. Such analytical power enables researchers to predict and forecast the values of dependent variables using the known values of independent variables.

First-generation regression techniques, including multiple regression, discriminant analysis, logistic regression, and analysis of variance, make a critical assumption of independence among dependent variables. This assumption significantly constrains their capacity to fully model complex interrelationships, notably in TF models where multiple output variables may interact (Haenlein & Kaplan, 2004). These traditional methods require revision to determine possible mediating or moderating effects among output variables. Addressing this limitation, Jöreskog (1973) introduced SEM as a more advanced, second-generation technique. SEM facilitates the simultaneous modelling of relationships among multiple dependent and independent constructs, including those not directly observable.

One fundamental limitation of first-generation regression methods is their reliance on the direct observability of all variables if real-world sampling experiments can readily capture all variable values (Haenlein & Kaplan, 2004). This approach excludes unobservable variables, often crucial latent constructs, from analysis. SEM, however, accommodates these latent constructs. Steinberg (2009) advocated for applying SEM in context-sensitive DF, highlighting its suitability for complex structural models in situation state estimation, as required in TF. SEM's capability accommodates non-linear and non-Gaussian factors and cyclical dependencies among model variables, whether latent or observable (Steinberg, 2009; Steinberg & Rogova, 2008).

Transversal (cross-sectional) SEM and longitudinal SEM are two distinct approaches to statistical analysis. While cross-sectional SEM examines relationships between variables at a single instance of time, longitudinal SEM, including Cross-Lagged Panel Model (CLPM) SEM (Hamaker et al., 2015) and autoregression techniques (Burant, 2022), focuses on analysing dynamic interrelationships and causal inferences over time. CLPM SEM is especially effective in longitudinal studies for identifying directional influences between variables across different time points, addressing temporal precedence, which is critical for causal inference (Hamaker et al., 2015). This method is popular in psychology and social sciences for examining the interplay of behaviour and attitude (Selig & Preacher, 2009).

Autoregression in SEM, on the other hand, concentrates on the self-influence of a variable over time, distinguishing short-term fluctuations from long-term trends, and is integral in time-series analysis and forecasting models (Bollen & Curran, 2004). Integrating autoregressive elements in SEM enhances the model's robustness, making it widely applicable in economics, finance, and behavioural sciences. The versatility of SEM in incorporating both cross-lagged and auto-regressive components underlines its effectiveness in exploring temporal dynamics across various research fields (Box et al., 2015).

Sohn and Moon (2003), Staphorst et al. (2013, 2014) and Dash and Paul (2021) collectively demonstrate the robust application of SEM in the field of TF. Sohn and Moon (2003) initially identified the limitations of traditional TF techniques in considering structural relationships among technology indicators and TF output metrics, advocating SEM as a superior alternative for modelling complex hierarchical relationships, as evident in their work on the TCSI. This approach was expanded by Staphorst et al. (2013, 2014), who posited that integrating context-sensitive DF within the TI process through SEM could significantly enhance the calibre of technological insights by refining interconnections among technology indicators tailored explicitly for TF applications.

Dash and Paul (2021) further corroborated this perspective, emphasising SEM's versatility in diverse social science disciplines, especially in technology management and predicting adoption trajectories. Their work extends beyond conceptual frameworks to empirical model construction using technology user data, highlighting the efficiency of PLS-SEM over CB-SEM in forecasting (Cepeda-Carrion, 2019). They propose future research directions that build upon the foundational work of Staphorst et al. (2016a), such as testing both composite-based and factor-based models in SEM, exploring complex consumer behaviour relationships, and integrating advanced model structures, thereby reinforcing SEM's significance in TF and its applicability in understanding consumer behaviour and technology adoption trends.

1.2.2 Data Fusion

Initially developed for military purposes in processing sensor data, DF has evolved into a multidisciplinary field, extending to environmental monitoring, healthcare, robotics, and financial systems (Hall & Llinas, 1997; Wald, 1997, 1999). This evolution reflects DF's capacity to integrate and synthesise data from various sources, producing superior-quality information tailored to specific applications (Khaleghi et al., 2013; Waltz & Llinas, 1990). The Joint Directors of Laboratories (JDL) DF model, a seminal framework in this field, segment the DF process into distinct levels, including data pre-processing, object refinement, and situation assessment, providing a structured approach to understanding and implementing DF (Steinberg & Bowman, 2017).

Context, synonymous with a situation, is crucial to the discipline of DF, defined as a set of relational connections. Context plays a pivotal role at each level of the DF process, including refining data alignment and association and enhancing situation state estimation (Steinberg, 2009). Recent advancements in context-sensitive DF techniques have emphasised refining knowledge generation by considering characteristics of exogenous context-related variables. This evolving focus highlights the integral role of context in effectively interpreting fused data (Steinberg, 2009).

Recent developments in the field of DF have seen significant advancements, particularly in integrating deep learning techniques. Li et al. (2022) provide a comprehensive review, emphasising the substantial progress made in traditional algorithms and the performance enhancements achieved through deep learning. These advancements are particularly evident in the fusion of different data modalities like spatio-spectral and spatio-temporal data. One prominent example is the application of deep learning in multimodal remote sensing DF to address the challenges of handling heterogeneous Earth observation data. Another is in the biomedical sector, where Stahlschmidt et al. (2022) propose using multimodal deep learning for DF in biomedical applications. Their work highlights the application of deep learning in combining various data types for enhanced analysis, particularly in complex biological systems.

However, the DF field also faces evolving challenges, particularly concerning data privacy, security, and the ethical implications of data usage. As technology advances, there is a growing need for robust frameworks for data governance and standards to ensure the responsible and ethical use of DF techniques (Koch, 2021). Future research in DF will likely focus on finding more efficient ways to handle the complexity and scale of DF applications across various domains.

1.2.3 Technology Forecasting

The pace of technological innovation, coupled with a shift towards globally inclusive and open innovation models, has significantly accelerated technological progress (Nyberg & Palmgren, 2011). This acceleration has, in turn, intensified the competition in global markets that focus on technology-driven products and services (Porter, 2007). For businesses, monitoring current trends and forecasting technological developments is crucial for survival and economic prosperity, providing a foundation for resilient infrastructure that meets the market's shifting demands (Porter, 2007). Furthermore, managing internal and external changes is essential for maintaining a sustainable competitive advantage in markets characterised by rapid technological evolution (Lichtenthaler, 2004).

TI is an indispensable component of technology management. It involves the meticulous aggregation of technology-related data, its conversion into insightful information, and the subsequent refinement into strategic intelligence that underpins decision-making (Chang et al., 2008; Lichtenthaler, 2004). This data includes technology indicators, such as technology's maturity and innovation levels, which are essential for the detailed assessment and evaluation of technologies throughout their lifecycle (Chang et al., 2008). By employing a methodology known as FTA, TI provides decision-makers with critical foresight. This methodological approach is central to the domain of TF and enhances the strategic management of technology by enabling well-informed decisions based on anticipatory insights (Porter, 2005).

TF is an intricate process that entails collecting and analysing information to discern and anticipate technological evolutions and the various contextual elements shaping them, such as national policy frameworks. Cho (2013) classifies TF methodologies into exploratory and normative techniques. Exploratory techniques, including Technology Forecasting using Data Envelopment Analysis (TFDEA), S-curve analysis, and trend extrapolation, function on the premise that technological advancement adheres to a discernible evolutionary trajectory. These methods leverage data analysis to forecast technological trends and patterns. Conversely, normative techniques, exemplified by methods like Delphi analysis and relevance trees, commence with a predefined technical objective. They then chart the steps or pathways to achieve this envisioned technological state (Cho, 2013).

Recent TF research advancements have yielded significant insights, particularly in SEM methodologies. Dash and Paul (2021) conducted an in-depth comparative analysis of two prevalent SEM techniques in TF: CB-SEM and PLS-SEM. Their investigation uncovered that PLS-SEM typically exhibits higher item loadings than CB-SEM. Furthermore, they found that employing Partial Least Squares (PLS) regression results in structural relationships that closely mirror those observed in CB-SEM. This research plays a pivotal role in underscoring the implications of selecting specific SEM methods on the precision and dependability of technology forecasts (Dash & Paul, 2021).

Kalem et al. (2020) conducted a study to forecast technological advancements in the mobile telecommunication sector, focusing on the advent of 5G technology. They aimed to predict the timeline for a prominent Turkish telecommunication company's deployment of 5G services. Their analysis, based on multiple forecast parameters influencing network investment decisions, led them to project August 2020 as a critical juncture. This research is pivotal as it exemplifies the practical application of TF in a dynamic industry. It provides valuable insights for strategic decision-making regarding technology adoption and investment strategies (Kalem et al., 2020).

1.2.4 National Research and Education Networks

NRENs have a rich historical lineage dating back to the advent of networked computing and the Internet era. Their origins trace back to the 1980s, a period characterised by a surging need for tailor-made, high-speed network infrastructures capable of accommodating the specialised demands of academic and research communities. The National Science Foundation Network (NSFNET) in the United States played a prominent role in shaping the modern Internet landscape (Leiner et al., 2009).

Inaugurated in 1985, NSFNET represented a pivotal development in networking by serving as a robust backbone that interlinked networks and university systems throughout the United States, facilitating rapid data exchange and fostering productive collaborations among researchers (Leiner et al., 2009). This dedicated NREN archetype gained swift recognition for its role in nurturing academic cooperation and elevating research capacities, prompting several nations to emulate it. As an illustrative case, the United Kingdom established its counterpart, JANET, in 1984. It has since become integral to the nation's digital research infrastructure (Cooper et al., 1991).

In the contemporary landscape, NRENs operate as purveyors of broadband network connectivity and services customised to cater to the unique requirements of research and educational communities within the confines of individual nations (GÉANT, 2022). These networks primarily hinge on the utilisation of fibre optic infrastructure, delivering advanced services at a cost-effective advantage compared to their commercial network counterparts. NRENs offer an array of services encompassing high-speed internet access bandwidth conducive to data research, provision of VPN services, facilitation of cloud storage and computing solutions, deployment of video conferencing facilities, provision of access to NRENs and Regional Research and Education Networks (RRENs), and furnishing identity and access management services.

From a technological vantage point, NRENs harness cutting-edge networking technologies to optimise network performance, enhance flexibility, and fortify resilience (GÉANT, 2022). Internet Protocol version 6 (IPv6) is pivotal in providing the requisite address space to accommodate the many devices and applications in academia and research, particularly for Internet of Things (IoT) applications (Aldowah et al., 2017; Al-Emran et al., 2020). Multiprotocol Label Switching (MPLS) further augments NREN efficiency by streamlining data traffic flow, a pivotal attribute for bandwidth-intensive and low-latency applications, as is frequently the case with research-related applications (Kompella et al., 2017; Rosen et al., 2001).

The advent of Software Defined Networking (SDN) constituted a paradigm shift in network management at NRENs by segregating control functions from forwarding functions, thereby enabling swift adaptation to the evolving difficulties of research and education environments (Bera et al., 2017; Priyadarsini & Bera, 2021). Additionally, NRENs frequently establish host Internet Exchange Points (IXPs), foster seamless data exchange, curtail latency, and bolster network resilience—a trifecta of imperatives underpinning the delivery of services within these highly specialised networks). The integration of these cutting-edge technologies underscores NRENs' commitment to catering to their user community's dynamic needs, propelling innovation, and expediting progress.

In nations where these capabilities have emerged and evolved, sometimes spread among several legal entities, NRENs typically assume the strategic role of serving as the foundational networking infrastructure and Internet Service Provider (ISP) for research and education connectivity. Notable examples of these entities include SANReN and the Tertiary Education and Research Network of South Africa (TENET) (SANREN, n.d.), JANET in the United Kingdom (Cooper et al., 1991), the Stichting Universitaire Rekencentrum Groningen Foundation Network (SURFnet) in the Netherlands, and Red de Interconexión de Recursos Informáticos (RedIRIS) in Spain.

Furthermore, complementing NRENs, RRENs are vital in fostering cross-border collaboration and connectivity. A noteworthy example of such collaborative efforts is GÉANT, which unites European NRENs in a harmonious network across the continent. The UbuntuNet Alliance (UA) takes centre stage in Africa, interconnecting NRENs in Eastern and Southern Africa to bolster education and research collaborations. In the United States, the landscape features the coexistence of multiple NRENs, including ESnet and Internet2, and state-level Research and Education Networks (RENs), such as the Kansas Research and Education Network (KanREN). It is worth noting that Internet2 also assumes the role of an RREN, championing collaboration and resource-sharing among research institutions throughout the United States (GÉANT, 2022).

Recent investigations within the African NREN ecosystem have concentrated on delineating the evolving roles undertaken by these entities. For instance, in conjunction with the West and Central African Research and Education Network (WACREN), researchers conducted a comprehensive study in 2019 to outline a strategic roadmap for NRENs in West and Central Africa (Kashefi et al., 2019). This endeavour involved systematically administering surveys dispersed across WACREN's expansive coverage region, enabling the research team to discern and articulate the region's distinct service requisites (Kashefi et al., 2019). The identified needs encompassed multi-faceted facets such as access to conferences, academic literature, research infrastructure (including libraries), video conferencing capabilities, collaborative software tools, and remote computing resources (Kashefi et al., 2019). Furthermore, the study illuminated the formidable challenges confronting researchers, specifically emphasising network connectivity issues that profoundly impact the progress of their research undertakings (Kashefi et al., 2019). Collectively, this research endeavours to offer a profound understanding of the exigencies and impediments faced by the academic and research community in West and Central Africa, thus serving as a foundational cornerstone for the development of bespoke NRENs tailored to address these unique requirements (Kashefi et al., 2019).

The arena of NRENs is currently undergoing rapid transformation, fuelled by technological advancements. This paradigm shift has ushered in innovative business models, cutting-edge infrastructure solutions, expanded service portfolios, and an enhanced landscape of international collaborations (GÉANT, 2022). NRENs inherently operate within an Information and Communication Technology (ICT) standards-driven environment. NRENs need to adeptly manoeuvre through a multi-faceted terrain that includes considerations such as bandwidth utilisation and contextual factors like regulatory mandates and government fiscal policies. Given the intricate interplay between technology and context-related dimensions within the domain of NRENs, it emerges as an ideal arena for a comprehensive examination of the capabilities and limitations of the TFSEMDF framework (Staphorst et al., 2013, 2014, 2016a, 2016b).

1.3 RESEARCH AIMS AND SCOPE

This study aimed to develop an enhanced TF framework, advancing beyond the typical constraints of conventional TF methods, which often overlook contextual variables in their forecasting models. In this endeavour, the study integrated PLS-SEM as a statistical approach to implement context-sensitive DF. This integration was pivotal in enabling the framework to effectively encompass a wide array of variables, both observed and unobserved (latent), while accounting for measurement errors and their interdependencies. The unique contribution of context-sensitive DF in this framework is its ability to enrich the forecasted outcomes with contextually relevant insights, thereby enhancing the accuracy and comprehensiveness of the TF process.

The study aimed to apply this innovative TFSEMDF framework within NRENs' multi-faceted and evolving sphere. This application included the development of both longitudinal and transversal (cross-sectional) model instantiations. The design of the longitudinal model instantiation aimed to forecast technological parameters over time, thereby capturing their dynamic evolution. In contrast, the transversal model instantiation focused on predicting the interrelations between a complex set of technology and contextual-related parameters at a single point in time, offering a snapshot of these intricate interactions.

The final aim of the study entailed an analysis of the strengths and weaknesses of the developed TFSEMDF framework. This evaluation employed various model instantiations specifically designed for NRENs' technology domain. The assessment aimed to rigorously examine the framework's efficacy and limitations in accurately capturing and forecasting this domain's complex interplays of technological and contextual variables. Such a comprehensive evaluation was pivotal in discerning the framework's applicability and robustness in addressing the intricate dynamics of NREN TF.

The study was limited in scope to building its model instantiations for the NREN technology domain using action research within SANReN (Gustavsen, 2008; SANReN, n.d.) and technology and context-related data sourced from TERENA (2011, 2012). Moreover, it incorporated limited academically validated relationships between model variables while creating model instantiations. An exploratory phase to directly glean insights from NRENs about the interplay between technology and contextual elements within this ecosystem was outside the scope of this study. Additionally, the investigation into the longitudinal model instantiation narrowed its scope to solely consider autoregressive modelling approaches, potentially omitting more complex or nuanced temporal dynamics that could become apparent using CLPM SEM.

1.4 Research Motivation

This study delved into developing, applying, and assessing the TFSEMDF framework within the NREN technology domain. A dual set of motivations, spanning business and academic interests, underpinned the study. Sons will explore these motivations.

1.4.1 Business Motivation

An NREN represents a pivotal national resource integral to the foundational framework of a country's research, education, and innovation sectors. Therefore, such institutional entities, which typically are bolstered by public investment through fiscal allocations, play a critical role in catalysing the advancement of national academic and research capabilities. In this context, possessing strategic foresight to guide NREN's technological trajectory and investment becomes imperative. Hence, foresight is instrumental in ensuring that an NREN can continually meet the escalating demands of its constituent communities, thereby facilitating the country's sustained competitiveness in the increasingly dynamic global knowledge economy.

Moreover, the strategic stewardship of an NREN extends beyond mere technological advancement. It encompasses the reasonable allocation and utilisation of public funds. Effective governance and foresight in managing these resources are paramount to balancing technological innovation and fiscal responsibility. This dual focus propels the NREN to evolve in alignment with global technological trends. It ensures that public investments yield tangible benefits, reinforcing the network's role as a catalyst for national development in research and education. In this vein, the NREN emerges as a critical nexus, harmonising technological progression with fiscal prudence to nurture an environment favourable to intellectual growth and innovation, ultimately enhancing the country's stature within the global academic and research community.

1.4.2 Academic Motivation

The current academic landscape reveals a limited but growing body of research that intersects SEM with DF and TF. Isolated inquiries into the application of SEM for DF, as notably undertaken by Steinberg (2009), alongside the utilisation of SEM in the domain of TF, as demonstrated in the studies by Sohn and Moon (2003) and Dash and Paul (2021), underscore the emergent nature of scholarly exploration in these fields. Despite these individual contributions, the integration of SEM-based DF techniques within the context of TF, a pivotal aspect of the TFSEMDF framework introduced in this study, still needs to be explored in academic discourse. This gap signifies a frontier in research, presenting opportunities for novel investigations that could significantly contribute to advancing methodologies in TF and DF.

This intersection represents a significant gap in the current academic fabric, where the interlacing of SEM's analytical capabilities with DF and TF presents a unique research opportunity. Given SEM's inherent complementary nature and mutual reinforcement potential in DF and TF, this unexplored nexus offers substantial academic intrigue. Delving into this intersection through the TFSEMDF framework promises to advance theoretical understanding and unveil new avenues for practical application. The exploration of this confluence is thus motivated by the potential to contribute a novel perspective to the existing body of knowledge, bridging critical gaps and fostering a deeper, more integrated understanding of these interconnected fields.

Furthermore, despite NRENs constituting a fundamental component of national research infrastructure and spanning several decades, academic inquiry into the technologies utilised, strategies implemented, and factors influencing NRENs still needs to be explored. The strategic significance of these entities within national research ecosystems, alongside their leading position in telecommunications technology development, underscores the need for rigorous research in this domain. NRENs often pioneer advanced technologies that subsequently permeate into commercial networks. Thus, comprehensive studies of NRENs are imperative, not merely to appreciate their current role but also to understand their potential as harbingers of future telecommunications advancements, thereby offering valuable insights into the evolution of academic and commercial networking landscapes.

1.5 Novel Contributions Emanating from the Study

1.5.1 Contributions to Academic Theory

This study made substantial theoretical contributions to several academic domains. It broadened the use of SEM in TF beyond the confines of forecasting the TCSI, as initially posited by Sohn and Moon (2003), to forecast a range of technology indicators within the complex NREN technology domain. Additionally, it showcased the novel application of context-sensitive DF in the field of TF, thereby evolving this methodology from its conventional usage in sensor fusion to a broader, more versatile capability that includes TF.

This research not only corroborated Steinberg and Rogova's (2008) assertion that SEM is a potent tool for enacting context-sensitive DF, but it also expanded its utility beyond its established domain in natural language processing, as proposed by Steinberg (2009). It ventured into the field of TF by formulating the TFSEMDF framework, as documented in the works of Staphorst et al. (2013, 2014, 2016a). The development of generic frameworks for SEM, context-sensitive DF, and the classification of technology indicators within TF preceded the TFSEMDF conceptual framework development.

Further, the study achieved a significant milestone by conducting longitudinal and transversal TF analyses for NRENs through the integration of technology and context-related data by developing autoregressive and cross-sectional NREN model instantiations of the TFSEMDF framework, as presented in Staphorst et al. (2014, 2016a). This comprehensive approach signified a leap in applying SEM in TF, demonstrating its versatility and effectiveness in handling diverse data types and forecasting scenarios.

The research underscored the robustness of PLS regression as an effective analytical tool for evaluating TFSEMDF model instantiations. Furthermore, it established that the extensive reliability and validity metrics initially developed for PLS-SEM (Staphorst, 2010; Vinzi et al., 2010) are applicable and productive when applied to the TFSEMDF framework. The study undertook a comprehensive theoretical and empirical investigation of the TFSEMDF framework's strengths and weaknesses (Staphorst et al., 2016b). This exploration revisited and analysed academically recognised strengths and limitations inherent in the constituent methodologies of the TFSEMDF framework, namely SEM, DF, and TF. This multi-faceted analysis provided critical insights into the framework's practical utility and identified key areas for further development and refinement.

1.5.2 Contributions to Engineering and Technology Management

The American Society for Engineering Management (ASEM) articulates that engineering management is fundamentally about strategically coordinating, planning, organising, and directing resources to oversee technological or systems-based activities (ASEM, n.d.-a). This discipline requires precise management functions such as strategic planning, organisational structuring, and resource allocation (Shah & Nowocin, 2015). The Engineering Management Body of Knowledge (EMBOK) highlights the need for professionals in this field to possess diverse skills developed through a framework established by ASEM alongside other experts (ASEM, n.d.-b; Shah & Nowocin, 2015). These skills are critical across several areas, including strategic management, technology management, and enhancing Research and Development (R&D) and innovation (Shah & Nowocin, 2015).

By furnishing a structured framework for predicting technology trends, the TFSEMDF framework, combined with the insights gleaned from this investigation, equips engineering and technology managers with potent tools for strategic decision-making. Managers must understand to comprehend and predict technology trends, as this proficiency is indispensable for effective future planning and navigating the dynamic landscape of technological advancements. Moreover, this knowledge directly influences the management of technology portfolios and the orchestration of R&D activities, ensuring that engineering and technology management professionals are impeccably poised to spearhead their organisations towards innovation and sustained growth within an ever-evolving milieu.

Additionally, the TFSEMDF framework resonates within NRENs, entities that propel research and educational progress through technology (Johnson et al., 2011; National Science Foundation, 2020; Nokia, n.d.). NRENs occupy the vanguard of innovation and collaboration within the research domain. They necessitate an engineering and technology management approach that closely aligns with the cadence of technological changes and effectively caters to the distinctive requisites of research and education communities. Applying the TFSEMDF framework enables NRENs to maintain a leading edge in technology by offering a strategic method for anticipating and adjusting to technological trends, which is crucial for NRENs as their continued success and sustainability rely on continuously incorporating new technologies to improve research and teaching methods (Johnson et al., 2011; National Science Foundation, 2020; Nokia, n.d.).

This study's findings are essential for NRENs' strategic planning and technology management. By profoundly understanding future technology trends, NRENs can make intelligent decisions about investing in infrastructure, choosing strategic partners, and deciding what services to offer. This forward-thinking approach is crucial for maintaining NRENs' leadership role in providing resources to the research and education communities (Chopra et al., 2019; Das, 2019).

1.5.3 Expanding the Forecasting Practitioner's Toolbox

The TFSEMDF framework developed through this research stands out for its generic design, which draws on foundational principles of SEM, context-sensitive DF, and TF. This design choice ensures the framework does not bind itself to any specific technology domain or operational entity, making it exceptionally versatile. For engineering and technology management practitioners, this translates to an adaptable tool suitable for creating model instantiations across diverse technology sectors. The TFSEMDF framework provides a robust baseline for constructing tailored forecasting model instantiations, particularly in domains where technology-related and context-related factors heavily influence technological evolution. This adaptability is crucial for practitioners who need to align TF activities with their operational fields' specific conditions and dynamic requirements, thereby enhancing the strategic decision-making process within their organisations.

In the practical application of this framework, the model instantiations for NRENs, tested using data from the TERENA Compendiums of 2011 and 2012, primarily reflect European NRENs' technological and operational contexts. While predominantly European, these Compendiums also include inputs from a handful of non-European NRENs, providing a somewhat broader perspective. Consequently, while the study has primarily calibrated the NREN model instantiations for European contexts, these models hold potential applicability for NRENs outside Europe with careful consideration and possible adjustments. This broader applicability is crucial for practitioners operating in global or non-European contexts. Practitioners can adapt the model instantiations by integrating local data that reflects their unique technological, economic, and regulatory environments. Such recalibration is essential to maintain the relevance and effectiveness of the model instantiations, ensuring that the forecasting outcomes remain robust and reflective of the specific conditions in different global regions.

Furthermore, the longitudinal TF capabilities of the autoregressive NREN model instantiation highlight an essential area for enhancement. This study's instantiation, while a pioneering effort, demonstrated limitations in its predictive efficacy, suggesting a need for improved model instantiation designs using advanced SEM techniques such as CLPM. This limitation implies an opportunity for practitioners to develop more sophisticated and accurate forecasting models that better accommodate the temporal dynamics and complex interdependencies characteristic of the NREN technology domain. Enhancing the model with robust longitudinal data handling capabilities would significantly improve the precision of forecasts, making the tool even more valuable for strategic planning and operational management in rapidly evolving technology landscapes. This advancement not only caters to the immediate forecasting needs but also sets a precedent for ongoing improvements and innovations in the field of TF, reinforcing the TFSEMDF framework's applicability and utility across varied technological domains.

1.5.4 Core Publications Originating from the Study

This study produced four core publications, each crucial in directly illuminating and addressing the research aims and objectives. These scholarly contributions added substantial depth and scope to the study and introduced novel viewpoints and problem-solving approaches that apply to this study. These publications spanned various subjects and employed varied methodologies, enhancing the comprehension of the subject matter. These influential works are as follows:

1. Staphorst et al. (2013) made a pioneering attempt to integrate DF and SEM for application in TF in the paper entitled "Structural Equation Modelling based Data Fusion for Technology Forecasting: A Generic Framework", presented at PICMET2013 in San Jose, United States. This paper introduced an initial version of the TFSEMDF framework, which is central to this study. Furthermore, it proposed a research methodology based on an autoregressive model instantiation to evaluate this framework's effectiveness in longitudinal forecasting, particularly in the context of NRENs.
2. In the paper titled "Structural Equation Modelling Based Data Fusion for Technology Forecasting: A National Research and Education Network Example," presented at PICMET2014 in Kanazawa, Japan, Staphorst et al. (2014) delved into the utilisation of the TFSEMDF framework within the context of the NREN technology sphere. This investigation adopted the autoregressive NREN model instantiation initially proposed by Staphorst et al. (2013), incorporating deductive reasoning derived from the action research in SANReN (Gustavsen, 2008; SANReN, n.d.) and integrating secondary data sourced from TERENA's NREN Compendiums (TERENA, 2011, 2012). The study harnessed many technology-related metrics as indicators within the model and applied PLS regression analysis to TERENA's dataset, confirming the model's indicator loadings and path coefficients. Additionally, it subjected the model's reliability and validity to rigorous scrutiny, employing the methodologies explained in Appendix B.
3. The article by Staphorst et al. (2016a), entitled "Technology Forecasting in the National Research and Education Network Technology Domain using Context Sensitive Data Fusion," published in the Journal for Technology Forecasting and Social Change, revisited and refined the generic TFSEMDF framework. This refined framework was applied to an updated cross-sectional NREN model instantiation, significantly improving upon the initial model instantiation considered in Staphorst et al. (2013) and Staphorst et al. (2014).
4. At PICMET2016 in Honolulu, Hawaii, Staphorst et al. (2016b) presented the paper "Technology Forecasting Using Structural Equation Modeling Based Data Fusion: Analysis of Strengths and Weaknesses Using a National Research and Education Network Example." This paper studied the inherent strengths and weaknesses of the TFSEMDF framework. Key strengths it uncovered included the framework's ability to integrate contextual information into forecasting, while a notable weakness is its high sensitivity to errors in structural model specification. The paper employed various model instantiations for the NREN technology domain, including a structurally

disarranged model instantiation (Staphorst et al., 2016b), using NREN-related secondary data from TENERA (2011, 2012).

1.5.5 Supplementary Publications Augmenting the Study

In addition to the primary publications directly addressing the central research problem, this study produced two supplementary papers. While these papers did not directly address the study's main research aims and objectives, they were crucial in advancing and refining the PLS-based SEM tools used in this study. While their focus was peripheral to the primary research objective, these ancillary works offered valuable insights and methodological enhancements bolstered the study. These supplementary papers are:

1. The paper authored by L. Staphorst et al. (2015), titled "Impact of Intellectual Property Rights from Publicly Financed Research and Development on Research Alliance Governance Mode Decisions," and presented at IAMOT2015 in Cape Town, South Africa, focused on a study to create a decision-making model using PLS-SEM for strategists within publicly financed R&D organisations. This model aimed to enable them to analyse and predict governance mode decisions for research alliances. A fundamental aspect of developing this framework involved acquiring proficiency in constructing intricate SEM path diagrams and utilising the SmartPLS tool (Ringle et al., 2022) to analyse such diagrams through PLS regression, as elaborated in Appendix A.
2. The article entitled "Impact of Intellectual Property Rights on the Governance Mode Decisions of Engineering Managers during the Establishment of Research Alliances with Publicly Funded Entities", published in the ASEM's Engineering Management Journal (EMJ) (Staphorst et al., 2017), expanded on Staphorst et al. (2015), by including a reliability and validity analysis of the PLS-SEM based governance decision-making model utilising the various metrics detailed in Appendix B of this study. These metrics were fundamental in evaluating the reliability and validity of the different NREN model instantiations of the TFSEMDF framework considered in this study.

1.6 Organisation of the Thesis

This thesis comprehensively explores the development, application, and assessment of the TFSEMDF framework, focusing on its implications for NRENs. It is structured as follows:

- **Chapter 1 - Introduction:** The opening chapter sets the stage for this research by introducing the central research problem and offering a comprehensive background on the current landscape of TF, the pivotal role of NRENs, the complexities inherent in SEM, and the foundational concepts of DF. It articulates the aims, scope, and driving motivations underpinning this research, laying a clear foundation for the study's trajectory. Additionally, the chapter delineates the novel contributions this research makes to TF and engineering and technology management, highlighting the unique insights and advancements it brings. The chapter culminates with a systematic overview of the thesis structure, providing a roadmap for the reader to navigate the subsequent chapters and understand the cohesive structure of the research presented.
- **Chapter 2 - Literature and Theory Review:** Literature and Theory Review: This chapter explores the theoretical foundations of TF, tracing its historical evolution, key indicators, and specific forecasting techniques utilising SEM. It also thoroughly examines SEM, covering its development, core principles, and applications, and delves into DF frameworks, highlighting their context-sensitive applications. The chapter then offers a comprehensive overview of NRENs, discussing their history and the extensive data available from TERENA and GÉANT Compendiums. An analysis and synthesis section further contextualises SEM, DF, and TF integration within NRENs, demonstrating how these methodologies enhance strategic decision-making and operational efficiencies. This structured discussion sets the stage for understanding the interplay of TF, SEM, and DF within the complex ecosystem of NRENs.
- **Chapter 3 - Research Phases and Objectives:** This chapter comprehensively describes the distinct phases of the research study, along with the objectives and outputs integral to each phase during the development, application, and assessment of the TFSEMDF framework. It traces the study's progression, highlighting the processes and methodologies employed at each stage. Additionally, the chapter includes a research roadmap, which further illustrates this progression, detailing the systematic approach and sequence of events that led to the results discussed in the subsequent chapters.
- **Chapter 4 - Phase 1: TFSEMDF Framework Development:** This chapter delves into the methodology employed in crafting the TFSEMDF framework, which integrates SEM, DF and technology indicator relational mapping for TF. It details the processes and approaches involved in this integration. The results section of the chapter presents the TFSEMDF framework that emerged from this synthesis, offering a comprehensive

view of its structure and components. Subsequently, the discussions section critically evaluates this developmental phase's outcomes, scrutinising the methodology's effectiveness and the framework's functional capabilities, providing a deep understanding of the TFSEMDF framework's potential applications, strengths and weaknesses.

- **Chapter 5 – Research Phase 2: TFSEMDF Framework Application:** This chapter considers the practical application of the TFSEMDF framework, detailing the implementation processes for the autoregressive and cross-sectional model instantiations within the context of NRENs. It showcases extensive results from regression analyses, providing a deep dive into the data-driven insights obtained. Additionally, the chapter critically discusses the utility and effectiveness of the TFSEMDF framework, specifically in the NREN technology domain. This analysis not only assesses the framework's applicability and performance but also offers insights into its potential improvements and impact within the field of TF for NRENs.
- **Chapter 6 – Research Phase 3: TFSEMDF Framework Assessment:** This chapter systematically evaluates the TFSEMDF framework, scrutinising its inherent strengths and weaknesses through a systematic analytical approach. The assessment is grounded in empirical results from various NREN model instantiations, such as the cross-sectional model instantiation considered in Chapter 5. The chapter encompasses a detailed discussion which forms a comprehensive evaluation of the framework, focusing on its robustness and applicability in the context of these model instantiations. This critical examination sheds light on the efficacy of the TFSEMDF framework and provides insights into its practical utility and relevance in NREN TF.
- **Chapter 7 – Research Conclusions and Future Work:** The final chapter consolidates the study's findings, addressing the conclusions from the three research phases. Additionally, this chapter considers related research and suggests directions for future work, identifying opportunities to develop further and refine the TFSEMDF framework. These proposed research avenues build upon the existing work, addressing identified gaps and potentially expanding the framework's applicability and effectiveness in future studies.
- **Appendix A - PLS Regression Analysis for TFSEMDF:** This appendix thoroughly examines PLS regression analysis within the context of SEM. It delves into the appropriateness of PLS regression for the TFSEMDF framework, providing a rationale for its selection and use. Furthermore, the appendix outlines the step-by-step process of

implementing PLS regression using the SmartPLS software (Ringle et al., 2022), based on and extending from Staphorst et al. (2015).

- **Appendix B - TFSEMDF Reliability and Validity Analysis:** The second appendix centres on the reliability and validity analysis of the TFSEMDF framework. It elaborates on the methods employed for evaluating the measurement component (outer model) and the structural component (inner model) in the context of PLS-SEM, expanding on Staphorst et al. (2017). This section explores the assessment techniques and criteria used to ensure the integrity and accuracy of the model's measurement and structural aspects. By detailing these processes, the appendix serves as a comprehensive guide for understanding and replicating the reliability and validity checks integral to the robust application of the TFSEMDF framework within PLS-SEM.
- **Appendix C - TERENA NREN Compendium Excerpts:** The last appendix of this study provides selected excerpts for illustrative purposes from the data provided in the 2011 and 2012 TERENA NREN Compendiums (TERENA, 2011, 2012), used as data sources for the study's PLS regression analyses for the various NREN model instantiations of the TFSEMDF framework constructed in this study. The excerpts from TERENA (2011) feature NREN funding sources and core traffic levels, while the excerpts from TERENA (2012) include the types and numbers of connected institutions.
- **Appendix D – Research Data Repository:** This appendix highlights the study's commitment to open science principles by discussing the hosting of the various research datasets used and generated by this study on Figshare. It outlines the Figshare private data collections containing the study's datasets, facilitating future scholarly exploration and collaboration.

1.7 CONCLUDING REMARKS

Chapter 1 of this thesis introduces the study's multi-faceted research problem, which encompasses three critical shortcomings: the absence of TF methodologies that effectively model complex hierarchical technology domains, the exclusion of context-related information in forecasting, and the deficiency of appropriate TF techniques tailored for the NRENs technology domain. This chapter establishes the study's background, articulates its aims, explores the research's motivations, and enumerates the significant contributions derived from developing, applying, and assessing the TFSEMDF framework within the NREN technology domain.

The chapter commenced with an introduction to the research problem, setting the stage for the study's focus on TF in the context of rapidly evolving technological landscapes. An extensive background section includes four key domains: SEM, DF, TF and NRENs. Each domain provided a foundational understanding of the key concepts and methodologies in these areas that were relevant to the study.

Next, the chapter articulates the research's aims and scope, outlining the specific objectives the study attempted to achieve within its investigative boundaries. A detailed exploration of the research motivation follows, examining both business and academic perspectives, thereby highlighting the study's significance in practical and scholarly contexts.

The chapter then presented the study's novel contributions, detailing advancements made in the academic fields of TF and NRENs. It also considers contributions to engineering and technology management alongside the value of TFSEMDF as a tool for forecasting practitioners. Next, the chapter considered the value of the TFSEMDF framework as a tool for forecasting practitioners. The chapter then provides an overview of the core and supplementary publications generated by the study, illustrating the research's academic impact and contribution to the existing body of knowledge for TF.

Chapter 1 concludes with a detailed overview of the thesis's organisation, detailing the subsequent chapters' and appendices' structure and contents. This overview serves as a comprehensive guide, delineating how the thesis progresses and detailing the development, application, and assessment of the TFSEMDF framework explored in this study.

CHAPTER 2 - LITERATURE AND THEORY REVIEW

2.1 INTRODUCTION

This chapter is dedicated to a rigorous literature and theory review, focusing on creating a solid foundation in SEM, DF and TF. These critical areas are the fundamental building blocks used in this study to develop the TFSEMDF framework. In addition, this chapter also rigorously examines NRENs, selected as the test technology domain used for applying the TFSEMDF framework to perform both longitudinal and transversal TF and to assess the framework's strengths and weaknesses.

The chapter commences with a comprehensive overview of SEM, encompassing its historical development, core principles, and the methodologies of the covariance-based and PLS regression approaches. The focus then shifts to the mathematical foundations of SEM, considering both critical theoretical concepts and notation standards for SEM path diagrams. A comparative analysis of transversal SEM, i.e. cross-sectional SEM, and longitudinal SEM techniques, like autoregression and CLPM SEM, follows. The exploration into SEM concludes by examining the relatively scarce yet crucial literature on applying SEM in TF and DF, providing critical insights into integrating these methodologies.

The chapter then addresses the field of DF. This exploration starts by delving into the history of DF, providing insight into its origins and development. An overview of critical concepts in DF follows, concentrating on the core principles and techniques fundamental to this field. The focus then shifts to the JDL Data Fusion Group (DFG) framework, initially developed for military applications. Additionally, the chapter examines a variety of DF methodologies, illustrating the diverse strategies employed in the field. The application of context-related information within DF receives special attention, underscoring the importance of context in enhancing the effectiveness of DF processes.

The chapter's exploration of TF commences with a historical review, mapping the evolution of the field and its methodologies. The discussion then moves to a thorough analysis of technology indicators and TF output metrics, underscoring their significance in the functionality of the TFSEMDF framework. This examination highlights how these elements are integral to practical technology trend analysis and forecasting. The chapter also extensively explores various TF methodologies, explicitly focusing on TFDEA. This detailed scrutiny of TFDEA and other methods illuminates their diverse applications and effectiveness across different forecasting scenarios, thus enriching the understanding of the TFSEMDF framework within the expansive landscape of TF. An investigation into the metrics for the measurement of the level of success achieved in TF ended this section.

Next, the literature and theory review chapter delves into NRENs. This exploration begins with a study of their history, shedding light on the evolution and role of NRENs within the global research, education, and innovation communities. The literature review then examines RREN and NREN organisations in Africa. The section concludes by reviewing the annual TERENA and GÉANT NREN Compendiums, valuable resources that contain extensive technology and context-related information from NRENs in Europe.

The chapter then transitions into its final section, which includes analysis and synthesis of the existing literature on SEM, DF, and TF within the NREN technology domain. This exploration involves synthesising key concepts and critically assessing existing frameworks, thematic insights, and future research directions. It comprehensively explains the methodologies' roles in enhancing strategic decision-making and operational efficiency within NREN ecosystems.

This chapter's literature and theory review, underpinning the development and application of the TFSEMDF framework, aligns with and builds upon the condensed versions of these topics previously published in Staphorst et al. (2013, 2014, 2015, 2016a, 2016b, 2017). These publications have laid the groundwork for the in-depth exploration presented in this chapter, ensuring a fundamental understanding of the critical components of the TFSEMDF framework and its application within the context of NRENs.

2.2 STRUCTURAL EQUATION MODELLING

SEM has emerged as a versatile and powerful statistical tool, offering second-generation modelling regression approaches to understanding complex relationships within data. Its applications span various research fields, providing insights into the intricate interplay of variables. The following sections delve into an overview and history of SEM, its application in DF and TF, the mathematical foundation of SEM, and the distinctions between longitudinal and transversal SEM. These sections aim to reveal the multifaceted nature of SEM, highlighting its evolution, methodological intricacies, and diverse applications in contemporary research.

2.2.1 History and Overview SEM

Haenlein and Kaplan (2004) argue that regression analysis encompasses a spectrum of statistical methodologies to model and scrutinise the relationships between dependent and independent variables within empirical datasets. Fundamental to this paradigm is the concept of regression functions, which attempt to delineate the fluctuations in dependent variables as functions of alterations in independent variables. This analytical concept facilitates forecasting future values for a dependent variable predicated upon the established values of independent variables, thus serving as an indispensable mechanism for predictive analyses across various disciplines (Haenlein & Kaplan, 2004).

Additionally, regression analysis can infer causal relationships between variables in specific contexts (Bollen & Curran, 2004). This capability is helpful in studies where understanding cause-and-effect dynamics is crucial. However, it is imperative to approach such inferences cautiously, considering the potential for unanticipated factors and the need for careful data interpretation. This multifaceted utility of regression analysis, from prediction and forecasting to causal inference, highlights its significance as an indispensable empirical research methodology (Bollen & Curran, 2004).

The traditional regression methodologies encompassing multiple regression, discriminant analysis, logistic regression, and variance analysis are aptly categorised first-generation techniques. This classification is primarily due to their inherent assumption of independence among multiple dependent variables, a perspective thoroughly analysed by Haenlein and Kaplan (2004). The principal drawback of this assumption is its restrictive nature, particularly in modelling the intricate interdependencies that often exist in real-world data, such as the interplay between two or more output variables in a TF model.

While first-generation regression techniques have played a pivotal role in the evolution of statistical analysis, their limitations in modelling complex interdependencies are evident. Jöreskog (1973) introduced CB-SEM as a second-generation technique to address these limitations. This innovative approach marked a significant advancement in statistical modelling, allowing for the simultaneous modelling of relationships among multiple dependent and independent constructs. CB-SEM is adept at unravelling complex interdependencies among variables, offering a more holistic view of the data. This technique goes beyond analysing isolated bivariate relationships. It establishes a framework for understanding the entire system of relationships within the dataset, thus offering a more accurate and comprehensive model of the underlying phenomena. The advent of CB-SEM as a second-generation technique provided a more robust and sophisticated tool for understanding the intricate relationships that characterise many datasets, thereby enhancing the accuracy and depth of statistical analysis (Dash & Paul, 2021).

First-generation regression techniques, such as multiple and logistic regression, are fundamentally limited by their assumption that all dependent and independent variables are directly observable (Haenlein & Kaplan, 2004). This assumption necessitates that the values of all variables be obtainable through empirical sampling, leading to the exclusion of unobservable variables, or latent constructs, from these models (Haenlein & Kaplan, 2004). This exclusion is a notable limitation, as latent constructs are often crucial in various research contexts. In contrast, SEM effectively incorporates latent constructs, overcoming the limitations of first-generation techniques (Haenlein & Kaplan, 2004). SEM's flexibility extends to handling non-linear and non-Gaussian factors and cyclical dependencies among variables, whether latent or observable, thus offering a more comprehensive approach to statistical modelling.

SEM differentiates between exogenous and endogenous latent constructs. Exogenous constructs are those variables that remain external to the model's internal dynamics, consistently acting as independent variables not influenced by other variables within the model (Haenlein & Kaplan, 2004). In contrast, endogenous constructs derive their definition from their interrelationships within the model, influencing and being influenced by other dependent or independent variables (Haenlein & Kaplan, 2004). This differentiation moves beyond the binary classification of dependent and independent variables, highlighting SEM's ability to intricately map and interpret complex pathways and interdependencies among variables (Haenlein & Kaplan, 2004).

Indicators represent latent constructs in SEM and fall into two distinct categories (Haenlein & Kaplan, 2004). A set of measured proxies characterises latent constructs with reflective indicators, often referred to as factors, that exhibit high correlations with the latent construct and other potential reflective indicators of the same construct. This high degree of correlation implies that these indicators effectively capture the variance in the unobserved latent variable (Haenlein & Kaplan, 2004). Conversely, a weighted amalgamation of indicators represents latent constructs with formative indicators, which do not necessarily demonstrate high correlations with the latent construct or among themselves. These formative indicators, each contributing a unique dimension, collectively define the breadth of the latent construct. Unlike reflective indicators, which each manifest the latent construct, formative indicators constitute distinct facets of the latent construct (Haenlein & Kaplan, 2004). This distinction between reflective and formative indicators is crucial in SEM for accurately conceptualising and measuring latent constructs, as it influences both the interpretation of the constructs and the statistical methods used for their analysis.

Jöreskog's seminal work in 1973 laid the foundation for the estimation of SEM parameters using covariance-based techniques, with the Linear Structural Relations (LISREL) program, developed by Jöreskog in 1975, becoming a prominent tool in this domain (Jöreskog, 1975). However, variance-based techniques, also known as component-based techniques, have gained considerable traction in the field (Haenlein & Kaplan, 2004). One notable variance-based technique is PLS regression, initially introduced by Wold as Non-Iterative Partial Least Squares (NIPALS) (Haenlein & Kaplan, 2004; Wold, 1975). PLS-SEM, evolving from this foundation, has been recognised as a critical second-generation technique, particularly for its efficacy in handling complex models and its suitability in exploratory research (Dash & Paul, 2021; Vinzi et al., 2010).

The methodological divergence between covariance-based techniques and PLS regression is noteworthy. Covariance-based approaches minimise the discrepancies between the sample covariance values and those forecasted by the model. This process involves estimating model parameters in a way that reproduces the covariance matrix of the observed measurements. In contrast, PLS regression, also called "Projections to Latent Structures," prioritises maximising the variance explained in the dependent variables by the independent variables (Haenlein & Kaplan, 2004). This distinction underscores each technique's unique strengths and applications, with PLS-SEM offering a more variance-focused approach, enhancing its utility in scenarios where prediction and exploratory analysis are paramount.

2.2.2 MATHEMATICAL FOUNDATION OF SEM

Path diagrams are a critical tool for conceptualising system models in SEM analysis. A generic SEM path diagram, such as the one presented in Figure 1, illustrates the configurations of exogenous and endogenous constructs, the path coefficients delineating the interconnections between these constructs, and the reflective and formative indicators, including their loadings on the constructs. This comprehensive representation originates from the work of Chin and Newsted (1999).

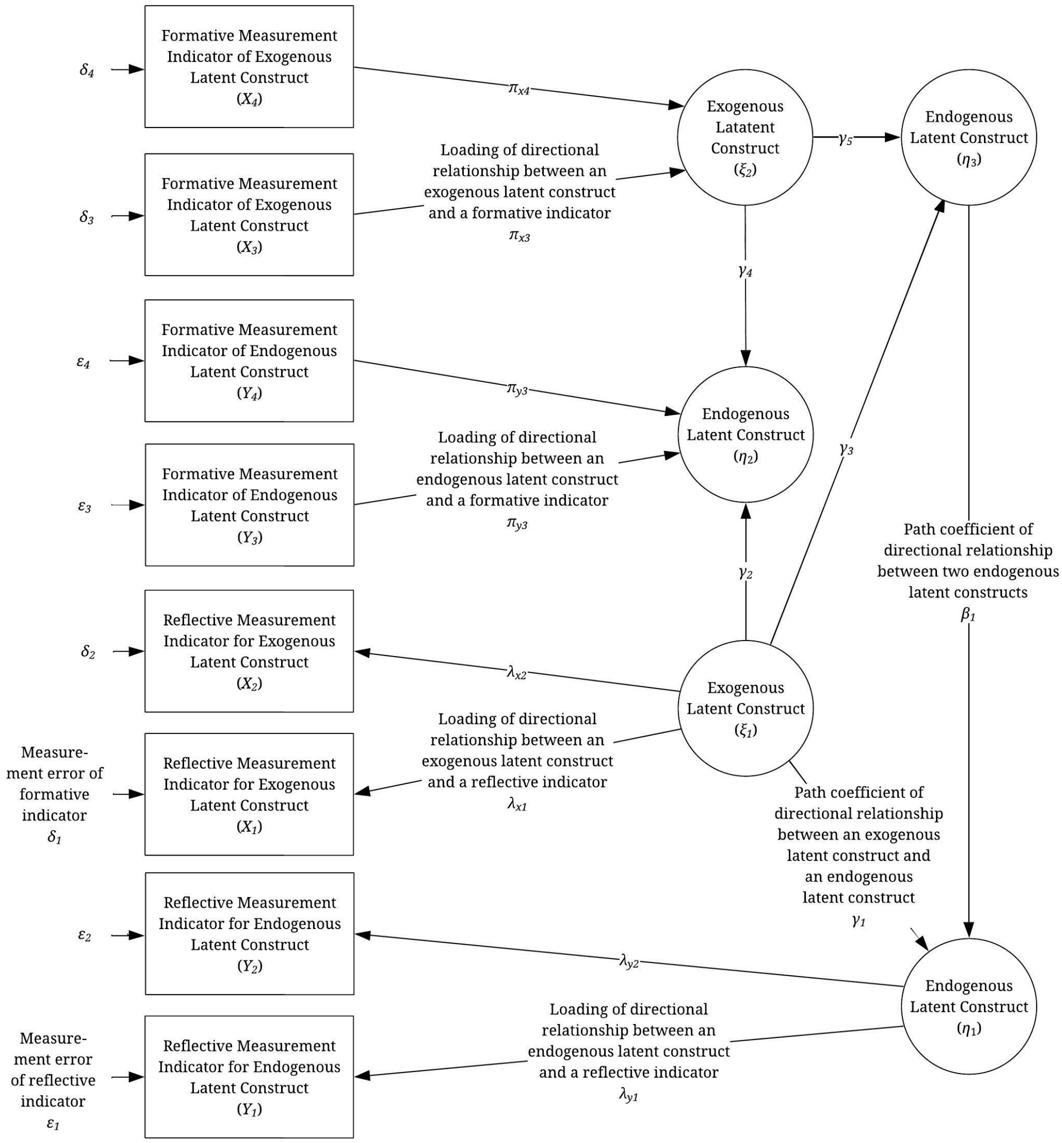


Figure 1: Generic SEM Path Diagram

When constructing such diagrams, it is imperative to adhere to established schematic conventions, as detailed by Haenlein & Kaplan (2004):

- Within the SEM framework, circles or ellipses depict constructs, symbolising the abstract nature of these elements. These constructs can be directly observable or latent.
- Measurement indicators are distinctly represented by squares or rectangles, differentiating them from the more abstract constructs.
- Single-headed arrows indicate directional relationships within the model. Specifically, in relationships between indicators and their associated constructs, arrows are directed towards reflective indicators, signifying the construct's influence on the indicator. Conversely, arrows point from formative indicators towards the constructs, indicating the indicators' contributory role in defining the construct.
- Non-directional relationships, which are less common, are denoted by double-headed arrows. This convention is occasionally employed to represent the variance of a variable, where a double-headed arrow loops from the variable back to itself. However, this study did make use of this specific convention.
- Each arrow in the diagram symbolises a parameter that can be free or fixed. Fixed parameters appear with explicit numerical values, while appropriate mathematical symbols represent free parameters, indicating their variable nature.
- Exogenous constructs refer to those not influenced by other constructs in a model instantiation. Conversely, constructs that other constructs influence in a model instantiation are endogenous.
- Directional arrows show the influence that endogenous or exogenous constructs exert on another endogenous construct by pointing from the influencing constructs to the influenced one.
- Directly measurable constructs are known as observable constructs or manifest constructs. The actual data points or measurements that researchers collect truthfully represent these constructs.
- Researchers refer to constructs that cannot be directly measured as latent constructs and estimate them from sets of either formative or reflective indicators.
- Reflective indicators are manifestations of their associated latent construct. Any changes in the latent construct should lead to changes in all its reflective indicators.
- Formative indicators are assumed to cause their associate latent construct. They are separate facets, or components, that jointly form the construct.

These schematic conventions are integral to the clarity and interpretability of SEM path diagrams, ensuring they accurately convey the hypothesised relationships and structural intricacies of the model under investigation. As shown in Figure 1, SEM-specific symbol conventions represent various elements within a model instantiation's path diagram. Below are the conventions as outlined by Staphorst (2010):

- ξ_n denotes the n^{th} exogenous latent construct, representing a variable that influences others within the model but is not itself influenced by other variables in the model.
- η_m symbolises the m^{th} endogenous latent construct, which is influenced by other variables within the model and potentially influences other endogenous latent constructs.
- X_i refers to the i^{th} measurement indicator, representing an observable variable associated with the n^{th} exogenous latent construct ξ_n.
- δ_i is the measurement error term associated with X_i, encompassing random and systematic errors attributable to the measurement method rather than the construct itself.
- Y_j represents the j^{th} measurement indicator, an observable variable associated with the m^{th} endogenous latent construct η_m.
- ε_j is the measurement error term associated with Y_j, comprising both random and systematic components.
- λ_{xi} denotes the loading of a directional relation between the n^{th} exogenous latent construct ξ_n and its i^{th} reflective indicator X_i.
- λ_{yj} represents the loading of a directional relation between the m^{th} endogenous latent construct η_m and its j^{th} reflective indicator Y_j.
- π_{xi} denotes the loading of a directional relation between the n^{th} exogenous latent construct ξ_n and its i^{th} formative indicator X_i.
- π_{yj} represents the loading of a directional relation between the m^{th} endogenous latent construct η_m and its j^{th} formative indicator Y_j.
- γ_c signifies the path coefficient of a directional relation between the m^{th} endogenous latent construct η_m and the n^{th} exogenous latent construct ξ_n, indicating the exogenous construct's influence on the endogenous construct.
- β_d denotes the path coefficient of a directional relation from the q^{th} to the p^{th} endogenous latent constructs, η_q and η_p, representing the influence of one endogenous construct on another.
- ζ_r represents the r^{th} disturbance term (or error term) in the r^{th} endogenous latent construct η_r. This term accounts for the variance in the endogenous latent construct not explained by the independent variables (not depicted in Figure 1).

These conventions facilitate a standardised representation of SEM models, allowing for clear communication and interpretation of the complex relationships and constructs within the model. Utilising these conventions for SEM, it is possible to construct five sets of structural equations that comprehensively represent the interrelationships inherent in an SEM model. Implementing matrix notation in this context is particularly advantageous, offering a structured and precise method to delineate these relationships. According to Staphorst (2010), the first set of structural equations describes the relationship between exogenous latent constructs and their reflective indicators, along with the associated measurement errors:

$$\boldsymbol{X} = \boldsymbol{\Lambda}_x \boldsymbol{\xi} + \boldsymbol{\delta} \quad 2.1$$

where the elements of matrices $\boldsymbol{X}$, $\boldsymbol{\Lambda}_x$, $\boldsymbol{\xi}$ and $\boldsymbol{\delta}$ correspond to X_i, λ_{xi}, ξ_n and δ_i, respectively, for all relevant values of i and n. The second set of equations articulates how endogenous latent constructs are functions of their reflective indicators and associated measurement errors (Staphorst, 2010):

$$\boldsymbol{Y} = \boldsymbol{\Lambda}_y \boldsymbol{\eta} + \boldsymbol{\varepsilon} \quad 2.2$$

where the elements of matrices $\boldsymbol{Y}$, $\boldsymbol{\Lambda}_y$, $\boldsymbol{\eta}$ and $\boldsymbol{\varepsilon}$ are Y_j, λ_{yj}, η_m and ε_j, respectively, for all pertinent values of j and m. The third set of equations addresses the relationships between exogenous latent constructs and formative indicators, along with measurement errors (Staphorst, 2010):

$$\boldsymbol{\xi} = \boldsymbol{\Pi}_x \boldsymbol{X} + \boldsymbol{\delta} \quad 2.3$$

where the elements of matrices $\boldsymbol{\xi}$, $\boldsymbol{\Pi}_x$, $\boldsymbol{X}$, and $\boldsymbol{\delta}$ are ξ_n, π_{xi}, X_i and δ_i, respectively, for all applicable values of n and i. The fourth set of equations focuses on the relationships between endogenous latent constructs and formative indicators, as well as measurement errors (Staphorst, 2010):

$$\boldsymbol{\eta} = \boldsymbol{\Pi}_y \boldsymbol{Y} + \boldsymbol{\varepsilon} \quad 2.4$$

where the elements of matrices $\boldsymbol{\eta}$, $\boldsymbol{\Pi}_y$, $\boldsymbol{Y}$ and $\boldsymbol{\varepsilon}$ are η_m, π_{yj}, Y_j and ε_j, respectively, for all relevant values of m and j. The final set of equations deals with the relationships between endogenous and exogenous latent constructs, including the associated measurement errors (Staphorst, 2010):

$$\boldsymbol{\eta} = \boldsymbol{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad 2.5$$

where the elements of matrices $\boldsymbol{\eta}$ (left-hand side of Equation (2.5)), $\boldsymbol{B}$, $\boldsymbol{\eta}$ (right-hand side of Equation (2.5)), $\boldsymbol{\Gamma}$, $\boldsymbol{\xi}$ and $\boldsymbol{\zeta}$ are η_q, β_d, η_p, γ_c, ξ_n and ζ_r, respectively, for all applicable values of q, d, p, c, n and r (Staphorst, 2010). $\boldsymbol{\eta}$ appears on both sides of Equation (2.5) as endogenous constructs may depend on one another.

These equations provide a rigorous mathematical framework for understanding and analysing the complex web of relationships among latent constructs and their reflective and formative indicators within an SEM model. Researchers can achieve estimation of the coefficients in these SEM equations using various methods, including CB-SEM and PLS-SEM (Dash & Paul, 2021).

2.2.3 Longitudinal vs Transversal SEM

Transversal (cross-sectional) SEM and longitudinal SEM represent two distinct approaches to statistical analysis, each with unique applications and implications. Cross-sectional SEM examines relationships between variables simultaneously, providing a snapshot of the interrelations among variables. This approach is beneficial in studies where the goal is to understand the correlational structure of variables at a specific moment, offering insights into the concurrent relationships among them. However, it does not account for changes or developments over time, which is critical in many research areas (Bollen & Curran, 2004).

Longitudinal SEM, on the other hand, delves into the dynamic interrelationships and causal inferences over time. This approach includes methodologies like the Cross-lagged Panel Model SEM (CLPM SEM) and autoregression techniques. CLPM SEM is particularly effective in longitudinal studies for identifying directional influences between variables across different time points. It addresses temporal precedence, which is crucial for establishing causal inferences. This method has been highly valued in psychology and social sciences for examining the interplay of behaviour and attitude over time (Hamaker et al., 2015; Selig & Preacher, 2009). Autoregression in SEM, meanwhile, focuses on the self-influence of a variable over time, distinguishing short-term fluctuations from long-term trends. This technique is integral in time-series analysis and forecasting models, where understanding the persistence and evolution of a variable is key (Bollen & Curran, 2004; Burant, 2022).

Another emerging longitudinal SEM technique is Dynamic Structural Equation Modelling (DSEM). DSEM, as described by Asparouhov et al. (2018), is a framework for analysing the evolution of observed and latent variables over time within structural equation models. Particularly suited for intensive longitudinal data, which involves multiple observations from several individuals over numerous time points, DSEM is a comprehensive integration of time-series and structural equation modelling techniques. The estimation of DSEM is conducted using Bayesian methods, specifically the Markov chain Monte Carlo Gibbs sampler and the Metropolis-Hastings sampler, making it versatile for longitudinal analyses of any duration and frequency of observations (Asparouhov et al., 2018; McNeish & Hamaker, 2020).

The versatility of SEM in incorporating both cross-lagged and autoregressive components underlines its effectiveness in exploring temporal dynamics across various research fields. By integrating these elements, SEM allows researchers to dissect and understand the complex temporal interplay among variables, offering a more comprehensive view of the phenomena under study. This integration is particularly beneficial in fields like economics, finance, and behavioural sciences, where understanding variables' immediate and long-term effects is essential (Box et al., 2015).

2.2.4 Application of SEM in DF and TF

Steinberg's research emphasises the suitability of SEM as a statistical tool for implementing DF in natural language processing, particularly noting its capability to incorporate context sensitivity in solving DF inferencing problems (Steinberg, 2009; Steinberg & Rogova, 2008). Steinberg (2009) defines a 'situation' or 'context' as a network of relationships representing a specific instantiated relation. In DF inferencing, context is crucial for refining ambiguous estimates, clarifying available data, and constraining processing during data acquisition, cueing, or fusion (Steinberg, 2009).

Moreover, Steinberg aligns the terminologies of DF and SEM. He suggests that variables in DF problems are equivalent to endogenous constructs in SEM, while context variables in DF resemble exogenous constructs in SEM. Additionally, traditional DF sensor measurements correspond to the reflective and formative indicators in SEM. This conceptual alignment clarifies the relationship between DF and SEM. It highlights the adaptability of SEM in complex inferencing scenarios, where it accounts for internal system dynamics and external context-related influences.

Sohn and Moon (2003) critically analysed the prevalent limitations in most TF techniques, particularly their oversight of the structural relationships between technology indicators and TF output metrics. In contrast, SEM offered a significant methodological advantage by facilitating the modelling of intricate hierarchical relationships between these technology indicators and TF output metrics. This capability of SEM extended beyond the scope of traditional TF techniques, which often needed to account for such complex interdependencies.

Sohn and Moon (2003) demonstrated the efficacy of SEM as a regression technique in evaluating multi-layered hierarchical models. This approach involved progressive aggregations and refinements of input technology indicator data, culminating in reliable TF output metrics statistical estimates. Specifically, their research utilised the TCSI metric, a market prospect indicator within the framework proposed by Watts and Porter (1997), as the primary TF output metric in their SEM model (Sohn & Moon, 2003). This application of SEM in their study underscored its versatility and effectiveness in handling complex, layered data structures, thereby providing a more accurate analysis of TF outputs.

2.3 DATA FUSION

DF is a dynamic and multifaceted field, continuously evolving and expanding its scope beyond its initial military applications to encompass a wide range of disciplines. DF has become an essential tool in the modern data-driven landscape, from its roots in military strategy to its current applications in diverse fields such as environmental monitoring and healthcare. The following sections provide a comprehensive overview of DF, detailing its evolution, the seminal JDL/DFG framework, various DF methodologies, and the critical role of context sensitivity in enhancing the process of data integration and analysis.

2.3.1 HISTORY AND OVERVIEW OF DF

DF has evolved substantially as an interdisciplinary field, drawing insights and methodologies from computer science, engineering and intelligent systems. Its foundational principle, rooted in the innate biological propensity to merge sensory inputs for enhanced survival, was adeptly transformed for use in technological contexts (Lahat et al., 2015). Integrating heterogeneous data sources was instrumental in advancing DF methodologies, enabling a deeper understanding and interpretation of complex phenomena.

Buchroithner (1998) and Wald (1997) conceptualised DF as a structured framework designed for synthesising data from diverse sources, initially developed within the military domain to enhance the generation of tactical knowledge through multi-layered sensor data processing. They described DF as a process that amalgamated data from varied origins, aiming to yield superior quality information, with the specific definition of 'superior quality' contingent upon the application context (Buchroithner, 1998; Wald, 1997).

The formalisation of DF, particularly in military and defence applications, represented a significant milestone in the field's evolution. During the late 1980s and early 1990s, the establishment of the JDL DF model introduced a structured approach to the fusion process. Hall and Llinas (1997) described that this model organised DF into various levels, including source preprocessing, object refinement, and situation assessment. This hierarchical framework clarified the DF process and enabled a more sophisticated approach to integrating diverse data sets. The introduction of the JDL model was crucial in laying a solid foundation for future advancements in DF, providing a flexible framework widely adopted and adapted for numerous applications beyond its original military and defence context.

In recent times, DF's application has extended to diverse fields, such as environmental monitoring, healthcare, robotics, and financial systems, as Wald (1997, 1999) and Hall and Llinas (1997) noted. This broadening of scope indicates DF's inherent versatility and capacity to assimilate and process data from many sources. The progression of DF into these varied domains reflects its fundamental ability to integrate and synthesise disparate data streams, thereby producing tailored information aligned to the requirements of various applications. This capability, as highlighted by Waltz & Llinas (1990) and Khaleghi et al. (2013), underscores the adaptability of DF techniques in addressing complex data integration challenges across various sectors. The transition from a military-centric focus to a universal application demonstrates the robustness of DF methodologies in handling diverse data types and formats, making it an indispensable tool in the modern data-driven landscape (Gagolewski, 2022).

Integrating Machine Learning (ML) techniques with DF methodologies marked a transformative phase in the field (Diez-Olivan et al., 2019; Meng et al., 2020). The advent of ML algorithms, especially those predicated on neural networks and deep learning paradigms, significantly augmented the proficiency of DF systems in managing and interpreting high-dimensional and intricate datasets (Diez-Olivan et al., 2019; Meng et al., 2020). These algorithms demonstrated a remarkable ability to discern patterns and relationships within data, thereby enabling more efficient and adaptable fusion processes. This synergy between ML and DF proved advantageous in environmental monitoring, healthcare, and intelligent transportation systems (Diez-Olivan et al., 2019; Meng et al., 2020). The sheer complexity and volume of data in these domains called for advanced analytical methods to derive meaningful insights from various data sources. (Lahat et al., 2015).

A salient development in this regard is the application of deep learning in multimodal remote sensing DF. This approach has been instrumental in addressing the complexities associated with heterogeneous Earth observation data. The work of Li et al. (2022) stands as a testament to this advancement, offering a comprehensive review that underscores the significant strides made in refining traditional algorithms and the notable enhancements in performance brought about by deep learning. This advancement is particularly noteworthy in the fusion of disparate data modalities, such as spatio-spectral and spatiotemporal data, where deep learning algorithms have demonstrated exceptional efficacy.

Another area where deep learning has significantly impacted is the biomedical sector. The research by Stahlschmidt et al. (2022) delves into utilising multimodal deep learning for DF in biomedical applications. Their study illuminates the potential of deep learning in amalgamating various types of data to facilitate enhanced analysis, especially in the context of complex biological systems. This development highlights how DF's nature is evolving, with practitioners harnessing cutting-edge machine-learning techniques to push the boundaries of data analysis and interpretation in diverse scientific and practical applications.

2.3.2 JDL/DFG FRAMEWORK

The JDL DF model, a cornerstone in the field of DF, systematically delineated the DF process into a series of distinct levels. As Steinberg and Bowman (2017) explained, this framework provides a structured methodology for comprehending and implementing DF, particularly in complex environments. Central to this model was the recognition by the JDL/DFG that DF, within a military context, involves a systematic aggregation and refinement of sensor data to generate high-quality tactical knowledge. This understanding established a standard structure for the multi-layered DF process, applicable across various military scenarios and implementations. The JDL/DFG articulated this structure through the definition of six levels of processing, each representing a progressive stage in the DF continuum (Steinberg & Bowman, 2017):

- **Level 0 - Signal/Feature/Subject Assessment:** This foundational level focuses on the initial assessment of raw signals, features, or subjects, setting the stage for more advanced processing.
- **Level 1 - Object Assessment:** At this level, the emphasis shifts to evaluating and characterising objects based on the processed data.
- **Level 2 - Situation Assessment:** This level involves synthesising information to understand and assess the broader situational context.
- **Level 3 - Impact Assessment:** Here, the focus is on evaluating the potential impacts or outcomes of the situation under analysis.
- **Level 4 - Process Refinement:** This advanced stage aims to refine the DF process, enhancing its efficiency and effectiveness.
- **Level 5 - User Refinement:** The final level centres on tailoring the DF process's output to the requirements of the end users.

Each level represented a progressive stage in the DF process, from the initial assessment of raw signals and features at Level 0 to tailor the output to meet specific user needs at Level 5. This hierarchical structuring clarified the DF process and enabled a more sophisticated approach to integrating diverse data sets.

The introduction of the JDL model was crucial in laying a solid foundation for future advancements in DF. It provided a flexible framework widely adopted and adapted for numerous applications beyond its original military and defence context. Various academic and defence publications extensively discuss the widespread use and acceptance of the model for categorising DF-related functions (Defense Technical Information Center, n.d.; Steinberg & Bowman, 2017).

The JDL DF model has provided a clear and comprehensive framework for DF, particularly in military applications. It facilitates a more systematic and practical approach to data integration and analysis. Its evolution and adaptation over the years reflect the dynamic nature of the field of DF and its growing importance in a wide range of applications (Steinberg & Bowman, 2017).

2.3.3 DF Methodologies

In the technical landscape of DF, methodologies typically stratify into three principal categories: low-level, intermediate-level, and high-level fusion (Castanedo, 2013). Low-level fusion, often termed data-level fusion, involves merging raw data from various sources, predominantly sensors. This foundational stage aims to create an enriched dataset, synthesised to be more informative and comprehensive than its constituent inputs. The significance of low-level fusion lies in its ability to effectively harness and transform unprocessed data, setting a critical foundation for subsequent, more sophisticated stages of analysis.

Intermediate, or feature-level fusion, marks a progression to a more refined phase within the DF spectrum (Castanedo, 2013). This level involves integrating distinct features extracted from the initial data sources. These features, encapsulating specific characteristics or attributes of the observed phenomena, play a vital role in distilling the essence of the data. The process of feature-level fusion is instrumental in accentuating the most salient aspects of the data, thereby facilitating a more targeted and insightful analysis.

At the apex of the DF hierarchy stands high-level fusion, commonly called decision-level fusion. This advanced tier distinguishes itself by synthesising outcomes or decisions from an array of algorithms or classifiers, each operating on distinct datasets or extracted features. High-level fusion epitomises the ultimate stage of the DF process, where the aggregated and processed information from the preceding levels is employed to inform and shape decisions or outcomes. This stage underscores DF methodologies' comprehensive and integrative nature, showcasing the field's capacity to converge diverse data streams into coherent, actionable insights (Castanedo, 2013).

DF utilises a spectrum of classic mathematical techniques, each foundational to amalgamating data from disparate sources. These techniques are pivotal in ensuring that the resultant fused information transcends the reliability and value of the individual data inputs. The array of techniques employed in DF is diverse, encompassing statistical methods, linear algebra and matrix methods, optimisation techniques, and information theory. Each of these methodologies plays a crucial role, as summarised below, in different aspects of the DF process, from initial data processing to the final synthesis of information:

- **Statistical Techniques:** Statistical methodologies form the bedrock of numerous DF techniques, offering robust frameworks for integrating diverse data sets. Central to this domain is Bayesian inference, which adopts a probabilistic stance towards DF. This approach is instrumental in assimilating prior knowledge and managing inherent uncertainties within the data. Bayesian methods have shown efficacy in sensor fusion, where they facilitate synthesising data from an array of sensors to deduce the state of a system or environment, a process eloquently described by Manyika and Durrant-Whyte (1994). Complementing Bayesian inference is the Kalman filter, a statistical tool renowned for its application in time series analysis and navigation systems. The Kalman filter excels in estimating the state of dynamic systems based on sequences of incomplete and noise-laden measurements, thereby providing a critical tool in the DF arsenal.
- **Linear Algebra and Matrix Methods:** In DF, linear algebra and matrix methodologies are indispensable, particularly in data manipulation and transformation tasks. Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) are pivotal techniques for feature extraction and dimensionality reduction. Jolliffe (2002) explained that these methods are essential in distilling and preparing data from multiple sources for fusion, enabling the identification of critical features within the data. This process significantly reduces data complexity and augments the efficiency of the fusion process, ensuring it retains only the most pertinent aspects of the data for analysis.
- **Optimisation Techniques:** DF optimisation combines data from disparate sources to achieve the most accurate and representative output. Techniques such as linear and non-linear programming optimise the weighting of different data sources within the fusion process. Quadratic programming, a specialised form of non-linear programming, is particularly useful in scenarios necessitating a balance between multiple objectives or constraints, thereby ensuring an optimal fusion outcome.
- **Information Theory:** Information theory offers tools for evaluating the informational value derived from various data sources. Techniques like entropy and mutual information quantify the information content and the interdependence among data sources. As explained by Cover & Thomas (2006), these measures are integral in assessing the contribution of each data source and strategising their amalgamation to maximise the overall information yield in the fused data. This aspect of DF is critical in discerning the most informative and relevant data combinations, thereby enhancing the quality and utility of the fusion process.

DF relies on a rich array of classical mathematical techniques, each pivotal in synthesising data from varied sources. These methodologies provide the theoretical bedrock upon which contemporary DF algorithms and methods rely and enhance the robustness and efficacy of these applications. The synergy of these diverse techniques in DF underscores their collective importance in advancing the field, enabling the extraction of more meaningful and insightful information from complex datasets. Integrating mathematical principles with DF processes is fundamental to the ongoing evolution and sophistication of data analysis in various domains.

2.3.4 Context Sensitive DF

In the discipline of DF, the integration of context, defined as a network of relational connections and synonymous with situational analysis, is increasingly recognised as a critical factor. Context plays a pivotal role in refining data alignment and association, particularly in environments characterised by large volumes or diverse data arrays. This approach aids in reducing ambiguity and misinterpretation, enabling DF systems to identify significant patterns and relationships more effectively. The recent advancements in context-sensitive DF techniques, which focus on the characteristics of exogenous context-related variables, underscore this importance. These techniques enhance the accuracy and actionability of insights derived from complex datasets (De Paola et al., 2016; Steinberg, 2009).

The strategic use of context-related information in DF represents a significant evolution in data analysis. It involves guiding the interpretation of data, mainly when sourced from diverse and multifaceted origins, to construct a coherent and meaningful understanding of the underlying phenomena. This evolution in DF methodologies enriches the depth and accuracy of the analysis, ensuring that conclusions derive from a comprehensive understanding of the data and its environmental context (De Paol et al., 2016).

This shift towards incorporating context sensitivity in DF methodologies marks a paradigmatic shift in data analysis. It reflects a broader trend towards a holistic approach, where interpreting data occurs through the lens of its contextual background. This perspective is not merely an operational enhancement. However, it represents a fundamental change in how data is approached and utilised across various research and application fields, paving the way for more insightful data analysis in the future (De Paol et al., 2016).

2.4 Technology Forecasting

The following section thoroughly examines TF, beginning with its historical development and significant contributions that shaped its trajectory. It then delves into the vital components of technology indicators and output metrics, which are essential in understanding and predicting technological trends. The subsequent sections dissect various TF methodologies, highlighting both exploratory and normative approaches and their applicability in today's technology-driven world. Special attention is given to TFDEA, discussing its methodological advancements and practical significance. The section concludes by addressing how success in TF is measured, focusing on the strategic utility of forecasting outcomes.

2.4.1 History and Overview of TF

TF is pivotal in technology management, serving as a cornerstone for strategic decision-making processes. This intricate process encompasses systematically collecting and analysing technology-related data, transforming it into informative insights and refining it into actionable knowledge. Such a comprehensive approach is essential for organisations to navigate the rapidly evolving technological landscape effectively, empowering them to make well-informed decisions regarding technology adoption, development, and investment strategies. The work of Lichtenthaler (2004) and Chang et al. (2008) emphasises the criticality of this process, highlighting the importance of converting raw data into a form that can inform and guide strategic decision-making in technology management.

Central to the TF process is the gathering of TI, including identifying and analysing technology indicators, such as technology maturity and innovation levels. These indicators are instrumental in thoroughly characterising and evaluating technologies throughout their lifecycle. They offer valuable insights into various technologies' potential impact, scalability, and market viability. Chang et al. (2008) emphasised the significance of these indicators in facilitating a comprehensive understanding of technologies, which is crucial for accurate forecasting. Additionally, the analytical process in TF often employs forward-looking methodologies like FTA, which equips decision-makers with valuable insights, enabling a more informed and strategic approach to technology management. Porter (2005) and Coates et al. (2001) have elaborated on the instrumental role of FTA in TF, highlighting its utility in providing a long-term perspective on technological trends and developments.

TF's strategic importance in technology management stands out due to its ability to anticipate technological disruptions and identify emerging opportunities. By systematically analysing current trends and projecting future developments, TF enables organisations to align their technology strategies with anticipated market and technological changes. This alignment is critical for maintaining competitive advantage and fostering innovation in an increasingly technology-driven world. Rohrbeck and Gemünden (2011) and Saritas and Aylen (2010) have explored this aspect, discussing how TF predicts the future of technology and equips organisations with the insights needed to navigate the complexities of technological change proactively. Thus, TF extends beyond mere prediction, encompassing the strategic integration of technology insights into organisational planning and decision-making processes.

The origins of TF are rooted in the seminal research conducted by the United States Department of Defense and The RAND Corporation during the 1950s (Cho & Daim, 2013). This era heralded the emergence of TF as a distinct academic and practical field, characterised by its focus on systematic studies aimed at predicting and influencing the trajectory of technological advancements. It was a foundational period in TF's history and established the groundwork for a comprehensive field dedicated to exploring and understanding future technological landscapes. Over the following decades, TF has evolved into a multifaceted field of inquiry, marked by the development of diverse methods and approaches. These methodologies have significantly contributed to the field's growth, transforming TF into a critical discipline for investigating and anticipating the future of technology and its implications for society (Cho & Daim, 2013; Feng et al., 2022).

In the aftermath of World War II, the strategic interests of the United States government in discerning technologies of substantial military significance catalysed the transformation of TF into a more formalised and structured discipline (National Research Council, 2010). The post-war period saw the development and refinement of various TF methodologies, each designed to predict and shape the course of technological advancement. These methodologies, pioneering in their approaches, established the foundational principles of contemporary TF practices. They underscored the necessity of a comprehensive analysis that extends beyond mere technological considerations to include the broader economic and social contexts in which these technologies are embedded and operate. As documented by the National Research Council (2010), this holistic approach to TF has been instrumental in shaping the modern landscape of TF, influencing its theoretical underpinnings and practical applications.

The utilisation of TF in the private sector experienced a notable increase during the 1960s and 1970s (National Research Council, 2010). This expansion beyond military and government applications into the private sector diversified the scope of TF and catalysed the development of varied methodologies. The advent of advanced computer capabilities during this period was pivotal, enabling the processing of larger data sets and supporting more data-intensive forecasting methodologies (National Research Council, 2010). Furthermore, the Internet's emergence and networking advancements significantly expanded the data available to forecasters, enhancing accessibility and facilitating the continuous evolution of TF. Today, the field is characterised by developing new techniques and refining traditional methods, reflecting its dynamic nature and ongoing adaptation to changing technological landscapes.

In contemporary times, TF has evolved into a versatile tool with applications across diverse sectors such as environmental monitoring, healthcare, and finance. This broadening of scope signifies an increasing acknowledgement of TF's critical role in informing strategic decision-making and shaping policy development across various domains. The integration of cutting-edge computational methodologies, notably ML and deep learning, has markedly augmented the efficacy of TF. These advanced techniques have endowed TF with the capacity to generate more precise predictions about future technological trends, thereby enhancing the strategic foresight of organisations and policymakers (Lahat et al., 2015). This evolution of TF underscores its adaptability and growing relevance in a rapidly advancing technological world.

2.4.2 Technology Indicators and TF Output Metrics

Technology indicators harness empirical data to assess characteristics that influence technological advancement and its subsequent commercialisation (Porter & Cunningham, 2004). These indicators are critical for gauging the progress and direction of technological innovation. Watts and Porter (1997) conceptualised technology indicators as empirical measures derived from generalised technological innovation and progression models, exemplified by the S-curve model. This model reflects the life cycle of a technology, from its nascent stage to maturity and eventual decline.

Nyberg and Plamgren (2011) refined this concept by positing that technological indicators are indices or statistical data characterising a technology's life cycle attributes. This characterisation empowers decision-makers to enact strategic initiatives with greater precision. According to

Grupp (1998), researchers typically categorise these indicators into three primary categories based on their functional intent: input, byput, and output indicators. Input indicators measure variables connected to the catalysts of technological progress, such as R&D investment and talent acquisition. Byput indicators, on the other hand, correspond to variables that relate to the intermediate phenomena within the technological progression, such as the development of prototypes and interim benchmarks. Output indicators gauge the ultimate advancements in processes or products, which can be qualitative, quantitative, or value-rated (Grupp, 1998; Nyberg & Plamgren, 2011).

The spectrum of sources for extracting technology indicators is broad, extending from patent databases and scholarly publications, indicative of formal innovation processes (Porter & Cunningham, 2004), to more informal indicators like industry rumours and financial market fluctuations, which can reflect market expectations and speculative valuations (Nyberg & Plamgren, 2011). Among the methodologies utilised for gathering and analysing technology indicators, bibliometrics has become exceedingly prominent. This method leverages the frequency of citations, publications, or patents as quantifiable proxies for technological activity and progress within a specific domain (Nyberg & Plamgren, 2011).

Many frameworks have emerged to systematically categorise technology indicators, highlighting the complexity and significance of this subject. Nyberg and Palmgren's seminal work in 2011 assumes a central role, providing a synthesis of the frameworks proposed by Watts and Porter (1997), Grupp (1998) and Chang (2007). In their exploration, Nyberg and Palmgren distil the fundamental principles of these frameworks, emphasising their essential contributions to the systematic understanding of technology indicators. Specifically, the Watts and Porter (1997) framework comprises three main categories:

- **Technology Life Cycle Status Indicators:** These metrics are rooted primarily in the S-curve model and serve as a barometer for assessing the advancement stage of technological development along its life cycle, as well as detailing the tempo of technological growth (Nyberg & Plamgren, 2011; Watts & Porter, 1997).
- **Innovation Context Receptivity Indicators:** These metrics delve into the evaluation of the surrounding technological ecosystem, encompassing factors such as the adequacy of supporting technologies and the evolution of standards and regulatory frameworks

pertinent to the technology under scrutiny (Nyberg & Plamgren, 2011; Watts & Porter, 1997).

- **Market Prospect Indicators:** This category of indicators scrutinises the potential commercial viability of the technology in question. In this domain, TF practitioners and researchers pay particular attention to facets such as potential application areas for the technology, considerations related to intellectual property, and the technology's competitive positioning in the market (Nyberg & Plamgren, 2011; Watts & Porter, 1997).

Grupp, recognised as the initiator of the comprehensive function-based categorisation of technology indicators into input, byput, and output indicators (Grupp, H., 1998), initially classified these three indicator types based on the specific stage within the technology's life cycle where the measurement took place. The following outlines the categorisation method:

- **Resource Indicators:** This category of input indicators quantifies the various potential expenditures allocated to research, development, and innovation activities (Grupp, 1998; Nyberg & Plamgren, 2011).
- **R&D Results Indicators:** This indicator type, representing an output-oriented perspective, is geared toward assessing qualitative, quantitative, or value-rated advancements achieved in production processes or products (Grupp, 1998; Nyberg & Plamgren, 2011).
- **Progress Indicators:** Falling within the byput metric framework and exemplified by the technometric indicator (Grupp, H., 1998), this indicator type evaluates sub-phenomena within the spectrum of technological progress. For instance, the technometric indicator quantifies features or product specifications, contributing to a more comprehensive understanding of technological advancement (Grupp, 1998; Nyberg & Plamgren, 2011).

Chang's (2007) Technology Indicator Ontology (TIO) offers a comprehensive framework for categorising technology indicators. Within the TIO, technology indicators are classified into two overarching categories, each encompassing several sub-groups, providing a structured and detailed taxonomy for the systematic analysis and assessment of technological phenomena:

- **Technology Development Indicators:** Within this expansive category, TF practitioners and researchers employ an array of measures to monitor and analyse the multifaceted aspects of technology, including its evolution, transformation, progress, and overarching trends, all from a technological standpoint (Chang, 2007; Nyberg & Plamgren, 2011).

- **Market Development Indicators:** This category encompasses a spectrum of indicators directly linked to the market's development dynamics and the potential application areas for the technology under scrutiny. These indicators span a wide range of domains, including sales performance, investment patterns, and the technology's industrial applications, all integral to assessing its market viability and commercial potential (Chang, 2007; Nyberg & Plamgren, 2011).

2.4.3 TF METHODOLOGIES

Cho (2013) categorises TF methodologies into two primary classes: Exploratory and Normative. Exploratory techniques, encompassing methods such as TFDEA, S-curve analysis, and trend extrapolation, are grounded in the assumption that technological progress follows a traceable and evolutionary path. These techniques employ robust data analysis to project future technology trends and discern patterns that inform strategic planning and policy development. Normative techniques, on the other hand, are characterised by a goal-oriented approach. Methods such as Delphi analysis and relevance trees are illustrative of this class, beginning with a clearly defined technological outcome in mind. They strategically map out the necessary actions, resources, and processes to materialise the targeted technology state, often engaging stakeholders to converge upon a consensus for the desired future (Cho, 2013).

Departing from the framework Cho (2013) presented, the National Research Council (2010) defines four distinct methodology types as central to TF: Judgmental or intuitive methods, Extrapolation and trend analysis, Models, and Scenarios and simulations. This classification presents a comprehensive spectrum of approaches, offering unique perspectives and tools for forecasting future technological developments.

According to the National Research Council (2010), judgmental or intuitive methods in TF encompass a variety of qualitative techniques that draw on the specialised knowledge and foresight of experts within specific technological domains. These methodologies, rooted in expert opinion and systematic consensus-building processes, are particularly valuable when empirical data is scarce or when the complexity of the technology requires nuanced understanding. Below are detailed descriptions of several prominent judgmental or intuitive methods (National Research Council, 2010):

- **Expert Opinion**: In this approach, commonly termed a "genius forecast," an expert predicts based on their knowledge and experience. While it can yield rapid insights, this method is inherently subjective and susceptible to personal bias, making it less reliable when used in isolation (National Research Council, 2010; von Karman, 1945).
- **Panel Consensus**: This method involves a group of experts and balances individual biases by aggregating diverse opinions. However, it risks being influenced by dominant personalities or groupthink, where the desire for harmony leads to consensus at the expense of alternative viewpoints (National Research Council, 2010).
- **The Delphi Method**: The Delphi Method, designed to forge consensus among experts through iterative rounds of anonymous questionnaires, mitigates some of the biases inherent in panel-based forecasts. This structured approach facilitates the refinement of individual forecasts into a collective prediction, aiming to narrow the spectrum of opinions and achieve a consensus on the most probable future outcomes (Dalkey et al., 1969; Rowe & Wright, 1999). The method's strengths lie in its flexibility, cost-effectiveness, and adaptability to various topics, making it particularly suitable for continuous forecasting efforts (Dalkey et al., 1969; National Research Council, 2010; Stewart, 1987).

Extrapolation and trend analysis methods in TF, as outlined by the National Research Council (2010), utilise historical data to project future technological developments. These methodologies become particularly effective in the presence of extensive datasets, facilitating the identification of existing patterns and trends in future predictions. Key among these extrapolation and trend analysis methods are the following (National Research Council, 2010):

- **Trend Extrapolation**: This method involves analysing historical data to identify critical trends projected into the future. For example, Moore's Law, which predicts that computational power tends to increase twofold approximately every two years, exemplifies trend extrapolation in TF. It demonstrates how historical patterns of technological advancement allow for the extrapolation of future developments (Moore, 1965; National Research Council, 2010).
- **Gompertz and Fisher-Pry Substitution Analysis**: These analyses stem from the observation that new technologies often follow a predictable S-shaped growth curve from

introduction to maturity and market saturation. By fitting historical data to these growth curves, forecasters can estimate when a technology might reach maturity and be poised for substitution by newer innovations (Fisher & Pry, 1970; National Research Council, 2010).

- **Analogies**: Analogic forecasting involves drawing parallels between past and current technologies to predict future developments. This method requires identifying and analysing similar historical situations or technologies and using them to project future trends. Green and Armstrong (2007) suggest a structured approach to analogy forecasting that involves defining the target, selecting relevant experts, and matching outcomes to derive forecasts (National Research Council, 2010).
- **Morphological Analysis**: Teoriya Resheniya Izobretatelskikh Zadach (TRIZ), a theory based on the laws of technological evolution, uses an analytical approach to project future technological developments. It involves studying the history and evolution of technology to predict future changes. The method emphasises several principles, including increasing ideality, non-uniform evolution of subsystems, and transition from macro- to microscale, to anticipate how technologies might evolve (Fey & Rivin, 2005; Kucharavy & De Guio, 2005; National Research Council, 2010).

In TF, diverse mathematical and computational approaches simulate and prognosticate technological progress. As the National Research Council (2010) articulates, these methods presuppose the presence of adequate data to construct and analyse sophisticated models that shed light on forthcoming trends. Integral to the architecture of TF, these methods synthesise data into a coherent forecast, underpinning strategic planning in technological innovation. Notable TF modelling approaches encompass the following:

- **Theory of Increasing Returns**: This modelling theory, contrasting with the traditional law of diminishing returns, posits that in technology and knowledge-based industries, the value of a product or service increases as it gains more users. Arthur (1996) noted that this phenomenon is particularly evident in networked markets, where early adopters and widespread usage can lead to a market "lock-in," making it challenging for newer technologies to displace established ones. This model is instrumental in understanding the dynamics of technological markets, especially those driven by network effects (National Research Council, 2010).

- **Chaos Theory and Artificial Neural Networks**: As applied to TF, Chaos Theory suggests that technological evolution can be non-linear and exhibit unpredictable behaviours, such as bifurcations and transient chaos. Wang et al. (1999) proposed using artificial neural networks to model these complex dynamics. Through pattern recognition of historical data, neural networks can identify underlying trends and potential phase transitions in technology development, offering a robust tool for predicting non-linear and disruptive technological changes (National Research Council, 2010).
- **Influence Diagrams**: Influence diagrams provide a graphical and mathematical representation of decision-making situations, particularly useful in modelling the evolution of cause-effect relationships and uncertainties in technology. Howard and Matheson (2005) described how these diagrams, as an extension of Bayesian networks, can be used to analyse the interdependencies and probabilistic outcomes in technological forecasting. They offer a structured approach to visualising decision processes and forecasting the impacts of various technology pathways (National Research Council, 2010).

The National Research Council (2010) asserts that scenarios and simulations offer a robust framework for deciphering complex interactions and conceptualising possible futures within TF. These approaches probe the intricate and frequently unforeseeable patterns of technological evolution, providing essential methodologies for mapping out the trajectory of emerging technologies. Notable techniques in this category include, but are not limited to, the following (National Research Council, 2010):

- **Scenario Planning**: Scenario planning, originating in the military and corporate strategy field, involves creating detailed narratives about alternative futures. Apologies for the oversight. Each scenario depicts a plausible future state and builds upon a foundation of current trends, historical data, and expert insights. Kahn's pioneering work at the RAND Corporation and later at the Hudson Institute exemplified this approach, encouraging strategic thinkers to consider a wide range of possibilities, including those that may seem initially implausible (Kahn, 1960; Kahn & Aron, 1962). In TF, scenario planning allows forecasters to explore various development paths of technologies and their potential impact on society and markets (National Research Council, 2010).
- **Back-casting**: Unlike traditional forecasting methods, back-casting begins with defining a desirable future state and then works backwards to the present, identifying the steps necessary to achieve that future state. This approach is beneficial in sustainable

development and innovation policy, as it focuses on long-term objectives and works backwards to understand the critical steps needed to reach those goals. Back-casting involves engaging stakeholders to articulate their future aspirations and then developing pathways to realise those visions, making it a participatory and goal-driven forecasting method (National Research Council, 2010; Robinson, 1990).

- **Dynamic Simulations and War Games**: Dynamic simulations and war games attempt to model complex systems and interactions within specified parameters. In military contexts, war games simulate battles to test strategies and tactics. These simulations allow stakeholders to test how new technologies might influence or disrupt existing systems in TF. These methods provide a dynamic environment where different actors (real people or simulated agents) interact based on predefined rules, allowing forecasters to observe the outcomes of technological interventions in various scenarios (Davis & Bigelow, 1998; National Research Council, 2010).

In addition to the four categories of TF methodologies the National Research Council (2010) has defined, it also recognises that the evolving landscape of TF has led to innovative techniques that address the dynamic nature of technological change. Some of the modern methodologies highlighted by the National Research Council (2010) that offer diverse perspectives and tools to anticipate future technological shifts include:

- **Prediction Markets**: Prediction markets treat forecasts about future events or trends as tradable assets in a virtual market. Participants buy and sell predictions based on their perceived likelihood of occurrence, leading to a market-generated forecast. This method capitalises on the wisdom of crowds, aggregating diverse opinions to update estimates in real time. Wolfers and Zitzewitz (2004) highlight the efficiency of prediction markets in aggregating information, though they acknowledge challenges in formulating some forecasting problems in market terms (National Research Council, 2010).
- **Alternate Reality Games (ARGs)**: ARGs create simulated environments where participants engage in scenarios that mimic potential real-world developments. These games blend elements of scenarios, war games, and computer simulations, allowing players to explore first-order and second-order effects of technological changes. ARGs like "World Without Oil" demonstrate how interactive simulations can reveal the impacts of various technological futures through immersive gameplay (Dator, 2009; National Research Council, 2010).

- **Online Forecasting Communities**: The Internet facilitates the formation of communities dedicated to continuous forecasting, such as Techcast. These platforms function as virtual think tanks, pooling expertise from diverse individuals to generate collective forecasts. The success of such communities hinges on the quality of participant contributions and the efficacy of integrating their collective judgment (National Research Council, 2010).
- **Obsolescence Forecasting**: This method involves envisioning scenarios where current technologies become obsolete due to new advancements. It requires forecasters to think about potential discontinuities and paradigm shifts in technology. This form of forecasting is akin to back-casting, as it starts with the end (i.e. the obsolescence of a specific technology) and works backwards to understand the factors that lead to it (Georghiou et al., 2008; National Research Council, 2010).

2.4.4 TFDEA: At the Forefront of TF Research

TFDEA has garnered significant attention and recognition within the TF academic community over the last two decades. Its prominence is rooted in the effective amalgamation of traditional Data Envelopment Analysis (DEA) principles with TF's unique requirements. This review delves into TFDEA's history, tracing its development from foundational concepts to its application in various technological domains.

The genesis of TFDEA traces back to the efforts of Charnes et al. (1978), who developed DEA to assess the proficiency of decision-making units in operational research. This breakthrough laid the foundation for later adaptations of DEA, particularly in TF. The speed of technological advancement and the complexity of managing and predicting its growth necessitated more dynamic and robust forecasting methods. This need led to the exploration and application of DEA in TF, marking a significant departure from traditional methods and setting the stage for the development of TFDEA.

TFDEA formally entered the academic community's domain in the early 2000s. Anderson et al. (2001) pioneered this area by presenting TFDEA in 2001, focusing on the enterprise database system market and applying DEA to assess and forecast its rate of change (Anderson et al., 2001). Subsequent enhancements to TFDEA were made by Anderson et al. (2002), who modified the methodology by altering assumptions about State-of-the-Art (SOA) technology at product release, allowing for a more iterative approach to measuring technological progress (Anderson et al., 2002). Inman's 2004 dissertation further detailed the steps and theoretical framework of TFDEA, providing a comprehensive guide for its application across various technological fields (Inman, 2004).

TFDEA is a complex process encompassing six key steps outlined by Durmuşoğlu & Dereli (2011). It begins with selecting relevant Decision-Making Units (DMUs), such as individual technologies or innovation projects, vital for defining the analysis scope. The second step involves determining specific inputs and outputs to measure technology performance, like R&D expenditure and market share. Constructing the DEA model with identified DMUs and their inputs and outputs is the next step, allowing for efficiency evaluation through mathematical programming. The fourth step extends the analysis longitudinally, providing a dynamic view of technological progression. Calculating the rate of change in efficiency scores over time is the fifth step, which is crucial for forecasting technological trajectories. Finally, interpreting and applying these results in a broader technical and market context concludes the TFDEA process, offering insights for strategic decision-making and policy formulation in technology management.

The practical application of TFDEA in empirical studies underscores its versatility and effectiveness in TF. For instance, Inman et al. (2005) applied TFDEA to forecast the introduction of United States fighter jets, demonstrating its superiority over traditional forecasting methods like linear regression (Inman et al., 2005). Another notable application by Anderson et al. (2008) involved TFDEA in wireless TF. Although their study faced challenges directly comparing forecasts with actual outcomes, it highlighted TFDEA's potential in managing complex technological predictions (Anderson et al., 2008). These and other studies demonstrate the practical utility of TFDEA in a range of technology areas, marking its evolution from a theoretical model to a valuable tool in both academic research and industry practices within TF.

Current enhancements in TFDEA primarily focus on addressing its intrinsic limitations and adapting the methodology to the rapidly evolving technological innovation landscape. A significant enhancement area is the integration of more dynamic models to account for the nonlinear and often discontinuous nature of technological progress. Bogetoft and Otto (2011) explored incorporating stochastic elements into the DEA framework to handle the uncertainties and variabilities inherent in technology development. Another notable direction includes network DEA models, as Kao and Hwang (2008) suggested, which consider the interdependencies and interactions among different technologies or components of a technological system (Kao & Hwang, 2008). This approach is particularly relevant for complex technologies where progress in one area can significantly impact others. Additionally, there is a growing interest in applying Artificial Intelligence (AI) and ML in the TFDEA framework, as these technologies can enhance forecasts' predictive accuracy and adaptability (Durmuşoğlu & Dereli, 2011).

2.4.5 MEASURING SUCCESS IN TF

Evaluating success in TF transcends mere accuracy, pivoting instead on the actionability and strategic utility of the forecast's outcomes (National Research Council, 2010). The primary objective of TF is to facilitate informed decision-making, rendering a forecast valuable if it leads to enhanced decision quality, such as optimising investment strategies, guiding research directions, or informing policy adjustments (Vanston, 2003). The inherent challenge in assessing a forecast's success lies in the retrospective nature of evaluating decision efficacy, underscoring the importance of careful forecast preparation to bolster decision-maker confidence in both the methodology and its implementation. The process of developing a TF involves three distinct phases (National Research Council, 2010):

1. **Problem Framing and Outcome Definition**: This initial phase involves articulating a clear and concise problem statement, often posed as a question. An example is the long-range Delphi forecast by RAND in 1964, which solicited predictions on urgent and feasible scientific breakthroughs over the next half-century (Gordon & Helmer, 1964).
2. **Data Collection and Analysis**: The credibility of data forms the cornerstone of any forecast. Vanston and Vanston (2004) suggest criteria for evaluating these data types, emphasising aspects such as currency, completeness, potential bias, gathering technique, statistical data relevancy, qualifications, bias, and balance for expert opinion data.
3. **Interpretation and Synthesis**: The final stage involves concluding the data using appropriate methodologies. Each methodology aligns with specific data types and yields distinct results. Vanston (2003) advocates for a multifaceted approach, proposing five perspectives, namely extrapolation, intuitive, pattern analysis, goal analysis, and counterpuncher, to enrich the forecasting process. This multi-pronged approach also echoes Martino's (1983) recommendation to consider a broad spectrum of dimensions

(technological, economic, managerial, political, social, cultural, intellectual, religious, and ecological), enhancing the robustness and reliability of the forecast.

The success of a TF is not solely contingent on its accuracy but on its capacity to inform and improve decision-making processes. Integrating diverse methodologies and comprehensive data analysis is pivotal in constructing forecasts that decision-makers can confidently rely upon.

2.5 National Research and Education Networks

This section presents a thorough literature review of NRENs, highlighting their pivotal role as connectivity and service providers for the global education, innovation, and research communities. The section commences with an extensive overview of the multifaceted roles of NRENs, followed by a detailed exploration of their historical development. The focus then shifts to the current state of NRENs and RRENs in Africa, underscoring their significance and impact in advancing research into cross-disciplinary challenges in this region. Lastly, the section offers an in-depth look at the TERENA and GÉANT NREN Compendiums, delving into their extensive contents and the associated data collection and processing methodologies.

2.5.1 History and Overview of NRENs

NRENs are vital to the global academic, research, and innovation sectors. They support a range of institutions, including Higher Education Institutions (HEIs), Technical and Vocational Education and Training (TVET) colleges, research organisations, innovation incubators, and schools in some cases. NRENs provide advanced, cost-effective, high-speed broadband services, primarily fibre optic connectivity. They enhance global research collaboration, granting access to digital libraries, learning management systems, and essential services like capacity building and roaming services like eduroam. NRENs negotiate favourable connectivity terms for their members (GÉANT, 2022; Melhem et al., 2021; SANReN, n.d.).

In addition to these services, NRENs are innovators and facilitators of cutting-edge services tailored to the unique requirements of the academic and research communities. These include virtual networking services, cloud storage and computing solutions, video conferencing facilities, and comprehensive identity and access management services. The infrastructure of NRENs incorporates advanced networking technologies like IPv6, which addresses the growing needs of IoT in academia and research domains (Aldowah et al., 2017; Al-Emran et al., 2020; Deering

& Hinden, 1998). Technologies such as MPLS (Kompella et al., 2017; Rosen et al., 2001) and SDN (Bera et al., 2017; Priyadarsini & Bera, 2021) enhance NREN efficiency and adaptability, optimising data traffic flow for bandwidth-intensive and low-latency applications crucial in research settings (GÉANT, 2022).

NRENs stand out as enablers of cross-disciplinary communities of practice, which is particularly vital in regions facing multifaceted challenges like Africa. They forge connections across diverse fields, including agriculture, bioinformatics, disaster mitigation, and telemedicine, fostering collaborative research to address pressing societal issues. The blend of adaptability and inclusiveness with their advanced technological capabilities underscores the pivotal role of NRENs in advancing education, scientific exploration, and sustainable development. Their contribution transcends the academic domain, supporting broader societal advancement and technological innovation (Foley, 2016; Melhem et al., 2021).

In nations where NREN capabilities have emerged, occasionally dispersed among multiple organisations in the public and private sectors, NRENs typically assume the strategic role of serving as the foundational networking infrastructure providers and ISPs for research and education communities. Notable examples of these entities include SANReN and TENET in South Africa (SANREN, n.d.), JANET in the United Kingdom (Cooper et al., 1991), SURFnet in the Netherlands, and RedIRIS in Spain (GÉANT, 2022).

Additionally, several RRENs have emerged to bolster NRENs within distinct geographic regions, facilitating cross-border collaboration and connectivity. An example of such collaborative efforts is GÉANT, which harmoniously unites European NRENs into a single network. UA takes centre stage in Africa, interconnecting NRENs in Eastern and Southern Africa to bolster education and research collaborations. In the United States, the landscape features the coexistence of multiple NRENs, including ESnet and Internet2, each dedicated to specific national sectors, while state-level RENs, such as KanREN, cater for specific geographical regions. It is worth noting that Internet2 also assumes the role of an RREN, championing collaboration and resource-sharing among research institutions throughout the United States (GÉANT, 2022).

The genesis of NRENs is deeply rooted in the collaborative spirit of the academic research community of the early 1970s. Initially conceived to share scarce and expensive computing resources, the United States led the charge in establishing the first national-level networks to serve the academic, research, and military communities. Prominent early NRENs in the United States include establishing the Computer Science Network (CSNET) in 1981, followed by NSFNet in 1985. In Europe, Norway's University Network (UNINET) emerged in 1976. These early networking initiatives revolutionised academic and research collaborations and laid the groundwork for the Internet as we know it. For example, the Canadian Academic and Research Networks (CAnet) and the Australian Academic and Research Network (AARNet) were instrumental in establishing their countries' first national internet backbone infrastructures, setting a precedent for the rest of the world (Twinomugisha, 2006).

The 1980s marked a significant expansion of NRENs, driven by the need for tailored, high-speed network infrastructures in academia. The launch of NSFNET in the United States in 1985 created a robust backbone linking various networks and university systems, facilitating rapid data exchange and research collaboration. NSFNET's success as a dedicated NREN model inspired its replication worldwide. The United Kingdom's JANET was established a year earlier, in 1984, and became a cornerstone of the nation's digital research infrastructure (Cooper et al., 1991). The influence of these networks on the global internet landscape was profound, demonstrating the capacity of NRENs to transform communication and research methodologies across continents (Leiner et al., 2009; Twinomugisha, 2006)

In parallel, other nations followed suit, establishing their NRENs to cater to the burgeoning demands of the academic and research sectors. The Swiss Academic and Research Network (SWITCH) initiated operations in 1987, while the Nordic University and Research Network (NORDUnet) began serving the Nordic countries in 1988. Canada inaugurated CAnet in 1990. These networks fostered national and regional connectivity and became integral to the broader tapestry of global research and education communications. As a result, the NRENs' role transcended their initial purpose, becoming a foundational element for the Internet's expansion and the facilitation of international academic collaboration (Leiner et al., 2009; Twinomugisha, 2006).

The history of NRENs in Africa closely intertwines with the Internet's early development. Pioneered by universities, the first notable instances include South Africa's UNINET establishing its first Transmission Control Protocol/Internet Protocol (TCP/IP) connection in 1991, the University of Zambia creating the Zambia Network (ZAMNET) in 1994, and Mozambique's Eduardo Mondlane University going online in 1995. UNINET was Africa's earliest physical academic and research network, later evolving into TENET. The first inter-nation academic network, East and Southern Africa Network (ESANET) was established in 1991, connecting universities across Uganda, Kenya, Zambia, and Zimbabwe, illustrating the significant role of NRENs in Africa's digital evolution.

NRENs exhibit wide-ranging capabilities and maturity levels, varying significantly from country to country. This diversity reflects the unique developmental paths and operational contexts of each NREN. The Greaves (2009) Capability Maturity Model (CMM), consisting of six distinct levels, offers a structured way to understand this diversity, categorising the developmental stages of NRENs from their inception to full maturity. Level 0 signifies no NREN or awareness of its need. Level 1 represents awareness of the lack of an NREN. At Level 2, structured conversations about establishing an NREN occur. Level 3 involves a formal commitment to create an NREN. Level 4 marks the presence of a coherent service offering, indicating the existence of an NREN. The NREN achieves level 5 when it establishes regional or global connectivity. Finally, Level 6 represents a mature NREN with advanced services and rich connections.

The Greaves (2009) CMM provides a comprehensive framework for charting the developmental journey of NRENs. It underscores the pivotal role of leadership at every stage of this progression, from inception to full maturity, emphasising the critical transitions that NRENs undergo as they evolve in capability and sophistication. (Greaves, 2009). In the context of implementing the TFSEMDF framework of this study in the NREN technology domain, using the capability maturity level as the TF output metric could be particularly interesting to the users of the framework, as this will allow for the assessment of the developmental trajectory and readiness of technologies in an NREN, providing valuable insights for strategic planning and decision-making.

2.5.2 Research and Education Networking in Africa

In the landscape of African digital infrastructure, RRENs and NRENs play a foundational role in enabling broadband connectivity for research, education, and innovation. These networks have been fundamental in coordinating broadband services for various institutions, extending beyond HEIs to encompass research organisations, innovation centres, TVET colleges, and sometimes even schools. Three RRENs support NRENs across the African continent, namely the Arab States Research and Education Networks (ASREN), WACREN and UA (Foley, 2016;

Melhem et al., 2021). These networks have facilitated connections between more than 20 African countries, linking them within the continent and to the European RREN, GÉANT. These connections have allowed African HEIs, research organisations, innovation hubs, and sometimes TVETs and schools to collaborate and access a wealth of resources and information from around the globe (Melhem et al., 2021).

While approximately 40 of the 54 African countries are associated with these RRENs, there is considerable variation in their capacity to leverage these networks, with only a few being considered mature. Notable examples of mature NRENs are SANReN and TENET in South Africa, the Kenya Education Network (KENET), and the Research and Education Network of Uganda (RENU) (Melhem et al., 2021). These networks have successfully established robust connectivity infrastructures, offering high-speed internet access and advanced digital services. The maturity of an NREN typically hinges on solid government backing and the formation of an organisation recognised and supported by both public and private HEIs. These organisations should have sufficient staffing and capability to manage administrative and technical matters, including negotiating connectivity contracts (Foley, 2016; Melhem et al., 2021).

The AfricaConnect3 project represents a significant leap forward in this domain, building upon the successes of AfricaConnect and AfricaConnect2 (AfricaConnect 3, n.d.; UbuntuNet Alliance, 2019). The project endeavours to expand high-speed internet connectivity and services across Africa with a €37.5 million contract agreement funded by the European Union and the African regional RENs. Specific goals include enhancing human capital development, improving access to e-infrastructure, developing dedicated services and applications, building human resource capacity, and raising awareness about the role of digital transformation in education and research. The project's scope underscores the strategic partnership between Africa and Europe, with GÉANT playing a critical role in network procurement and coordination for the African RENs (AfricaConnect 3, n.d.; UbuntuNet Alliance, 2019).

2.5.3 TERENA AND GÉANT NREN COMPENDIUMS

The GÉANT Association is a collaborative organisation that drives advanced network development and supports European research and education networking (GÉANT, 2022). The GÉANT network officially launched in November 2000. It originated as a collaboration project involving several European countries and organisations. GÉANT was initially developed and managed by Delivery of Advanced Network Technology to Europe (DANTE) in partnership with NRENs across Europe, and it represented a significant step in advancing research and education networking infrastructure on the continent. Similarly, TERENA was an organisation dedicated to promoting and developing high-quality international networking infrastructure across Europe's research and education community. TERENA played a significant role in fostering collaboration and technical innovation among its member NRENs. TERENA and DANTE merged their activities to form the GÉANT Association in 2015 (GÉANT, 2022).

The creation of the GÉANT Association marked a significant step in unifying these efforts, streamlining resources, and enhancing capabilities to support the European research and education community more effectively. This consolidation represented a strategic effort to foster advanced networking, collaboration, and innovation on a continental scale. The GÉANT Association continue to oversee the GÉANT network project while carrying forward the objectives and activities of both TERENA and DANTE (GÉANT, 2022).

The GÉANT Compendium of NRENs in Europe, known before 2015 as the TERENA Compendium on European NRENs, is an annual report that offers an extensive overview of the role and status of NRENs in supporting the scientific and academic community. While currently focusing on presenting critical data collected from the 43 European NRENs interconnected by the pan-European GÉANT network, past editions of this report did include data collected from RRENs and NRENs from other parts of the globe, including Africa. The latest report, the 2022 Compendium, was compiled from data provided by 40 out of the 43 European NRENs (GÉANT, 2022).

The GÉANT Compendium of NRENs measures a wide range of technical and context-related metrics to provide a comprehensive overview of the status and performance of NRENs in Europe. Some of these critical metrics include (GÉANT, 2022).:

1. **Organisational Aspects**: This includes data on the NRENs' budget allocations, staff numbers, and overall organisational structure. It explores the funding sources for NRENs, including government support, member contributions, or both.
2. **Network and Traffic**: Metrics related to the volume of data traffic handled by the NRENs, the capacity of their backbone and access networks, and the typical connectivity speeds provided to different user types, such as universities and research institutions.

3. **Service Portfolio**: Analysis of the range of services NRENs offer beyond their core role as connectivity providers. These services include cloud services, trust and identity services, and educational content and service support.
4. **User Base**: Information about the end users of NREN services, including the types of institutions served (such as universities, research institutes, schools, etc.) and the extent of their engagement with NREN services.
5. **Involvement in European Commission-Funded Projects:** Data on NRENs' participation in European Commission-funded projects indicate their role in broader European research and education initiatives.
6. **Growth Trends and Forecasts**: Projections and expectations for traffic growth and network capacity expansion over the coming years, especially considering the impact of events like the COVID-19 pandemic.
7. **Digital Education Services**: Metrics related to the support provided for digital education, including developing new services that facilitate online teaching or ease the administrative burden on educational institutions.
8. **Intercontinental Connectivity**: Information about the NRENs' global reach, including data on international bandwidth and collaborations with R&E networks outside Europe.

These technical and context-related metrics collectively provide rich insights into the NRENs' operational efficiency, service diversity, and strategic direction, enabling them to align their services with the emerging needs of the research and education community. This comprehensive overview aids in identifying trends, forecasting future developments, and facilitating informed decision-making that caters to the evolving landscape of academic and research networking (GÉANT, 2022).

Data for the GÉANT Compendium of NRENs is collected through a comprehensive and collaborative process, primarily based on responses to an annual survey sent to all NRENs in Europe. The critical steps in this data collection process include (GÉANT, 2022):

1. **Annual Survey Distribution**: The GÉANT Compendium team annually distributes a detailed survey questionnaire to all European NRENs. This survey gathers extensive information about each NREN's network, services, users, and organisational aspects.
2. **Guidance by Subject Specialists**: Subject matter experts from within the GÉANT project, such as members of the GN4-3 project, supervise the crafting of survey questions.

These specialists ensure that the survey accurately captures the diverse aspects of NREN operations and services.

3. **Collection of Responses**: NRENs respond to the survey with detailed information about their network infrastructure, service portfolio, organisational structure, user base, and other relevant data for the specified period (usually the previous year).
4. **Inclusion of Supplementary Data**: Besides the survey responses, the Compendium may include publicly available data, internal GÉANT data, and information from other relevant surveys. This supplementary data helps fill gaps and provides additional insights, particularly in trust and identity (T&I) and educational services.
5. **Analysis of Data**: The subject specialists and the Compendium team analyse the collected data. This analysis involves reviewing responses, verifying data accuracy, and synthesising information to present a coherent overview of the NREN landscape.
6. **Summarisation and Reporting**: The main findings and data points are then summarised and compiled into the Compendium report. This report extensively portrays the NRENs, highlighting trends, changes, and critical metrics.
7. **Online Compendium Version**: The data from the Compendium surveys, both current and past, are also made available through an online version of the Compendium, providing an accessible resource for further analysis and reference.

This structured data collection and analysis process ensures that the Compendium offers a detailed and accurate representation of the European NRENs, their capabilities, and their technological evolution over time. The approach adopted in gathering and interpreting this data underscores the reliability of the Compendium's findings, enhancing its value as a strategic resource for decision-making in the NREN community (GÉANT, 2022).

These Compendiums, particularly its 2011 and 2012 editions (TERENA, 2011, 2012), were the secondary data source for the TFSEMDF model instantiation development and PLS regression analysis during this study's second and third phases. The extensive and detailed dataset, amassed from a significant sample of NRENs through the rigorous and systematic data collection process described above, provides a wealth of technical and context-related measurements. These measurements readily serve as indicator data for applying and assessing the TFSEMDF framework in the NREN technology domain. Appendix C of this study provide illustrative examples of the data provided in the 20111 and 2012 TERENA NREN Compendiums, showing excerpts of the NREN funding sources and core traffic levels for TERENA (2011) and the types and numbers of institutions connected for TERENA (2012).

2.6 Integration of Literature on SEM, DF, TF and NRENs

The following sections explore the interdisciplinary integration of literature on SEM, DF, and TF within the NREN technology domain. This integration process entails synthesising and critically assessing concepts from the literature resources studied in Sections 2.2 to 2.5, exploring thematic insights and suggesting future research directions. Together, they evaluate how the SEM, DF, and TF concepts and methodologies considered in these literature resources can enhance strategic decision-making and operations at NRENs while offering a nuanced understanding of the diverse technology-related and context-related indicators within the complex NREN ecosystem.

2.6.1 Synthesis from SEM, DF, TF and NREN Literature

This section synthesises critical concepts for existing literature on SEM, DF, TF and NRENs. The synthesis explores how these methodologies, as described in the literature in Section 2.2 to Section 2.5, enhance strategic decision-making and operational efficiency within NRENs, highlighting critical insights from prior studies. This synthesis lays the foundation for more detailed critical and thematic evaluations in the following sections, informing the broader understanding of how SEM, DF, and TF can contribute to the NREN technology domain.

SEM provides a sophisticated statistical methodology for deciphering complex relationships and managing latent variables, which is integral in studying complex environments such as NRENs, laden with technology-related and context-related influencing factors. SEM effectively addresses the intricate relationships within NRENs, where the complexity of technological interactions often obscures direct measurements of factors influencing network efficiency and user satisfaction. The PLS regression technique applied to SEM offers a straightforward approach capable of testing hypotheses and complex latent and observable construct structures using small datasets, thereby allowing for a profound understanding of the dynamics within the complex NREN technology domain.

DF potentially plays a crucial role in harmonising diverse technology and context-related data sources within the NREN technology domain, encompassing network infrastructure measures, government influence, and core network traffic levels, all of which are vital for comprehensive strategic decision-making and operational management (Staphorst et al., 2016a). Implementing DF through SEM, as Steinberg (2009) proposed in the natural language processing domain, facilitates using datasets that include varied data types, such as quantitative traffic data and qualitative user feedback, essential for holistic network design and operations. DF's capabilities enable the structured combination of disparate data, yielding significant insights about the infrastructure, connectivity capabilities, service portfolios, utilisation, and reach of NRENs (Staphorst et al., 2016a). This comprehensive data integration facilitated by DF is crucial for decision-makers to understand the network's relevance and sustainability, aiding informed decisions regarding upgrades, service portfolio changes, cybersecurity measures, and resource allocations.

TF, which Sohn and Moon (2003) demonstrated can be achieved using SEM, adds a forward-looking component to the interdisciplinary toolkit, offering methods to predict future technological trends and their impacts on NRENs. Through traditional forecasting methods and advanced approaches like scenario planning and the Delphi technique, TF provides a systematic means to anticipate technological advancements and potential disruptions. This predictive insight proves invaluable for strategic planning within NRENs, ensuring they respond to current trends while preparing proactively for future technological integration. Applying TF through SEM-based DF could firmly ground these forecasting efforts, employing empirical technology and contextual data processed by robust analytical frameworks, thereby enhancing the accuracy and relevance of predictions for NREN strategic management.

The pioneering work of Staphorst et al. (2013, 2014, 2016a, 2016b) exemplifies the collaborative synergy between SEM, DF, and TF through the development of the TFSEMDF framework, tailored for application within the NREN technology domain. This framework leverages the strengths of each methodology to address the unique challenges faced by NRENs, offering a comprehensive framework to construct TF model instantiations that enhance current operational efficiency and strategically position NRENs for future challenges and technological shifts. Staphorst et al. (2013, 2014, 2016a, 2016b) have demonstrated through their research how the integrated application of SEM, DF, and TF can provide a robust mechanism for analysing, predicting and enhancing technological change at NRENs, thereby ensuring their relevance in supporting high-level research and education outcomes in a rapidly evolving technological landscape.

This synthesis highlighted the potential for integrating SEM, DF, and TF within the NREN technology domain while revealing complexity and opportunity areas. The following section builds upon these insights by critically analysing the existing literature considered in this study, identifying gaps, and deepening the understanding of how these methodologies can more effectively align with the multifaceted demands of the NREN technology domain.

2.6.2 Critical Analysis of SEM, DF, TF and NREN Literature

This section builds upon the previous synthesis by critically analysing the existing literature on SEM-based DF for TF within NRENs considered in Sections 2.2 to 2.5. The critical analysis reveals key insights alongside notable gaps. It addresses how researchers have applied these methodologies, where they fall short and suggests improvements to better serve NRENs' complex and evolving needs.

Integrating SEM, DF, and TF methodologies within NREN-related studies is a prominent shortfall in the existing literature. While SEM provides a robust statistical foundation capable of illuminating complex relationships between various NREN-related constructs, both technological and contextual, it often operates in isolation from DF and TF. This separation markedly limits the analytical depth achievable, particularly in scenarios characterised by the complex, multi-dimensional environments inherent to NRENs. The potential benefits of a holistic integration of these methodologies still need to be explored in academic research and practical applications, indicating a significant gap in the current corpus of knowledge. This lack of integrated application hinders the ability to leverage complete insights from multifaceted NREN indicator data, reducing the effectiveness of TF in strategic decision-making and operational optimisation within these networks (Staphorst et al., 2016a).

The failure to integrate these methodologies primarily arises from a traditional focus within each discipline that prioritises methodological rigour and purity over practical, interdisciplinary applications. Current frameworks predominantly applying SEM, DF, and TF in isolation often fail to capture modern NRENs' interconnected and intricate nature adequately. Although many of these frameworks handle data complexity with a degree of competence, they falter when tasked with forecasting future technological impacts while incorporating technological and contextual information. For example, while SEM adeptly models complex relationships within datasets, it does not natively support the real-time, continuous data integration that DF facilitates and that longitudinal TF requires. Similarly, while TF provides valuable foresight into potential technological change, it often lacks the structural and empirical grounding that SEM analyses offer. Consequently, the isolated application of these methodologies can result in a fragmented understanding of NRENs, failing to effectively provide a comprehensive framework that addresses their complex technological and contextual ecosystems.

Moreover, a critical evaluation of how existing frameworks handle the complexities of NRENs reveals that many frameworks inadequately adapt to the pace of technological advancement and contextual factors inherent in this domain. NRENs require frameworks that comprehend and analyse current technologies and anticipate and prepare for future developments, considering context-related information. Traditional TF methods, such as scenario planning and the Delphi technique, offer insights into forthcoming trends but typically remain disconnected from the operational and structural insights provided by DF and SEM, respectively. This disconnection can lead to strategic misalignments and inefficiencies in strategic and operational planning, emphasising the necessity for a more adaptive and integrated approach to managing technology change at NRENs. Existing frameworks must evolve to provide a more holistic, anticipatory, and context-informed foundation that can better support the complex, evolving needs of educational and research institutions relying on these networks (Staphorst et al., 2016a)

Steinberg's (2009) research on applying SEM in DF within natural language processing settings and Sohn and Moon's (2003) use of SEM in evaluating TF models illustrate preliminary efforts to integrate these methodologies. Steinberg's approach notably included context sensitivity in DF inferencing, leveraging SEM to model complex relationships within DF scenarios (Steinberg, 2009). Similarly, Sohn and Moon utilised SEM to articulate the hierarchical relationships crucial for TF in multi-layered data (Sohn & Moon, 2003). However, while these applications demonstrate the versatility of SEM in DF and TF, they do not extend to a comprehensive tripartite integration involving SEM, DF and TF. Furthermore, none of these studies conducted an empirical evaluation within the NREN technology domain, highlighting a significant gap in the literature where a complete integration of these methodologies could potentially enhance strategic and operational decision-making at these entities.

The future direction requires a paradigm shift toward more integrative and comprehensive frameworks that unify SEM and DF for TF within a singular, cohesive framework tailored to the unique demands of NRENs. Such frameworks must adeptly synthesise data from various sources and apply this integrated data to predict and navigate future technological shifts effectively. By doing so, NRENs can maintain robust, adaptable, and forward-looking operations, ensuring they adequately support the intricate and dynamic needs of the research and education sectors they serve. The innovative work by Staphorst et al. (2013, 2014, 2016a, 2016b) in developing the TFSEMDF framework represents a significant advancement in this direction, offering a compelling example and a potential template for future research and practical application in this critical area. This framework bridges the gaps identified and sets a precedent for future research to build upon, thereby providing a foundation from which model instantiations can be developed for the specific purpose of TF in the NREN domain, assisting these entities in evolving and adapting to an ever-changing technological landscape.

From the critical analysis presented in this section, thematic insights emerge on how integrating SEM, DF, and TF could produce methodologies that positively impact NRENs' strategic decision-making and operations. The thematic exploration in the next section highlights critical challenges and opportunities, setting the stage for proposing new frameworks for managing and forecasting technological trends in NRENs.

2.6.3 Thematic Insights from SEM, DF, TF and NREN Literature

This section explores thematic insights from integrating SEM, context-sensitive DF, and TF methodologies within the context of NRENs. These insights emerge from the existing literature considered in Section 2.2 to Section 2.5 and illuminate the challenges and opportunities intrinsic to the complex technology domain of NRENs. The subsequent themes highlight critical aspects of data complexity, modelling accuracy, predictive foresight, and methodological integration, contributing to a nuanced understanding of strategic decision-making and operational efficiency within NRENs.

The first prominent theme that emerges across the disciplines of SEM, DF, and TF in the context of NRENs is the challenge of data complexity. NRENs accumulate diverse data types, from quantitative network metrics to qualitative user feedback. Integrating these data sources, including technological and contextual data, exposes data quality and alignment issues, posing substantial challenges for precise modelling and insightful analysis. Staphorst et al. (2016a) have shown how the TFSEMDF framework can effectively tackle these issues by leveraging the strengths of each methodology to enhance data coherence and utility. This integration offers a robust approach to managing and analysing complex indicator data environments, which is essential for accurate forecasting in the NREN technology domain. This approach not only improves the reliability of the data but also ensures that comprehensive and precise data analysis informs decision-making processes at these entities (Staphorst et al., 2016a).

Another critical theme centres on the accuracy of modelling techniques used in the NREN technology domain. Accurate network capacity and capability predictions, service portfolio scope, and user interactions are pivotal for strategic decision-making at NRENs. SEM provides a substantial foundation for these analyses because it can uncover latent variables and model intricate relationships within complex systems. However, integrating SEM with DF to perform TF, as pioneered by Staphorst et al. (2013, 2014, 2016a, 2016b) in creating the TFSEMDF framework, enhances these techniques' predictive power and accuracy. This integration ensures a deeper understanding of the interactions between network elements, contextual drivers, and user behaviour, leading to more effective management and planning strategies. By employing SEM-based context-sensitive DF, which systematically combines data from multiple sources, including technology-related and context-related sources, to achieve more accurate and valuable information, the TFSEMDF framework provides predictive capabilities exceeding traditional TF methodologies (Staphorst et al., 2016a).

The theme of predictiveness in the literature is also significant, especially in how it impacts NREN strategic planning. Effective predictive models are crucial for NRENs, providing foresight into technology adoption trends and potential network expansions or constraints. The interdisciplinary approach of combining SEM, DF, and TF methodologies to create the TFSEMDF framework offers a unique advantage. This approach, both transversally and longitudinally, predicts technology needs and anticipates the challenges of new technology integration within existing educational and research infrastructures while incorporating the effects of contextual drivers in the complex NREN ecosystem. Doing so enables proactive management of resources and strategic planning, which is critical in maintaining and enhancing NRENs' service quality, thereby ensuring their continued relevance to the research and education communities. Staphorst et al. (2013, 2014, 2016a, 2016b) highlight the potential of TFSEMDF model instantiations to improve TF in the NREN technology domain, providing a tool to be more dynamic and responsive in approaching future technological challenges.

While Steinberg (2009) and Sohn and Moon (2003) provided early attempts at integrating SEM with DF and SEM with TF, respectively, the lack of broad integration of SEM, DF, and TF also emerges as a theme. It highlights a significant gap in existing frameworks, often treating these methodologies as isolated or loosely connected approaches. Staphorst et al. (2013, 2014, 2016a, 2016b) pioneered this area, offering a new perspective on synergistically applying these methodologies to enhance the understanding and management of NRENs. The development of the TFSEMDF framework demonstrates how integrated approaches can lead to better decision-making and more effective management of technology within NRENs. This integrated approach aids in the practical management of network resources. It fosters a deeper understanding of the theoretical facets of network operations, user behaviour, and technological change, promoting a more holistic view of NRENs' strategic and operational management (Staphorst et al., 2016a).

A focused exploration of future research opportunities is essential to leverage the thematic insights drawn from the existing literature considered in this study. The subsequent section delves into these potential areas of investigation emanating from the existing literature, particularly emphasising the integration of SEM, DF, and TF methodologies within complex technology domains such as NRENs.

2.6.4 Research Prospects from SEM, DF, TF and NREN Literature

The synthesis, critical analysis and theme extraction for existing literature reviewed in this study on SEM, DF, and TF in the context of NRENs unveils numerous avenues for future research, particularly in developing specific SEM constructs that integrate DF insights. Such integration would enhance the analytical capabilities of SEM by incorporating a more diverse array of data types, such as real-time network analytics and user feedback, which are crucial for comprehensively understanding NREN dynamics. The work of Staphorst et al. (2016a) provides a foundational framework for this integration, demonstrating how enriched data inputs can significantly refine model accuracy and predictive power. Future research could focus on operationalising these concepts by creating tailored SEM constructs that leverage DF processes to capture and model the complex interactions within NRENs more effectively, thus providing deeper insights into technology adoption patterns and network usage behaviours (Staphorst et al., 2016a).

Another promising area for future research involves developing DSEM models, specifically those capable of comprehensively providing longitudinal TF capabilities, that can adapt to new data from NREN operations. These models would be particularly beneficial in environments where network conditions and technology use continually change. By incorporating adaptive algorithms that update model parameters based on new data in real-time, these DSEM models could provide NREN managers with the tools to make more responsive and effective decisions. This adaptability is crucial for sustaining the relevance and accuracy of the models, ensuring they reflect the current state of the network and its users. Exploring these methodological innovations will push the boundaries of current research and significantly enhance the practical applications of SEM-based DF for TF within the complex ecosystem of NRENs (Staphorst et al., 2016a) and other domains with similar complexities.

Exploring interdisciplinary approaches that combine insights from SEM, DF and TF with other methodologies or domains, such as network theory or complex systems analysis, could also yield novel frameworks and applications that improve NREN strategic management. This cross-pollination of methodologies could foster innovative solutions tailored to the evolving challenges and opportunities within the NREN technology domain.

This section detailed potential avenues for future research emanating from the preceding integrative synthesis and critical analysis of the existing literature considered in Section 2.2 to Section 2.5. Possible areas for future research originating from this study's development, application and assessment of the TFSEMDF framework in the NREN technology domain (Staphorst et al., 2013, 2014, 2016a, 2016b) are presented in Section 7.5

2.7 CONCLUDING REMARKS

This literature and theory review presented in this chapter covered an array of critical theoretical domains central to this research. The examination of SEM began with a historical perspective, delving into the evolution of the methodology and its pivotal concepts. The discussion then progressed to a granular level with the mathematical foundations of SEM, providing the groundwork for understanding its analytical capabilities and its subsequent application in the diverse fields of DF and TF.

The chapter methodically explored the field of DF, charting its historical development and elaborating on the foundational concepts that define this intricate field. It analysed the JDL/DFG, detailing its structural composition and operational application within DF. Furthermore, the chapter assessed various DF methodologies, highlighting the critical role that context-related information plays in enhancing the accuracy and efficacy of DF practices.

The section on TF offered a chronological overview, charting the field's inception and maturation over time. It sheds light on the essential technology indicators and TF output metrics instrumental in forecasting practices. An extensive review of TF methodologies, including TFDEA, was presented, revealing the field's expansive nature and the nuances of its research frontiers. The examination of TF methodologies underscored their strategic value in predicting future technological trends. It informed the discussion on measuring success in the field, which is pivotal for evaluating the efficacy of forecasting models and approaches.

The comprehensive study of NRENs underscored their historical significance and pivotal role in the research, education, and innovation landscapes. This section then explored the RREN and NREN ecosystems in Africa. The review concluded with an overview of the TERENA and GÉANT NREN Compendiums, which serve as extensive knowledge repositories, providing insight into European NRENs' technological, operational, and strategic attributes.

The analysis and synthesis of the existing literature on SEM, DF, and TF within the NREN technology domain provided an understanding of how these methodologies can enhance strategic decision-making and operational efficiency. This section began with a preliminary synthesis of key concepts and a critical assessment of existing frameworks, thematic insights, and future research directions. Together, these components offered valuable insights into the challenges and opportunities within the complex NREN ecosystem, laying the groundwork for developing robust frameworks tailored to the evolving demands of these networks.

This chapter established a platform for developing, applying, and assessing the TFSEMDF framework within the context of NRENs. It serves not only as a testament to the breadth of research conducted but also as a crucial reference point for the subsequent analysis and discussions that formed the core of this study.

CHAPTER 3 – RESEARCH PHASES AND OBJECTIVES

3.1 INTRODUCTION

The study unfolded in three methodically structured phases within the technology domain of NRENs, each addressing a distinct aspect of the TFSEMDF framework, i.e., framework development, application, and assessment. In its first phase, expanded from Staphorst et al. (2013), the study attempted the integration of SEM and DF for TF, culminating in the TFSEMDF framework's development. This phase focused on the innovative integration of various data sources, both technology-related and context-related, employing context-sensitive DF techniques to enhance TF's reliability and depth.

The second phase, partly previously presented in Staphorst et al. (2014) and Staphorst et al. (2016a), involved the practical application and refinement of the TFSEMDF framework. It commenced with implementing an autoregressive model instantiation for the NREN technology domain, utilising insights from action research within SANReN (Gustavsson, 2008; SANReN, n.d.) and data from TERENA's NREN Compendiums. The subsequent effort further developed cross-sectional NREN model instantiation, integrating additional findings and theoretical inputs, enhancing the framework's scope and efficacy.

Initially conceptualised by Staphorst et al. (2016b), the third phase marked the study's culmination with a comprehensive evaluation of the TFSEMDF framework's strengths and weaknesses. This critical examination provided insight into the framework's capacity to integrate context-related information effectively and identified areas of potential inaccuracy in structural model specification.

The following chapter concisely presents this research progression, outlining each phase's objectives, approaches, and scholarly publications while highlighting the subsequent research efforts that followed these earlier publications. This structured overview offers a cohesive narrative of the study's evolution and academic contributions to the field of TF within the NREN technology domain.

3.2 Research Roadmap

This study methodically progressed through the research roadmap depicted in Figure 2, beginning with integrating generic frameworks for SEM, context-sensitive DF and technology indicator relational mapping for TF, thereby creating the TFSEMDF framework.

Figure 2: Research Roadmap Depicting Phases with Objectives and Outputs

The research then transitioned to applying this framework for longitudinal and transversal forecasting in NRENs' technologically advanced and contextually nuanced domain. Finally, the study focused on a critical evaluation of the framework's strengths and weaknesses. These three distinct research phases each included specific objectives and associated outcomes, as depicted in Figure 2. By organising the research into these phases, the study addressed each aspect comprehensively, from conceptual development to empirical application and critical evaluation.

3.3 RESEARCH PHASES, OBJECTIVES AND OUTPUTS

The following sections provide a comprehensive overview of the focus and research objectives characterising each distinct phase of the study. These sections also pinpoint the specific publications by Staphorst et al. (2013, 2014, 2016a, 2016b) that emerged as outputs corresponding to each research objective during the study. Additionally, these sections highlight the alignment of the research phases with the chapters in the remainder of this thesis. This alignment aims to offer a coherent and systematic narrative of the research trajectory undertaken and the scholarly contributions that ensued from it.

3.3.1 RESEARCH PHASE ONE: TFSEMDF FRAMEWORK DEVELOPMENT

The first phase of the study, focusing on developing the TFSEMDF framework, was anchored by two key research objectives as outlined in Figure 5: Research Objective 1(a) and Research Objective 1(b). Research Objective 1(a) entailed a comprehensive review of the theoretical underpinnings of SEM, DF and TF. The aim was to abstract essential elements and relationships from these methodologies, thereby creating refined generic framework extractions that capture the core principles of each.

Research Objective 1(b) progressed to examining the intersections among these abstracted generic frameworks. The goal was to synergise generic frameworks for SEM, context-sensitive DF and technology indicator relational mapping for TF into a cohesive single framework for SEM-based context-sensitive DF for TF. It involved synthesising critical elements from each methodology, culminating in forming the integrated TFSEMDF framework, which embodies the combined strengths and functionalities of SEM, context-sensitive DF and technology indicator relational mapping for TF.

The methodology employed in this phase, along with the results obtained and the subsequent discussion of these findings, was initially published in Staphorst et al. (2013), then improved in Staphorst et al. (2014, 2016a) and finally refined in this study. Chapter 4 of the thesis centres on the TFSEMDF framework development phase of the study. Furthermore, Section 7.2.1 comprehensively covers conclusions from the framework development effort.

3.3.2 Research Phase Two: TFSEMDF Framework Application

The study's second phase aimed to apply the TFSEMDF framework practically within the NREN technology domain. This phase focused on two key research objectives.

The first objective for this phase, Research Objective 2(a) in Figure 5, was to apply the TFSEMDF framework using an autoregressive NREN model instantiation for longitudinal forecasting. The study involved creating the autoregressive NREN model instantiation based on insights gleaned from action research in SANReN (Gustavsen, 2008; SANReN, n.d.). This step was followed by extracting technology indicators and context-related data from the 2011 and 2012 TERENA NREN Compendiums (TERENA, 2011, 2012) to conduct regression analysis of the model instantiation using PLS-SEM, as per Appendix A's expansion of the process initially defined in Staphorst et al. (2015). The process also included verifying the reliability and validity of the PLS-SEM results, as outlined in Appendix B's expansion of the approach initially presented in Staphorst et al. (2017.

The subsequent objective, Research Objective 2(b) in Figure 5, focused on applying the TFSEMDF framework to develop and evaluate a cross-sectional NREN model instantiation for transversal forecasting. Achieving this objective involved constructing the cross-sectional NREN model instantiation drawing on literature, followed by extracting relevant data from TERENA Compendiums to perform regression analysis using PLS-SEM according to Appendix A's expansion of the process initially defined in Staphorst et al. (2015) and then confirming the reliability and validity of these results using Appendix B's expansion of the approach described in Staphorst et al. (2017).

Chapter 5 of the thesis considers these two research objectives' methodology, results, and discussion. This phase's findings were initially published in Staphorst et al. (2014), refined in Staphorst et al. (2016), and subsequently employed in the study examining the strengths and weaknesses of the TFSEMDF framework presented in Staphorst et al. (2016b). Section 7.2.2 provides concluding remarks on these results, offering insights into the practical applications and efficacy of the TFSEMDF framework within the NREN technology domain.

3.3.3 Research Phase Three: TFSEMDF Framework Assessment

The third phase of this study entailed a comprehensive evaluation of the inherent strengths and weaknesses of the TFSEMDF framework, empirically tested within the NREN technology domain. This evaluative phase addressed two specific research objectives.

The first objective, Research Objective 3(a) in Figure 5, was devoted to an in-depth exploration of the intrinsic strengths of the TFSEMDF framework. Initially, the study engaged in a detailed theoretical analysis to identify the core strengths inherent in the TFSEMDF framework's constituent methodologies, i.e. SEM, DF and TF. Following this theoretical groundwork, the study applied phase two’s cross-sectional NREN model instantiation and a structurally disarranged NREN model instantiation as empirical instruments. This analysis aimed to validate the theoretical strengths through practical application and observation within the context of the NREN technology domain.

Research Objective 3(b), as depicted in Figure 5, focused on uncovering the inherent weaknesses of the TFSEMDF framework. This effort involved a theoretical investigation into characteristics of the constituent methodologies, i.e. SEM, DF and TF, of the TFSEMDF framework that might present limitations or challenges. An empirical assessment used phase two’s cross-sectional NREN model instantiation and structurally disarranged NREN model instantiation to evaluate the identified TFSEMDF weaknesses within the context of the NREN technology domain.

Chapter 6 of the thesis presents the methodology, results, and discussion surrounding these two pivotal research objectives. The findings from this phase, initially documented in Staphorst et al. (2016b), offer a thorough critique and assessment of the TFSEMDF framework's strengths and weaknesses in the NREN technology domain. Section 7.2.3 presents conclusions from examining the TFSEMDF framework's strengths and weaknesses.

3.4 CONCLUDING REMARKS

Chapter 3 of this thesis provided a brief overview of the study's design, organised into three distinct phases, each addressing a specific component of the TFSEMDF framework within the NREN technology domain. The chapter detailed the approach for each phase, clearly outlining the steps and methods used to achieve the research objectives.

The chapter provides a structured roadmap for the study, detailing the approach and execution captured in this thesis. This structure enhances the ease of navigating through the thesis and aids in comprehending the logical progression of the research. Chapter 3 also highlights the academic contributions of each phase, serving as a reference to the tangible outcomes of the study.

CHAPTER 4 –RESEARCH PHASE 1: TFSEMDF FRAMEWORK DEVELOPMENT

4.1 INTRODUCTION

This chapter delves into the foundational phase of the study, detailing the formulation of the TFSEMDF framework by addressing the foundational Research Objectives 1(a) and 1(b). Objective 1(a) focussed on extracting generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF. Concurrently, Objective 1(b) undertook the intricate process of integrating these generic frameworks into the proposed TFSEMDF framework.

The chapter details the research methods employed to achieve Research Objectives 1(a) and 1(b) defined for the first phase. It then presents the findings for each objective. Finally, the chapter critically evaluates these findings, assessing their impact and relevance in supporting the research objectives. This analysis aims to determine the research methods' efficacy and the conclusions' reliability.

The contents of this chapter are the culmination of an iterative process of refinement, encapsulating the enhancements and improvements of the concepts and methodologies presented by Staphorst et al. (2013, 2014, 2016a). The methods, results, and critical discussions presented herein include significant expansions on the work presented in these foundational publications, reflecting the progressive development of the TFSEMDF framework through evolving insight and augmented through scholarly underpinnings.

4.2 RESEARCH METHODOLOGY

The forthcoming sections provide an in-depth overview of the research methodologies used in this study phase. Section 4.2.1 presents a detailed exploration of the research methodology applied in developing generic frameworks for SEM, DF, and technology indicator relational mapping for TF, which aligns with Research Objective 1(a). Section 4.2.2 discusses the research methodology used to integrate these frameworks into the comprehensive TFSEMDF framework, thereby addressing Research Objective 1(b).

4.2.1 Generic Framework Abstractions for SEM, DF and TF

The methodology employed to achieve Research Objective 1(a), which focused on developing generic frameworks for SEM, DF and technology indicator relational mapping for TF, encompassed a structured two-step process. First, the study conducted an expansive literature review to explore the fields of SEM, DF, and TF, as presented in Chapter 2. This step included thematic analysis (Braun & Clarke, 2006), which was instrumental in delving into seminal theories, methodologies, and applications within these fields. The thematic analysis allowed for a systematic identification and interpretation of patterns and themes in the literature (Ryan & Bernard, 2003). This process then entailed abstracting generic frameworks based on the body of literature presented in Chapter 2, utilising inductive reasoning to synthesise and distil the core principles of SEM, DF and TF into comprehensive generic frameworks.

The first step of the research methodology for Research Objective 1(a) involved a comprehensive literature review that probed the theoretical foundations and practical applications of SEM, DF, and TF, as documented in Sections 2.2, 2.3, and 2.4, respectively. It entailed an analytical appraisal of the seminal and more recent scholarly contributions that form the conceptual bedrock of these fields. The review included a critical exploration of the mathematical foundations of SEM (Staphorst, 2010), as described in Section 2.2.2. Section 2.2.2's examination of the construction and notation of SEM path diagrams, as introduced by Chin and Newsted (1999) and expanded upon by Staphorst (2010), was also pivotal. Further, the literature illuminated the application of SEM in the domain of TF, as demonstrated by Sohn and Moon (2003) in Section 2.2.4, and its capacity to facilitate context-sensitive DF (Steinberg, 2009; Steinberg & Rogova, 2008), as discussed in sections 2.2.4 and 2.3.4. An in-depth review was also undertaken into the multi-levelled JDL/DFG framework as detailed by Steinberg and Bowman (2017) in Section 2.3.2, as well as Grupp's (1998) segmentation of TF technology indicators into input, byput, and output indicators, as outlined in Section 2.4.2.

In the second step of the research methodology for Research Objective 1(a), the study applied inductive reasoning to distil and synthesise generic frameworks for SEM, DF, and technology indicator relational mapping for TF, informed by the insights from the literature reviewed in the first step. It also employed conceptual framework development (Miles & Huberman, 1994) and grounded theory (Charmaz, 2006) methodologies to construct generic frameworks for SEM, DF and technology indicator relational mapping for TF that encapsulated the critical theoretical and practical elements of the respective fields, while also capturing the inherent structures typical of models created within these domains. During this framework generalisation process, the guiding principle was to identify structural similarities and opportunities for harmonisation across SEM, DF and TF, thereby setting a strategic path toward integrating these frameworks into a cohesive TFSEMDF model.

4.2.2 Merging Frameworks for SEM, DF and TF to Create TFSEMDF

The methodology for achieving Research Objective 1(b) focused on integrating the previously developed generic frameworks for SEM, DF, and TF into a cohesive framework, TFSEMDF. This objective was approached through a two-step process, reflecting analytical rigour and methodological precision. The first step involved a detailed comparative analysis (Glasser, 1965; Smith, 2015) of the existing frameworks to identify overlapping concepts and complementary elements. This analysis played a critical role in understanding how to combine these frameworks effectively. The second step revolved around the actual synthesis of the TFSEMDF framework, which entailed carefully merging the identified commonalities and harmonising various aspects of SEM, DF, and TF into a unified model. This process included the standardisation of terminologies and notations to ensure consistency and clarity in the newly developed framework, ultimately creating a tool that effectively combines the strengths of each methodology for comprehensive SEM-based DF in the context of TF.

The first step in Research Objective 1(b)'s methodology was the detailed comparative analysis (Glasser, 1965; Smith, 2015) of the generic frameworks for SEM, DF and TF. Central to this step was the focus on cross-disciplinary integration, which closely examined each framework's methodologies and theoretical constructs. The comparative analysis, a method well-established in the literature (Glasser, 1965; Smith, 2015), was conducted to identify where these frameworks intersect and overlap, setting the stage for their integration. By drawing on the principles and methodologies inherent in SEM, DF, and TF, this step aimed to highlight synergistic potentials and conceptual commonalities. The approach of cross-disciplinary integration, as advocated by Johnson and Onwuegbuzie (2004), further enriched this analysis. These findings then guided the development of the TFSEMDF framework, ensuring that subsequent integration builds on a comprehensive understanding of how to combine these frameworks effectively.

In the second step of the methodology for Research Objective 1(b), the focus shifted to synthesising the TFSEMDF framework, leveraging the insights gleaned from the comparative and cross-disciplinary analysis conducted in the first step. This phase involved a sophisticated process of framework unification, integrating the identified commonalities and complementary elements from SEM, DF, and technology indicator relational mapping for TF. For guidance on this integration process, the methodology drew upon the principles of framework unification as discussed in the works on integrative research methods, such as those by Mingers and Brocklesby (1997), who explore the amalgamation of disparate methodologies in complex systems. The process involved aligning theoretical concepts and harmonising terminologies and notations, ensuring a seamless and coherent structure for the TFSEMDF framework. The resultant TFSEMDF framework emerged as a comprehensive, multifaceted tool adept at addressing the complex requirements of SEM-based DF in the context of TF and reflecting the synergistic potential of integrating diverse methodological perspectives.

4.3 RESULTS

The subsequent sections focus on the results obtained during this study phase. Section 4.3.1 presents the results from developing generic frameworks for SEM, DF, and technology indicator relational mapping for TF, in line with Research Objective 1(a). Section 4.3.2 then details the results of merging these frameworks into the generic TFSEMDF framework, achieving the goals for Research Objective 1(b).

4.3.1 GENERIC FRAMEWORK ABSTRACTIONS FOR SEM, DF AND TF

4.3.1.1 Generic Framework for SEM

Section 2.2 outlines the extensive literature review and analytical appraisal of SEM as part of this research phase. A key emergent theme from this analytical appraisal was the critical distinction of variables emphasised into exogenous and endogenous latent constructs that underpin SEM theory. SEM's theoretical architecture extends beyond the traditional dichotomy of dependent and independent variables typical in regression analysis (Haenlein & Kaplan, 2004). It explicitly focuses on exogenous constructs as independent variables, which do not depend on any relational dynamics within the model. It also considers endogenous constructs, which serve as dependent or independent variables based on their interactions with other variables in the model (Staphorst, 2010).

As outlined in the literature review of Section 2.2, SEM also caters for indicators that act as proxies to represent latent constructs (Haenlein & Kaplan, 2004). These indicators fall into either reflective or formative categories. Reflective indicators typically show high correlations with the latent construct and other reflective indicators, effectively representing the variance in the unobserved variable and implying a causal direction from the latent construct to its indicators. In contrast, latent constructs with formative indicators consist of a weighted combination of indicators that do not necessarily exhibit high correlations with the latent construct or each other. Such formative indicators capture different dimensions of the latent construct, indicating a causal influence from the indicators to the construct (Haenlein & Kaplan, 2004).

Figure 1 in Section 2.2.2 presents the archetypal configuration of an SEM path diagram, a prevalent tool for illustrating SEM models. Such diagrams characteristically feature a constellation of exogenous and endogenous constructs, interconnected by path coefficients, and are further defined by associated reflective and formative indicators, each with respective loadings (Staphorst et al., 2013). Upon analysis, one can discern that researchers can arbitrarily group the indicators and constructs within SEM path diagrams to align with their specific investigative objectives. Constructs, for instance, might be grouped according to shared functional characteristics or by their locational pertinence in the analysed system (Staphorst et al., 2013).

As outlined in the research methodology defined to achieve Research Objective 1(a), conceptual framework development (Miles & Huberman, 1994) to create a generic SEM framework employed a condensed version of grounded theory (Charmaz, 2006). This iterative process involved determining sensible groupings for the constituent indicators, constructs, and interconnections within SEM path diagrams. Starting without any predefined generic framework for SEM in mind, the study reviewed several arbitrarily selected example SEM path diagrams from literature (Vinzi et al., 2010) to encode collections of indicators and constructs into groupings that could potentially match critical structures in the generic frameworks for context-sensitive DF and technology indicator relational mapping for TF.

As the generic frameworks for context-sensitive DF and technology indicator relational mapping for TF evolved (see Sections 4.3.1.2 and 4.3.1.3, respectively), the generic SEM framework's chosen grouping approach was adapted and compared against the evolving structures of the generic frameworks for DF and TF. The culmination of this iterative process was a simplistic grouping approach that entails grouping exogenous and endogenous constructs into context and technology-related constructs with their associated indicators, respectively. Furthermore, the analysis determined that these groupings need to be stratified such that context and technology-related constructs of akin nature or complexity fall in the same tier (or level, using DF terminology), in line with the layering approach of the JDL/DFG framework for DF, detailed in Section 2.3.2. Applying this methodical grouping and stratification process yielded the generalised SEM framework illustrated in Figure 3.

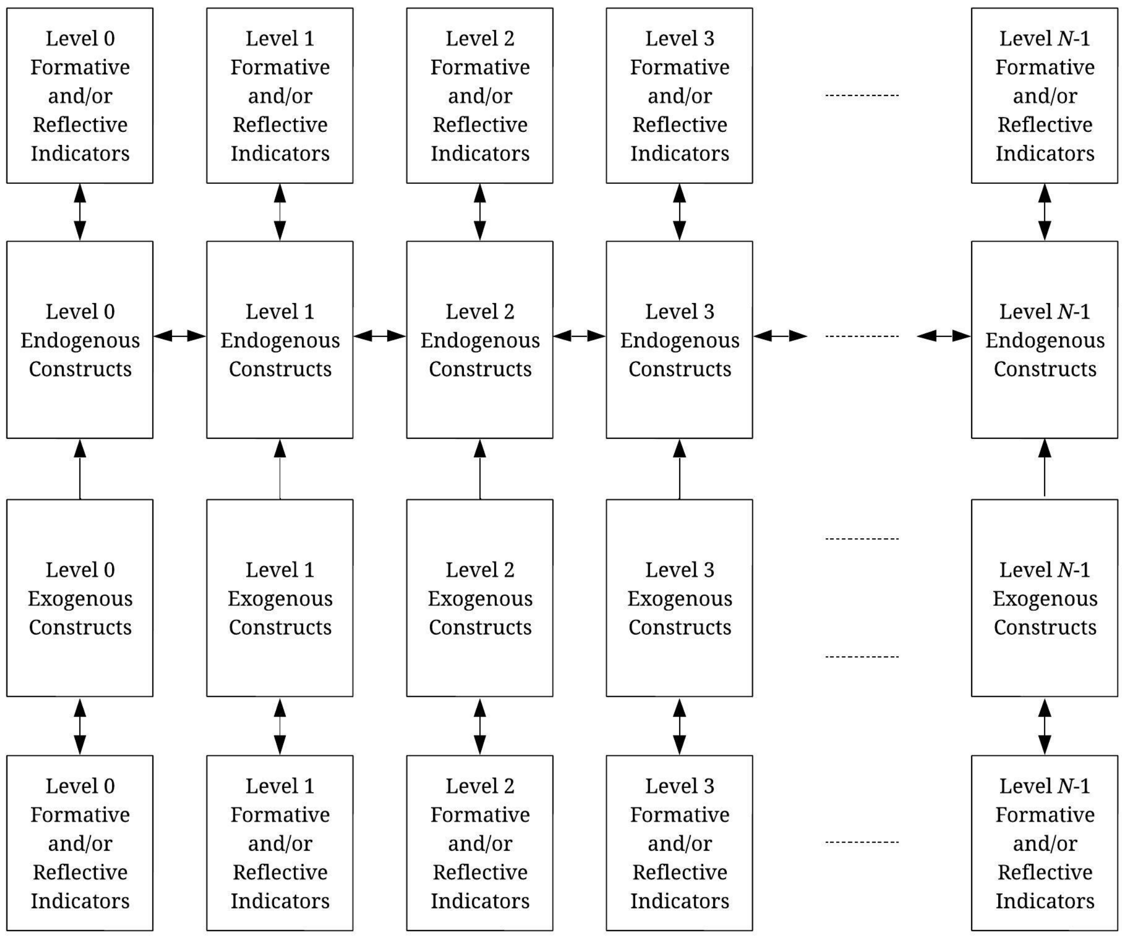


Figure 3: Generic Framework for SEM

As shown in Figure 3, the generic SEM framework comprises N stratified layers, each encompassing an aggregate of constructs and indicators. Within this schema, N is not a fixed quantity but a variable integer, indicative of the requisite number of stratification layers (or levels as per DF terminology), which is contingent upon the specifications set forth by the framework user. Importantly, there is a correspondence between the N levels of this SEM framework and the N

strata within the generic context-sensitive DF framework, as depicted in Figure 4. This parallel extends to the generic framework employed for the technology indicator relational mapping for TF, showcased in Figure 5, and is further integrated within the synthesised TFSEMDF framework (Figure 6). This alignment underscores a systematic and interdisciplinary approach to structuring these frameworks.

4.3.1.2 Generic Framework for Context-Sensitive DF

Following the research methodology defined for Research Objective 1(a), an extensive literature review was undertaken on DF, as presented in Section 2.3. From this review, the central theme that emerged is that DF is conceptually a stratified framework designed for the multi-layered refinement and integration of estimates of problem variables derived from various measurements, either directly or indirectly observable (Steinberg, 2009; Steinberg & Rogova, 2008). Within a military context, as recognised by the JDL/DFG, the process of DF involves systematically aggregating and enhancing sensor data to generate high-quality tactical knowledge (Steinberg, 2009; Steinberg & Rogova, 2008).

In congruence with the aim of this research objective to perform conceptual framework development (Miles & Huberman, 1994) to create a generic framework for context-sensitive DF, Section 2.3.2 critically evaluated the JDL/DFG's efforts to standardise the structure of the DF process across diverse military applications. This effort culminated in the JDL/DFG defining six distinct DF processing levels, as outlined in Section 2.3.2. To reiterate, the structured levels of DF analysis comprise the following: Level 0, which is concerned with signal, feature, and subject assessment; Level 1, with object assessment; Level 2, with situation assessment; Level 3, with impact assessment; Level 4, with process refinement; and Level 5, with user refinement.

Considering its established status, straightforwardness, and broad adoption, this study modified the foundational JDL/DFG framework to formulate a generic framework for context-sensitive DF. Although the specific DF level definitions articulated by the JDL/DFG are predominantly tailored to military applications and do not seamlessly transition to the use of DF in TF, the essential concept of progressive aggregation and refinement of measurement indicator data is integral to the TFSEMDF framework developed in this study. Consequently, the level definitions within the JDL/DFG framework have been abstracted, culminating in this study's generic framework for context-sensitive DF, depicted in Figure 4.

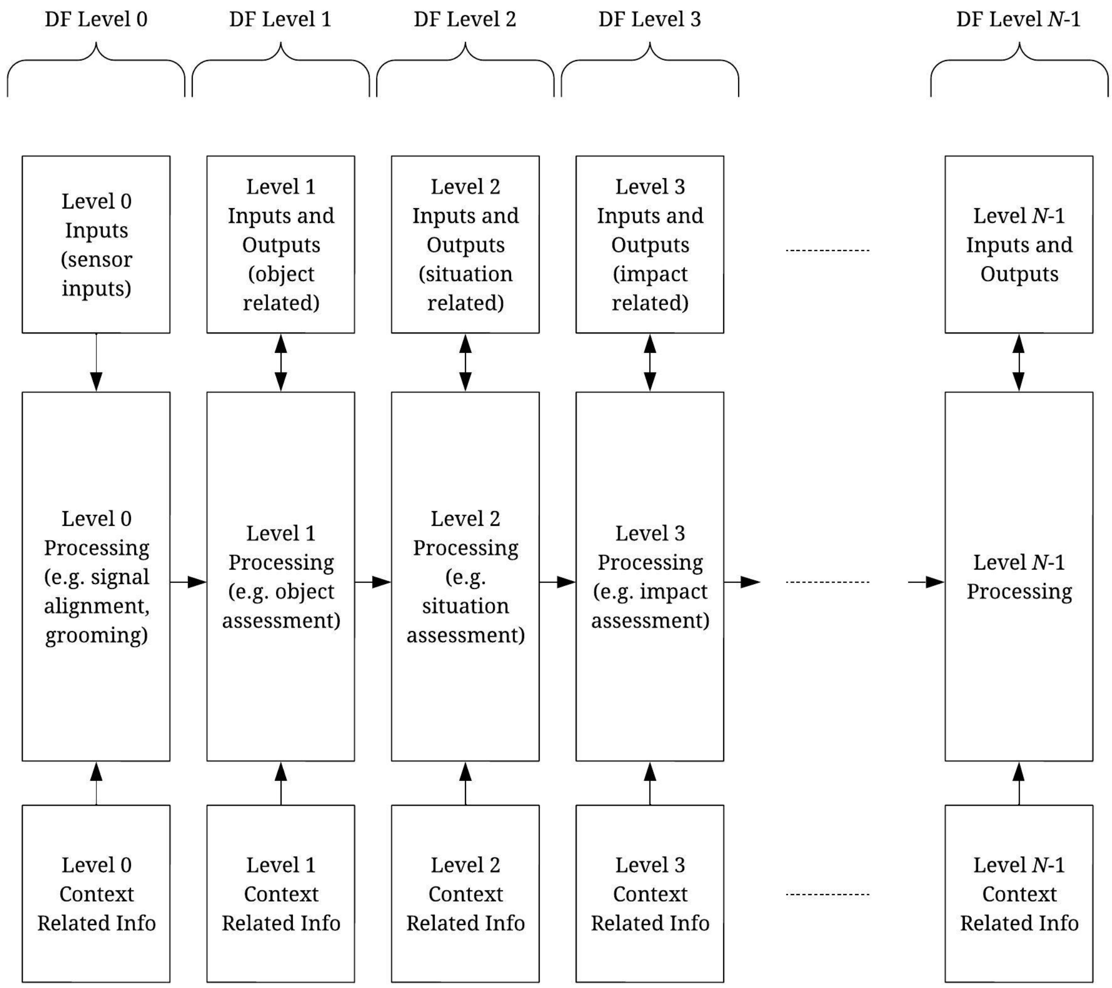


Figure 4: Generic Framework for Context-Sensitive DF

As evident from Figure 4, the generic context-sensitive DF framework's hierarchical structure comprises N levels of aggregation. Here, N represents a variable integer, signifying the number of tiers necessary for effective aggregation and refinement, tailored according to the requirements of the framework's user. This flexible design allows for incorporating relevant context-related information at any specified level within the DF, ensuring flexibility to a wide range of technology scenarios and operational needs. Notably, the N levels in this DF framework are analogous to those in the SEM framework, the generic framework for technology indicator relational mapping for TF and the integrated TFSEMDF framework, as depicted in Figure 3, Figure 5 and Figure 6, respectively.

4.3.1.3 Generic Framework for Technology Indicator Relational Mapping for TF

Section 2.4.2's literature review, part of the research methodology employed to achieve Research Objective 1(a), presents a literature study of frameworks for the relational mapping of technology indicators in TF. Notably, the Watts and Porter (1997) framework described three principal categories of technology indicator metrics: Technology Life Cycle Status Indicators, which gauge the maturity and growth rate of a technology; Innovation Context Receptivity Indicators, which assess the technological ecosystem's readiness; and Market Prospect Indicators, which evaluate the commercial potential and market positioning of the technology (Nyberg & Palmgren, 2011; Watts & Porter, 1997). Conversely, Grupp (1998) categorises these indicators into three main types based on their functional role: Input indicators, which relate to the drivers of technological progress; Byput indicators, which are associated with the intermediary phenomena of technological evolution; and Output indicators, which pertain to the qualitative, quantitative, or value-assessed progress in process or product development.

After critically assessing the technology indicator frameworks presented in Section 2.4.2, this study selected the framework from Grupp (1998) as the basis of the study's conceptual framework development (Miles & Huberman, 1994) effort, which culminated in the general TF technology indicator relational mapping framework depicted in Figure 5. This generic framework categorises input, byput, and output technology indicators based on shared characteristics or complexity levels. Additionally, it highlights the existence of intricate interconnections among these three types of indicators.

For the scope of this study, input technology indicators are pivotal in driving technology processes at Level 0. Byput technology indicators, on the other hand, are crucial in signalling subphenomena that occur between Level x-1 and Level x, where x ranges from 1 to N-1. In this framework, N serves as a flexible integer selected by the framework's user to align with the specific characteristics of the investigated TF scenario. Crucially, the N levels in this generic framework parallel those in the generic SEM framework illustrated in Figure 3, the generic context-sensitive DF outlined in Figure 4, and the synthesised TFSEMDF framework depicted in Figure 6. This alignment underscores a coherent and systematic approach across these frameworks, facilitating a unified methodology in their application to various analytical contexts within TF.

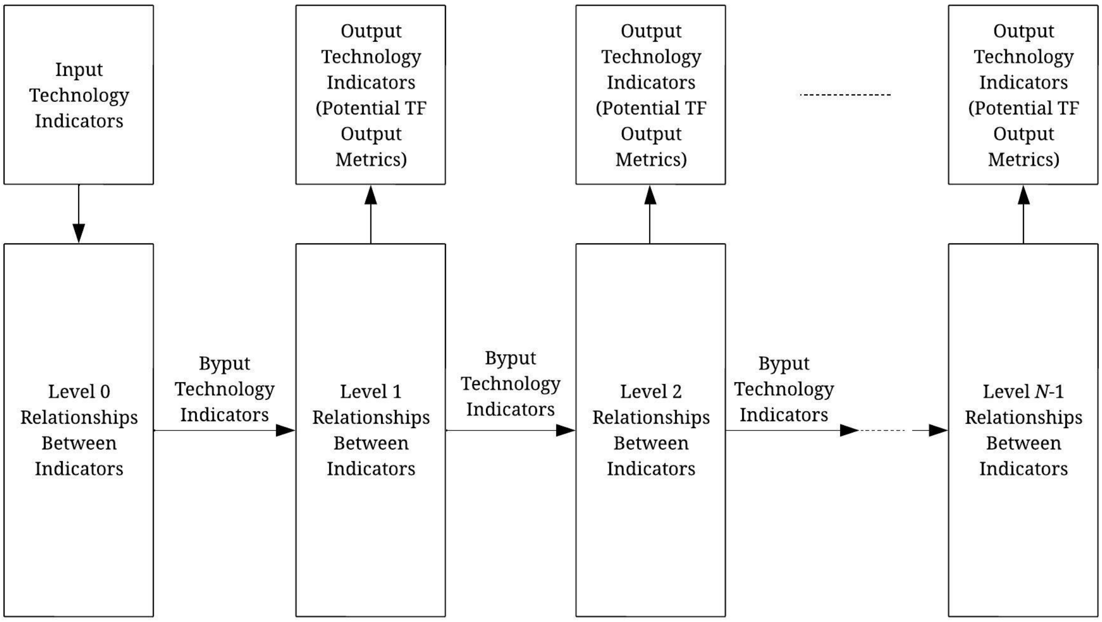


Figure 5: Generic Framework for Technology Indicator Relational Mapping for TF

Lastly, output technology indicators are instrumental in reflecting the progress related to products or processes at each respective Level x, for x values spanning from 1 to N-1. This conceptual framework underpins the analysis and interpretation of technology indicators within the context of this study, providing a structured approach to understanding their roles and interactions across different levels of technological development.

4.3.2 Merging Frameworks for SEM, DF and TF to Create TFSEMDF

Employing both comparative analysis (Glasser, 1965; Smith, 2015) and cross-disciplinary integration (Johnson & Onwuegbuzie, 2004), tools forming part of the research methodology for Research Objective 1(b), Steinberg and Rogova (2008) and Steinberg (2009) demonstrated that SEM is particularly well-suited for implementing DF, as it can simultaneously model multiple dependent and independent constructs. As outlined in Section 2.2.4, SEM facilitates the incorporation of context-sensitivity into DF inferencing problems, where "context" is defined as a set of relationships, each representing a specific instantiated relation. This process is crucial in refining estimates, interpreting data, and constraining processing during DF data acquisition, cueing, or fusion. Furthermore, Steinberg and Rogova (2008) and Steinberg (2009) align DF terminology with SEM, identifying DF problem variables as SEM endogenous constructs, context variables as SEM exogenous constructs, and traditional DF sensor measurements as reflective and formative indicators within SEM. This conceptual alignment underscores the compatibility and potential of SEM in enhancing the analytical depth and accuracy of DF analysis by integrating complex variable relationships and context-related factors.

Sohn and Moon (2003) effectively showcased SEM as a robust regression technique capable of assessing multi-layered hierarchical models, with this methodology comprehensively detailed in Section 2.2.4. By applying comparative analysis methodologies (Glasser, 1965; Smith, 2015) and harnessing cross-disciplinary perspectives (Johnson & Onwuegbuzie, 2004), their research not only highlighted the utility of SEM in the systematic aggregation and refinement of input technology indicator data but also confirmed its capacity to yield a statistically reliable estimate of the Technology Commercialization Success Index (TCSI) metric, crucial for TF applications.

Sohn and Moon (2003) demonstrated the efficacy of SEM as a regression technique for evaluating multi-layered hierarchical models, as discussed in Section 2.2.4. Their research highlighted the utility of SEM in the systematic aggregation and refinement of input technology indicator data. It confirmed its capacity to yield a statistically reliable estimate of the TCSI metric, which is crucial for TF applications. Furthermore, their approach underlines the adaptability of SEM in bridging diverse data sets and theoretical constructs, thereby enhancing the analytical depth and scope of research in TF contexts.

Building upon the application of SEM for TF by Sohn and Moon (2003) and extending it through the context-sensitive DF approach of Steinberg and Rogova (2008) and Steinberg (2009), this study introduces the novel TFSEMDF framework depicted in Figure 6 for SEM-based context-sensitive DF in TF (Staphorst et al., 2013, 2014, 2016a, 2016b). This framework was developed through cross-disciplinary integration (Johnson & Onwuegbuzie, 2004) and framework unification (Mingers & Brocklesby, 1997). It manifests by integrating the generic technology indicator relational mapping framework for TF from Figure 5, based on Grupp's (1998) definitions, with the general context-sensitive DF framework presented in Figure 4 while applying the SEM construct grouping and layering approach from the generic SEM framework in Figure 3. TFSEMDF provides a comprehensive and structured approach to SEM-based context-sensitive DF for TF.

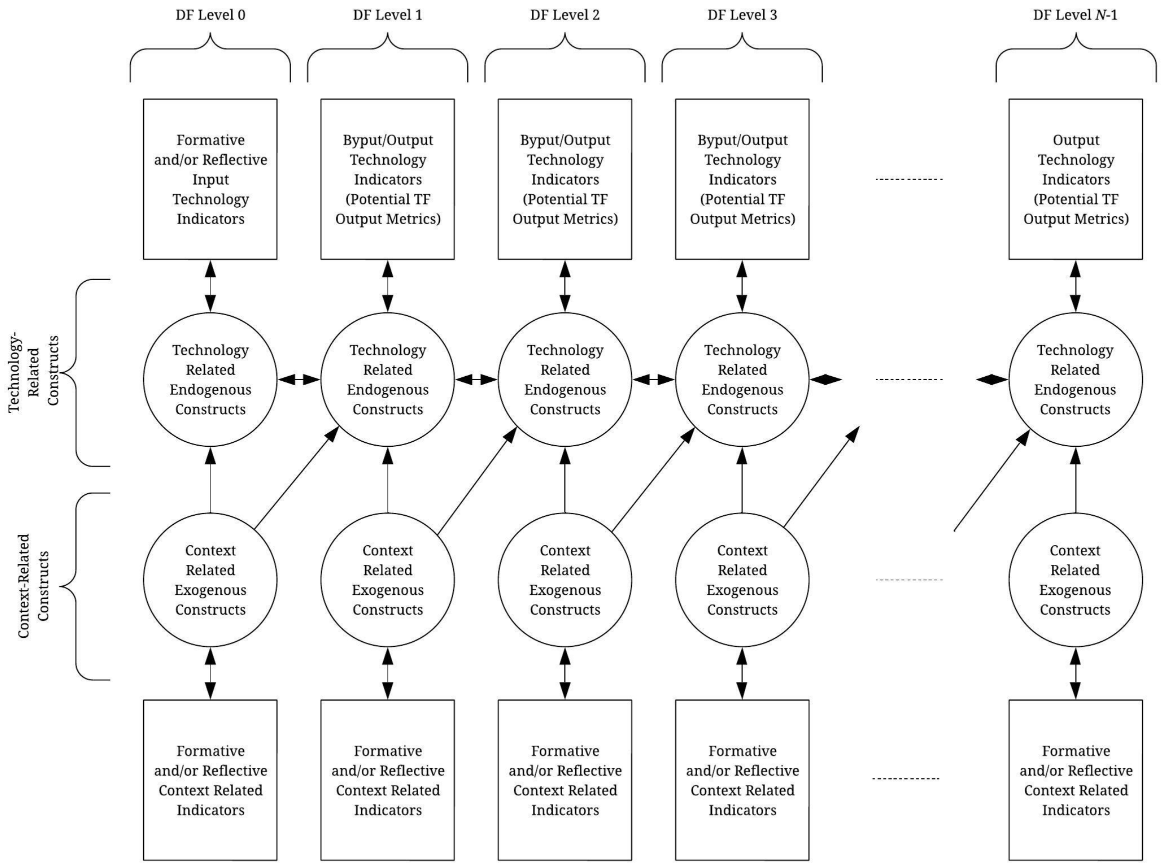


Figure 6: TFSEMDF Framework

The proposed TFSEMDF framework performs multi-layered aggregation and refinement of technology and context-related information, executed through SEM-based DF processing at DF Levels 0 to *N*-1, where the framework user selects *N*. The determination of the number of levels, *N*, which is analogous to the number of levels in the generic SEM framework, the generic context-sensitive DF framework and the generic technology indicator relational mapping framework for TF, depends on the intricacy of the examined technology domain and constraints related to time and budget. Additionally, considering potential diminishing returns arising from further aggregation and refinement plays a pivotal role in defining the optimal number of levels *N*, as outlined by Staphorst et al. (2013, 2014, 2016a, 2016b). This approach ensures a balanced evaluation of the depth of analysis against practical limitations and diminishing marginal benefits, thereby guiding the establishment of an appropriate and adequate level of data processing within the TFSEMDF framework.

In the TFSEMDF framework, input technology indicators (Grupp, 1998; Nyberg & Palmgren, 2011) and context-related indicators (Steinberg, 2009; Steinberg & Rogova, 2008) serve as inputs to technology-related endogenous constructs and context-related exogenous constructs, respectively. The framework utilises bi-directional interconnections between indicators and constructs, as well as among multiple constructs, adhering to the SEM path diagram conventions described in Staphorst (2010) and presented in Section 2.2.2. This structure represents positive and negative correlations between constructs and acknowledges that indicators can be reflective or formative.

The TFSEMDF framework accommodates all existing technology indicator types defined in Section 2.4.2 as latent or formative indicators for endogenous and exogenous constructs within the framework. It employs bypass and output indicators for endogenous constructs. In contrast, the framework's output metrics, designed to inform decision-making processes, are derived from output metrics related to endogenous constructs in the SEM model. Additionally, external environment-related indicators, including context-related indicators, contribute to exogenous constructs, enhancing context sensitivity in the DF process. Sohn and Moon (2003) provide an illustrative example using the TCSI metric, a market prospect indicator conceptualised by Watts and Porter (1997), as their SEM model's primary TF output metric.

To comprehend the functionality of the TFSEMDF framework, consider the aggregation and refinement processes from DF Level 0 to DF Level 1. In this transition, regression analysis outputs for technology-related exogenous constructs at DF Level 0 formatively or reflectively influence the technology-related endogenous constructs at DF Level 1. Additionally, results from regression analysis concerning the context-related exogenous constructs at DF Level 0 are instrumental in shaping both context-related exogenous and technology-related endogenous constructs at DF Level 1. Moreover, the outcomes of regression analyses for context-related exogenous constructs at DF Level 1 also influence technology-related endogenous constructs at the same level. At DF Level 1, technology indicators relevant to the constructs are potentially chosen as TF output metrics. Alternatively, they may function as byput technology indicators (Grupp, 1998; Nyberg & Palmgren, 2011) if further aggregation and refinement are necessary at subsequent DF levels. The pattern of aggregation and refinement that occurs as the framework progresses from DF Level x-1 to DF Level x (for x = 1, 2, 3, ..., N-1) follows a similar interconnected structure as observed from DF Level 0 to DF Level 1, with the distinction that constructs at DF Level x-1 contribute to the formation of constructs at DF Level x. This approach ensures a consistent and systematic analysis progression across the TFSEMDF framework's various DF levels.

Constructing a model instantiation of the TFSEMDF framework, as illustrated in Figure 6, initially requires the definition of a set of technology-focused endogenous constructs and context-specific exogenous constructs, each accompanied by their respective technology and context-related measurement indicators. These indicators should accurately reflect the key characteristics of the technology domain under study. Following this, a series of hypothesised relationships are established between these constructs, drawing from diverse sources, including theoretical frameworks, action research (Gustavsen, 2008), everyday wisdom, and speculative insights. The final phase of the SEM model development involves the application of empirical data to each of these measurement indicators, using PLS regression analysis as outlined in Staphorst (2010) and further detailed in Appendix A. This process aims to determine the significance and impact of the measurement indicators (including indicator loadings) and to assess the strength and validity of the hypothesised relationships (including path coefficients), thus providing a comprehensive and empirically supported representation of the framework.

In exploratory studies using SEM, evaluating the significance of hypothesised relationships within the model instantiation is the predominant focus, with the final phase of model construction typically marking the culmination of these investigations, as observed by Staphorst (2010). For the proposed TFSEMDF framework, this hypothesis testing stage becomes instrumental in assessing the effects of evolving contextual elements, like technology policy shifts, on the technology-centric endogenous constructs within the specified technology domain. Consequently, the TFSEMDF framework offers a robust mechanism for predicting the dynamic interplay between technology-driven endogenous constructs and the various external factors, encapsulating an analytical tool capable of forecasting the influence of context-specific exogenous constructs on technological parameters. This approach enhances understanding of the underlying relationships and aids in strategic decision-making within technology management and policy formulation.

The construction of a model instantiation within the proposed framework, utilising the outlined SEM model building methodology, facilitates the forecasting of TF output metrics through a technique aptly termed SEM post-processing, as illustrated by Vinzi et al. (2010). This methodology entails the integration of existing context and technology indicator data (input/byput) from a singular metric measurement instance into the structural equations formulated in the SEM model instantiation (referenced in Section 2.2.2). The process progresses by resolving these equations to deduce the unknown TF output metrics, achieved by computing the corresponding output technology indicator values. This approach enhances the model's predictive accuracy and enriches the framework's interpretative power, enabling a sophisticated analysis of technology forecasts within the studied domain.

4.4 DISCUSSION

The following two sections discuss the results obtained from the research conducted in this study phase. Section 4.4.1 discusses the results of developing generic frameworks for SEM, DF, and technology indicator relational mapping for TF in alignment with Research Objective 1(a). Subsequently, Section 4.4.2 discusses the results related to integrating these frameworks into the generic TFSEMDF framework, fulfilling Research Objective 1(b).

4.4.1 GENERIC FRAMEWORK ABSTRACTIONS FOR SEM, DF AND TF

The achievement of Research Objective 1(a) marked the creation of generic frameworks for SEM, DF, and technology indicator relational mapping for TF. This work contributes to the domain of analytical modelling by providing structured methodologies tailored to each of these distinct fields. The SEM framework introduced a simplistic approach to grouping and layering constructs, the DF framework adapts the DF process to incorporate context and allow for broader applications, and the framework for the relational mapping of technology indicators offers a structured and interrelated organisation of these metrics for use in TF.

The development of the generic SEM framework within this study primarily emphasised the grouping and layering of constructs, marking a distinct adaptation from traditional SEM methodologies. The framework's approach to systematically categorise constructs into context or technology-related groups enhances the structuring of SEM path diagrams. Additionally, the layering of constructs within the framework is a critical feature. Constructs are strategically arranged in layers reflecting their complexity or functional roles within the modelled system. This layering process visualises hierarchical relationships among the constructs, contributing to a more organised and interpretable analysis. By grouping similar or functionally related constructs on the same layer, the framework provides a streamlined approach to understanding the dynamics within SEM models, which is particularly beneficial in complex analytical scenarios.

Incorporating SEM's robust handling of exogenous and endogenous constructs and formative and reflective indicators, the generic SEM framework showcases its capability to represent complex variable relationships within a model. The framework effectively distinguishes between exogenous constructs, which act as independent variables not influenced by the model's internal dynamics, and endogenous constructs, shaped by their interactions within the model. Additionally, integrating both formative and reflective indicators enriches the framework's ability to depict various dimensions and influences of latent constructs, thereby enhancing the model's comprehensiveness and applicability in diverse research settings.

The generic context-sensitive DF framework developed in this study marks an evolution in the DF field by generalising the DF levels and incorporating context into the fusion process. This approach facilitates the application of DF across a broader spectrum of areas beyond its traditional military use. The framework offers a more flexible and universally applicable structure by moving away from the specific JDL/DFG level naming convention. This change is crucial in making DF accessible and relevant to a broader range of applications where the requirements and nature of data differ markedly from military contexts.

Incorporating context into the fusion process is another pivotal aspect of this framework. Context-related information is critical in refining and interpreting data, particularly in complex environments where many external factors may influence data. By integrating context sensitivity into the DF process, the framework enhances the relevance and accuracy of the fusion outcomes. This integration is essential in scenarios where understanding the environment or background conditions is as crucial as the data.

The framework for technology indicator relational mapping for TF, informed by the research of Watts and Porter (1997) and Grupp (1998), establishes a method for segmenting and interconnecting technology indicators. This framework organises indicators into three distinct types based on their functional roles within the technological process: input, byput and output. By classifying indicators in this manner, the framework aids in identifying and understanding the different stages and dimensions of technological development, providing a clearer picture of how various elements contribute to and affect the progression of technology.

The research methodology for achieving Research Objective 1(a) has successfully created the generic frameworks for SEM, DF and TF. These frameworks form the foundational basis for the subsequent development of the TFSEMDF framework. The study has effectively abstracted and synthesised key aspects of SEM, DF and TF into comprehensive, stand-alone generic frameworks through a detailed and systematic approach. Each of these frameworks embodies the essential elements of its respective field and offers the flexibility and depth required for the intricate modelling and analysis process in TF using DF.

4.4.2 Merging Frameworks for SEM, DF and TF

Creating the TFSEMDF framework, achieved under Research Objective 1(b), signifies the successful unification of the generic frameworks for SEM, DF and technology indicator relational mapping for TF. This unified framework responds explicitly to the requirement for a methodological approach that facilitates TF through the lens of context-sensitive DF, implemented via SEM. The TFSEMDF framework is a direct solution to this need, effectively bridging the methodologies of SEM, DF, and TF to create a comprehensive tool for advanced analytical processes in TF.

The development of the TFSEMDF framework aimed to effectively model complex multi-layer systems and, hence, cater for the increasingly sophisticated demands of TF. A standout feature of this framework is its adeptness at processing technology-related and context-related indicator data. This dual capability is crucial for collating and analysing data from diverse sources, which is fundamental for gaining a holistic understanding of the complex systems prevalent in TF undertaking. Significantly, the inclusion of context-related data in the TFSEMDF framework elevates the quality of TF outputs, as it allows for a more comprehensive and detailed interpretation of technological trends and patterns.

Furthermore, the TFSEMDF framework's structure facilitates the systematic processing of data at multiple levels, determined by the user based on the complexity of the technology domain and other constraints such as time and budget. This flexibility in defining the number of levels in the DF process allows for a tailored approach to data analysis, accommodating the specific needs of different applications and system complexities. Furthermore, the TFSEMDF framework addresses the potential diminishing returns that might arise from further aggregation and refinement, ensuring an optimal balance between the depth of analysis and practical limitations.

TFSEMDF employs bi-directional interconnections between indicators and constructs following the conventions of SEM path diagrams. This feature facilitates a detailed and nuanced portrayal of the interactions between various constructs, effectively capturing positive and negative correlations. Such a design enhances the analytical capabilities of the framework, enabling a deeper exploration and understanding of the intricate dynamics present in both technology and context-related data. This feature is handy in revealing the multifaceted relationships and influences within the data, thereby enriching the quality of analysis and insights obtained from the TFSEMDF framework.

Diverging from conventional TF methodologies, the proposed TFSEMDF framework uniquely addresses complex and hierarchical structural relationships between technology indicators and TF output metrics. It extends its analytical capacity to include non-linear and non-Gaussian factors and cyclical dependencies among model variables, ranging from latent to directly observable. Staphorst et al. (2013) elaborate that this comprehensive approach allows for more accurate modelling of real-world systems within the TF context. Moreover, this framework integrates all technology indicators identified in Section 2.4.2. It effectively utilises bypass and output indicators for endogenous constructs, which facilitates the derivation of meaningful output metrics from these constructs. This capability makes the TFSEMDF framework an all-encompassing tool for handling various indicator data and enhances its utility in decision-making processes.

The development of the TFSEMDF framework marks the successful accomplishment of Research Objective 1(b), representing a significant milestone in this study. The resultant outcome is a multidisciplinary framework that adeptly models complex hierarchical relationships inherent in modern-day systems during TF. This reflects the TFSEMDF framework's ability to adeptly navigate and integrate the complexities associated with TF, providing a comprehensive tool for analysing and forecasting in environments where technology and context-related factors are deeply intertwined.

4.5 CONCLUDING REMARKS

This chapter explored the study's first phase, which focused on developing the TFSEMDF framework. It commenced with an introduction that outlined the research phase's scope. Next, the research methodology section thoroughly explained the processes for first developing generic frameworks for SEM, context-sensitive DF and the relational mapping of technology indicators for TF and then combining these generic frameworks to derive the TFSEMDF framework.

The results section presented in this chapter provided an exposition of the development of the generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF, followed by showcasing the proposed integration that constitutes the derived TFSEMDF framework. This integration exemplifies the transition from distinct theoretical frameworks to a coherent functional amalgamation that encapsulates each constituent theoretical methodology's complexities while highlighting their synergistic potential in practical applications.

The discussion section considered the outcomes of the efforts to create generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF. Subsequently, the discussion delves deeply into the results of integrating these generic frameworks to derive the TFSEMDF framework. TFSEMDF's capability to incorporate context-related information into the TF process takes centre stage in this discussion. This exploration provides insights into how the framework leverages this context-related information integration to refine TF output metrics.

Section 7.2.1 provides the conclusions for the work presented in Chapter 4. This section summarises the key findings from the integration of SEM, context-sensitive DF and TF to synthesise the TFSEMDF framework, offering a retrospective view of the efforts and outcomes detailed in this phase.

The TFSEMDF framework developed in this chapter is applied in the second research phase, as presented in Chapter 5, to implement and analyse autoregressive and cross-sectional NREN model instantiations. In Chapter 6, which details the third research phase, the strengths and weaknesses of the TFSEMDF framework are considered using baseline and structurally disarranged NREN model instantiations.

CHAPTER 5 – RESEARCH PHASE 2: TFSEMDF FRAMEWORK APPLICATION

5.1 INTRODUCTION

This chapter is dedicated to the study's second phase, which focuses on the practical application of the TFSEMDF framework. This phase first addressed Research Objective 2(a), which involved the construction and PLS regression analysis of the autoregressive model instantiation for NRENs within the TFSEMDF framework to accomplish longitudinal forecasting. Next, the focus shifted to Research Objective 2(b), which concentrated on the construction and PLS regression analysis of the cross-sectional NREN model instantiation to accomplish transversal forecasting.

The chapter commences by outlining the methodologies applied to achieve Research Objectives 2(a) and 2(b). It then details the results obtained from using these methodologies. The analysis within this chapter culminates in a critical discussion of the results, evaluating the efficacy and precision of the TFSEMDF framework as evidenced by the outcomes of the autoregressive and cross-sectional NREN model instantiations.

This chapter builds upon and refines the research methodologies and findings initially introduced by Staphorst et al. (2014) concerning developing and examining the autoregressive NREN model instantiation. It also extends the methodological advancements and detailed analyses from Staphorst et al. (2016a), which focused on the cross-sectional NREN model instantiation. These foundational studies provide the empirical and theoretical basis for the comprehensive discussions in this chapter, effectively integrating insights from both longitudinal and transversal forecasting perspectives within the TFSEMDF framework in the context of NRENs.

5.2 RESEARCH METHODOLOGY

The following sections offer a comprehensive overview of the research methodologies applied during this study phase. Section 5.2.1 explores the methods used for constructing and analysing the autoregressive NREN model instantiation of the TFSEMDF framework, aligning with Research Objective 2(a). Furthermore, Section 5.2.2 discusses the approaches taken in constructing and analysing the cross-sectional NREN model instantiation of the TFSEMDF framework, which corresponds to Research Objective 2(b).

5.2.1 AUTOREGRESSIVE NREN MODEL INSTANTIATION

5.2.1.1 Autoregressive NREN Model Instantiation Construction

Research Objective 2(a) centred on constructing and analysing an autoregressive NREN model instantiation for longitudinal forecasting, employing PLS regression as detailed in Appendix A. This endeavour began with developing an autoregressive NREN model instantiation of the TFSEMDF framework shown in Figure 6. While adhering to the procedures outlined in Section 4.3.2, the model emerged by applying insights gleaned from TERENA (2011, 2012) and GÉANT (2022). The model was further informed by the knowledge acquired from the author's action research (Gustavsen, 2008) during his directorship at SANReN (SANReN, n.d.) between 2013 and 2021. As the director of SANReN, the author conducted this action research through a rigorous approach involving active engagement with specific real-world NREN-related problems, collaborative data collection within SANReN, and analysis to inform practical interventions, all aimed at achieving meaningful and sustainable improvements within SANReN. This approach fostered a dynamic feedback loop between research and practice, facilitating informed decision-making and positive change at SANReN. Note that the model construction approach did not rely on additional hypotheses from peer-reviewed literature.

Figure 7 depicts the autoregressive NREN model instantiation of the TFSEMDF framework, adopting N=3 context-sensitive DF structure. Each level distinctively addresses specific aspects of the NREN: Level 0 dedicates itself to examining the NREN infrastructure, providing a foundational analysis of the physical and technological underpinnings. Level 1 focuses on NREN services, delving into the network's offerings to the education, research, and innovation communities. Finally, Level 2 concentrates on NREN reach, exploring the extent of the network's penetration into the communities it services (GÉANT, 2022).

Within Level 0 of the autoregressive NREN model instantiation, the sole technology-related endogenous construct identified is *NREN Infrastructure Capability* (η_1). This construct is strategically defined to assess the extent of the NREN's preference for owning versus leasing fibre optic cable infrastructure, as detailed in GÉANT (2022). NREN experts generally theorise that an NREN's infrastructure capability is intricately associated with two formative input technology indicators constituting the construct. The first indicator is tasked with measuring the length of dark fibre-owned optical cabling infrastructure (labelled as Y_1, along with its indicator loading π_{y1}), and the second focuses on the quantity of rented managed circuits (indicated as Y_2, with its respective indicator loading π_{y2}). This dual-indicator approach intends to comprehensively model the strategic approach of NRENs to infrastructure management, capturing both ownership and leasing dimensions.

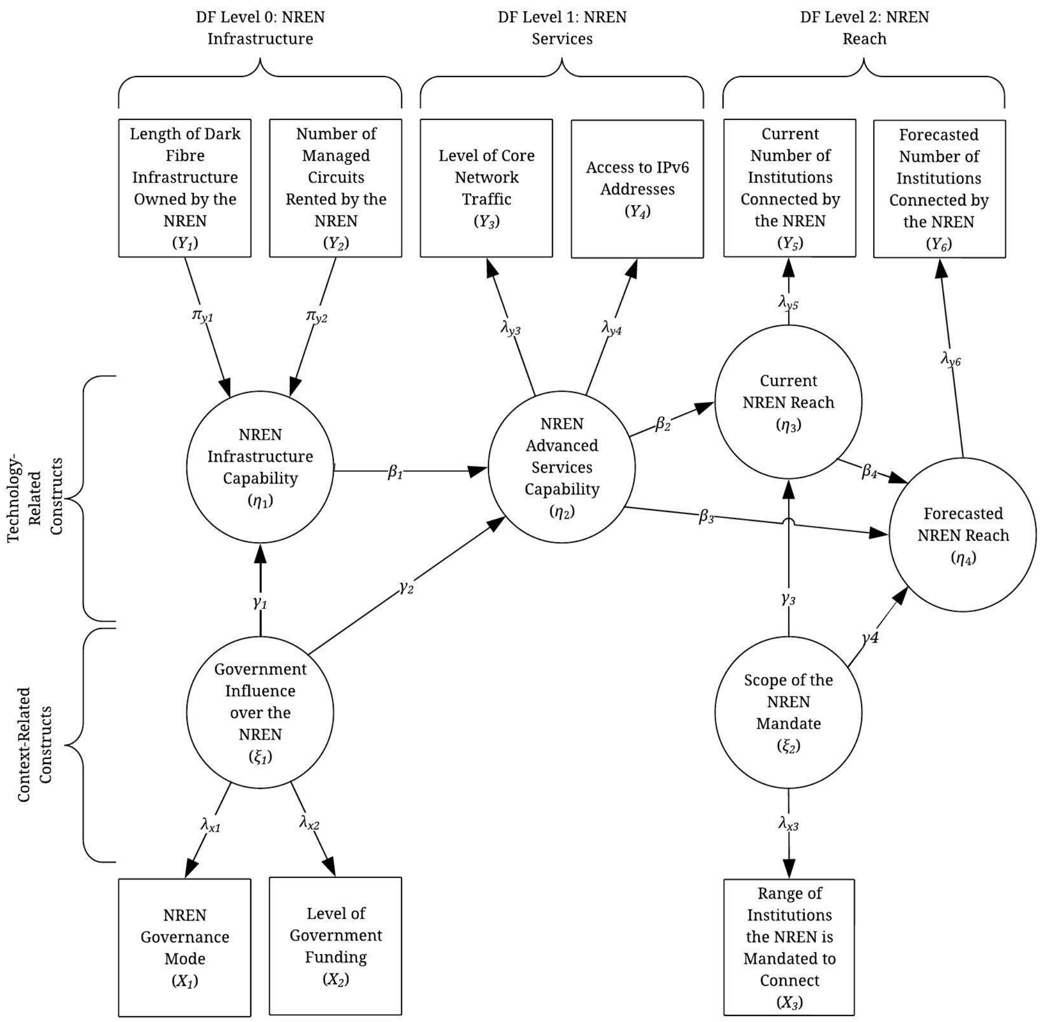


Figure 7: Autoregressive NREN Model Instantiation of the TFSEMDF Framework

Also situated at Level 0 is the context-specific exogenous construct titled *Government Influence over the NREN (ξ_1)*, equipped with two reflective indicators. These indicators independently represent the construct: the first measures the governance mode of the NREN (notated as X_1 with indicator loading λ_{x1}), which spans a spectrum from total government control to no government involvement, as discussed in GÉANT (2022) and Staphorst (2010). The second indicator evaluates the level of government funding provided to the NREN (denoted as X_2 with indicator loading λ_{x2}). Path coefficient γ_1 expresses the postulated positive association between *Government Influence over the NREN (ξ_1)* and *NREN Infrastructure Capability (η_1)*. This relationship quantifies the influence of government actions and policies on the strategic decisions regarding NREN infrastructure, providing insights into how governance and funding levels impact the NREN's operational and strategic choices.

Level 1 of the autoregressive NREN model instantiation introduces the endogenous construct *NREN Advanced Services Capability (η_2)*, which encapsulates the NREN's proficiency in providing an array of sophisticated services beyond primary Internet offerings, such as Lightpaths and Science Gateways (GÉANT, 2022). Although this level does not designate an exogenous context-related construct, the hypothesis suggests that the constructs at Level 0, which gauge government influence on the NREN and the NREN's strategic choice between owning and leasing fibre optic cable infrastructure, exert a positive influence on the NREN's advanced service provision capabilities. Figure 7 quantifies these conjectured connections with path coefficients γ_2 and β_1. The hypothesis suggests that reducing government oversight and greater control over infrastructure resources bolsters the NREN's capacity to innovate and supply advanced services (GÉANT, 2022). This supposition rests on the premise that enhanced autonomy in infrastructure management will likely propel the development and dissemination of cutting-edge NREN services.

At Level 1, two reflective byput technology indicators operationalise *NREN Advanced Services Capability (η_2)*. The first indicator quantifies the volume of network traffic within the core infrastructure of the NREN (labelled as Y_3 with indicator loading λ_{y3}). In contrast, the second gauges the proportion of users accessing IPv6 addresses (Deering & Hinden, 1998), denoted as Y_4 with indicator loading λ_{y4}. Recognition that the scarcity of traditional Internet Protocol version 4 (IPv4) addresses poses a considerable obstacle to the innovation of new services at NRENs (GÉANT, 2022; TERENA, 2012) informed the selection of the latter indicator. The utilisation of IPv6 adoption as an indicator measures the NREN’s forward-looking capacity to surmount such limitations and facilitate the growth of new and advanced services (Deering & Hinden, 1998).

Level 2 of the autoregressive NREN model instantiation dedicates itself to evaluating the network's reach, which indirectly measures the NREN's impact on its user communities, as described in GÉANT (2022) and SANReN (n.d.). A unique context-related exogenous construct, named *Scope of the NREN Mandate (ξ_2)*, is introduced at this juncture. It is quantified using a single reflective measurement indicator that tallies various institutions authorised to receive NREN connectivity (notated as X_3 with indicator loading λ_{x3}). This indicator tracks the mandate's scope, which may vary from a limited spectrum, exclusively encompassing higher education and research institutions, to a broad spectrum that includes a vast range of public institutions and potentially certain entities from the private sector. The measure indicates the NREN's operational expanse and potential influence across different research and educational landscapes.

In the architecture of Level 2 of the autoregressive NREN model instantiation, two essential technology-related endogenous constructs are introduced: *Current NREN Reach (η_3)* and *Forecasted NREN Reach (η_4)* (SANReN, n.d.). The first, *Current NREN Reach (η_3)*, is operationalised by a byput technology indicator that enumerates the existing number of institutions with NREN connectivity (specified as Y_5 with indicator loading λ_{y5}). The second, *Forecasted NREN Reach (η_4)*, is characterised by an output technology indicator that anticipates the number of institutions expected to be connected to the NREN in the future (specified as Y_6 with indicator loading λ_{y6}). This latter indicator is pivotal, as it serves as the model's TF output metric. The conceptualisation of these constructs, with *Current NREN Reach (η_3)* informing the *Forecasted NREN Reach (η_4),* confers upon the model instantiation its autoregressive capacity (Burant, 2022), which is essential for conducting longitudinal forecasting and modelling the evolution of the NREN's reach over time.

The model hypothesises a positive relationship between the Level 1 construct, NREN *Advanced Services Capability (η_2)*, and the current and forecasted NREN reach constructs, represented by path coefficients β_2 and β_3, respectively. This correlation is supported by observations from SANReN, indicating that expanding the advanced services portfolio of an NREN leads to increased demand for NREN connectivity from entities not yet connected (SANReN, n.d.). The perception that a broad array of advanced services signifies the NREN's maturity and reliability, as Greaves (2009) postulated, influences such demand. Additionally, the construct *Scope of the NREN Mandate (ξ_2)* is presumed to positively impact current and forecasted NREN reach, as indicated by path coefficients γ_3 and γ_4. Lastly, in the context of SANReN, it has been observed that current NREN reach positively influences forecasted NREN reach (denoted by path coefficient β_4), suggesting an inherent expectation that existing connectivity trends provide insights into future network growth, as per SANReN (n.d.).

5.2.1.2 Autoregressive NREN Model Instantiation Research Propositions

The hypothesised relationships within the autoregressive NREN model instantiation, as discussed in Section 5.2.1.1 and depicted in Figure 7, lead to a series of research propositions. The study meticulously evaluated these propositions through PLS regression analysis in Section 5.3.1. Figure 7 and Table 4 detail the alignment between these research propositions and the various paths in the NREN model instantiation.

1. **Research Proposition H1:** A positive correlation exists between NREN infrastructure capability and the level of government influence over the NREN. This proposition postulates that higher government involvement correlates with a more pronounced focus on developing and enhancing NREN infrastructure.
2. **Research Proposition H2:** A positive correlation exists between an NREN's advanced services capability and its preference for owning rather than leasing fibre optic cable infrastructure. The assumption is that ownership of infrastructure often correlates with the ability to provide a more extensive and robust range of NREN advanced services.
3. **Research Proposition H3:** There is a positive correlation between the NREN's advanced services capability and the level of government influence. This proposition suggests that increased government influence often aligns with more substantial NREN advanced service offerings.
4. **Research Proposition H4:** The current NREN reach positively correlates with the NREN's advanced services capability. This proposition assumes that an extensive range of NREN advanced services is often associated with a broader current network reach.
5. **Research Proposition H5:** A positive correlation is evident between the forecasted NREN reach and the NREN's advanced services capability. Here, the conjecture is that the spectrum of advanced services an NREN offers correlates to expectations of an expanded future network reach.
6. **Research Proposition H6:** There is a positive correlation between the scope of an NREN's mandate and its current reach. The proposition conjectures that a broader NREN mandate for connecting a diverse range of institutions coincides with a wider current network reach.

7. **Research Proposition H7:** The scope of the NREN's mandate positively correlates with the forecasted NREN reach. This proposition suggests that an inclusive mandate is frequently associated with the potential for future growth in the network's reach.
8. **Research Proposition H8:** A positive correlation exists between the current NREN reach and the forecasted NREN reach. This proposition assumes that a more extensive current network reach suggests a more extensive projected future reach.

5.2.1.3 Autoregressive NREN Model Instantiation Indicator Data

Through PLS regression analysis, the study determined the indicator loadings and path coefficients for the model depicted in Figure 7 by utilising secondary data sourced from TERENA's NREN Compendiums for the years 2011 (TERENA, 2011) and 2012 (TERENA, 2012). The autoregressive NREN model instantiation's indicator data composition, derived from these comprehensive TERENA NREN Compendiums, is briefly summarised in Table 1. These Compendiums reflect substantial participation from the NREN community, with 61 NRENs contributing data for the 2011 edition (TERENA, 2011) and 54 NRENs participating in the 2012 survey (TERENA, 2012). Section C.2 of Appendix C provides selected excerpts of the data employed in Table 1 from TERENA (2011, 2012) for illustrative purposes. Appendix D provides directions to access the Figshare collection containing these Compendiums as part of the study's research data repository.

This study utilised SmartPLS software (Ringle et al., 2022) to analyse the autoregressive NREN model instantiation shown in Figure 7. Specifically, SmartPLS facilitated the calculation of indicator loadings and path coefficients using PLS regression (see Section 5.3.1.1), executing the process detailed in Appendix A. Given the variability in scaling methods and data ranges in TERENA's data collection (TENRENA, 2011, 2012), SmartPLS's data normalisation capability was particularly crucial. Additionally, the software was used to assess the reliability and validity of the model instantiation, following the process and criteria set out in Appendix B, with these results discussed in Section 5.3.1.2. It is important to note that the dataset for this analysis was limited to inputs from 27 NRENs that provided complete survey responses, as specified in Table 1. To mitigate the impact of missing data, SmartPLS employed a mean replacement algorithm, a method advocated by Ringle et al. (2022), to enhance the robustness of the analysis in the face of incomplete data.

Table 1: Autoregressive NREN Model Instantiation Indicator Data Composition

Autoregressive NREN Model Instantiation Technology or Context-Related Indicator	**Indicator Data Composition**
NREN Governance Mode (X_1)	Extracted from the online profiles of the respondent NRENs of the 2011 Compendium (TERENA, 2011) using the following scaling: The NREN is a government agency or part of a ministry = 3 Government appoints at least half of the NREN's governing body = 2 Indirect relationship between the NREN and government = 1 No formal relationship between the NREN and government = 0
Level of Government Funding (X_2)	Level of government funding (as a percentage of total funding) received by respondent NRENS, as summarised in Graphs 6.4.2 and 6.4.3 in the 2011 NREN Compendium (TERENA, 2011)
Range of Institutions the NREN is Mandated to Connect (X_3)	The sum of the institution types in Table 2.2.1 of the 2011 NREN Compendium (TERENA, 2011) supported by respondent NRENs
Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)	The total length of dark fibre [in kilometres] owned by respondent NRENs as summarised in Table 3.6.3 of the 2011 NREN Compendium (TERENA, 2011)
Number of Managed Circuits Rented by the NREN (Y_2)	The total number of managed circuits rented by respondent NRENs as summarised in Table 3.3.2 of the 2011 NREN Compendium (TERENA, 2011)

Level of Core Network Traffic (Y_3)	Annual level (measured in Terabytes (TB) per year) of traffic sent on to the backbone networks of respondent NRENs, as measured by T_1+T_4 in Graphs 4.2.1 and 4.2.2 in the 2011 NREN Compendium (TERENA, 2011)
Access to IPv6 Addresses (Y_4)	The sum of the percentage of users from respondent NRENs that have both IPv4 and IPv6 addresses with the percentage of users that only have IPv6 addresses, as listed in Tables 4.6.1 and 4.6.2 in the 2011 NREN Compendium (TERENA, 2011)
Current Number of Institutions Connected by the NREN (Y_5)	The sum of the institutions in Table 2.2.1 of the 2011 NREN Compendium (TERENA, 2011) connected by respondent NRENs
Forecasted Number of Institutions Connected by the NREN (Y_6)	The sum of the institutions in Table 2.2.1 of the 2012 NREN Compendium (TERENA, 2012) connected by respondent NRENs

5.2.2 Cross-Sectional NREN Model Instantiation

5.2.2.1 Cross-Sectional NREN Model Instantiation Construction

Research Objective 2(b) focused on constructing and analysing a cross-sectional NREN model instantiation for transversal forecasting through PLS regression. The initial step towards this objective involved the development of an appropriate cross-sectional model instantiation of the TFSEMDF, as depicted in Figure 6. Following the process detailed in Section 4.3.2, the study integrated insights from TERENA (2011, 2012) and GÉANT (2022) with the knowledge acquired from the author's action research (Gustavsen, 2008) during their tenure as Director of SANReN (SANReN, n.d.) from 2013 to 2021. The approach also encompassed hypotheses formulated and validated in peer-reviewed literature, improving the process employed during the construction of the autoregressive NREN model instantiation considered in Research Objective 2(a). Notably, the deployment of this cross-sectional NREN model instantiation in the study, as expounded in the analysis results in Section 5.3.2, was solely for forecasting the relational dynamics between technology-centric endogenous constructs and context-related exogenous constructs at specific time instances. The model's design inherently restricts its forecasting capability to cross-sectional, single-time-point analyses, precluding its application for longitudinal forecasting.

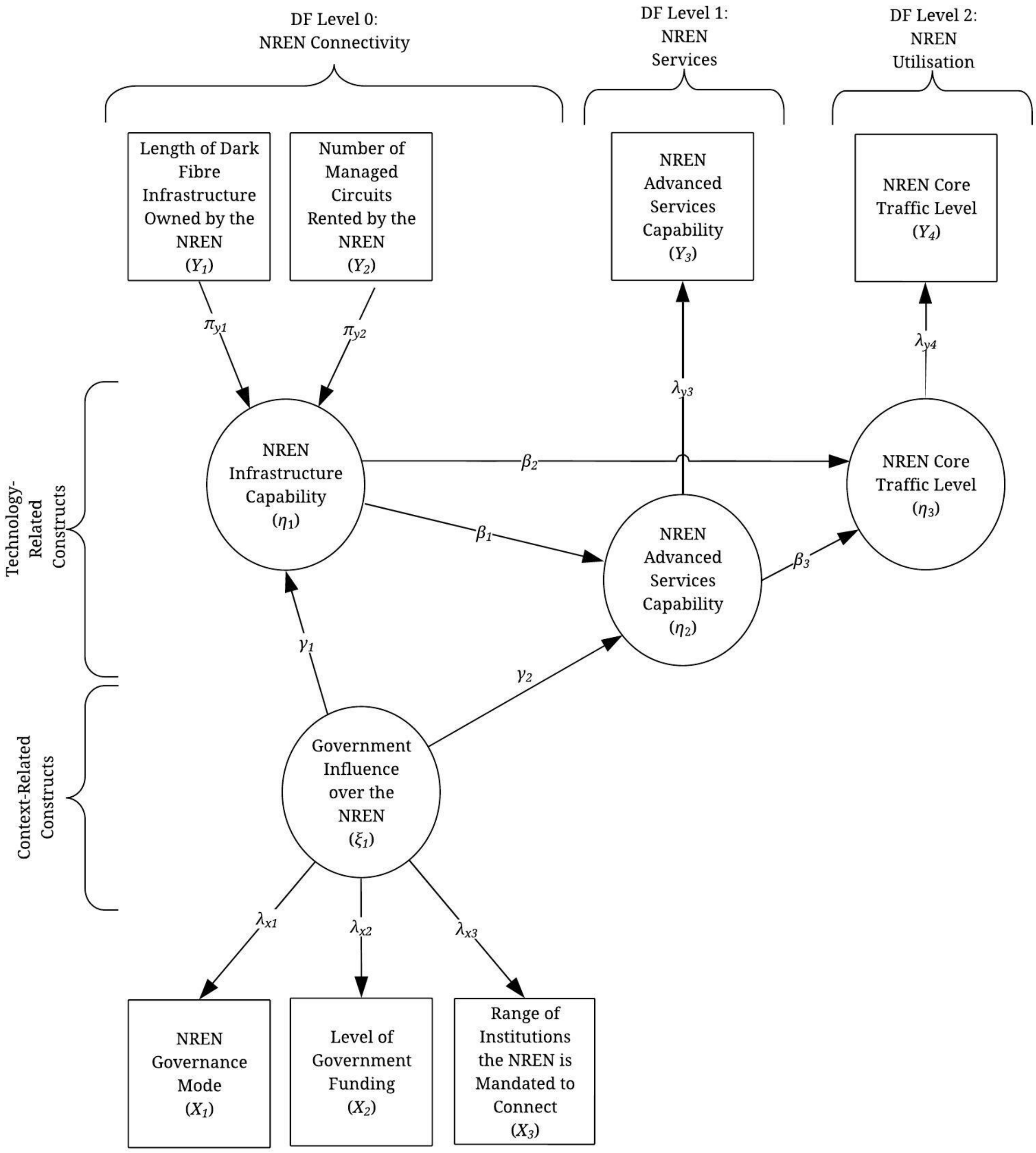


Figure 8: Cross-Sectional NREN Model Instantiation of the TFSEMDF Framework

Figure 8 displays the cross-sectional NREN model instantiation of the TFSEMDF framework, structured using N=3 distinct DF levels: Level 0 concentrates on NREN Connectivity, emphasising the infrastructure the NREN provides for delivering advanced services. Level 1 shifts the focus to NREN Services, exploring the suite of advanced services offered to users to maximise the use of the NREN-provided infrastructure. Finally, Level 2 addresses NREN Utilization, assessing how users employ the available advanced services. This multi-level approach enables a comprehensive evaluation of the NREN, from its foundational infrastructure and service offerings to end users' ultimate utilisation of these services.

The NREN Connectivity level within the cross-sectional NREN model instantiation, as illustrated in Figure 8, effectively consolidates the first six layers of the 7-layered Open Systems Interconnection (OSI) model (Zimmerman, 1980), covering Layer 1 (Physical) through to Layer 6 (Presentation layer). This aggregation captures comprehensive network functionalities, from physical infrastructure to data representation and management protocols. Conversely, the NREN Services level is aligned with the OSI model's 7th layer (the Application layer), focusing on the network-related application services directly accessible to users. Additionally, Bauer and Patrick (2004) expanded the OSI model to include Layers 8 to 10, encompassing Human-Computer Interaction (HCI) aspects. Within this extended framework, the NREN Utilisation level is a conceptual embodiment of these additional HCI-focused layers, representing the network's user engagement and interaction dimensions.

Level 0 in the cross-sectional NREN model instantiation, focused on infrastructure-related technology metrics, identified a single endogenous construct: *NREN Infrastructure Capability* (η_1). This construct intends to quantify the NREN's investment in dark fibre infrastructure and managed circuits (GÉANT, 2022; TENERA, 2011; TERENA, 2012). Dark fibre is described as fibre infrastructure wholly owned by the NREN or secured under a long-term Indefeasible Right of Use (IRU) (GÉANT, 2022). Conversely, managed circuits refer to fibre infrastructure owned by third parties but with bandwidth services on this fibre leased by the NREN.

Based on insights from TENREA (2011, 2012), two formative input technology indicators comprehensively represent the *NREN Infrastructure Capability* (η_1). The first indicator, *Length of Dark Fibre Infrastructure Owned by the NREN* (Y_1) with an associated indicator loading π_{y1}, measures the total length of dark fibre infrastructure owned by the NREN. The second, labelled as *Number of Managed Circuits Rented by the NREN* (Y_2) with indicator loading π_{y2}, quantifies the number of managed circuits rented by the NREN. This dual-indicator approach aims to provide a holistic view of the NREN's infrastructure investment strategy, encompassing ownership and leasing dimensions.

Additionally, at Level 0 of the model, a context-specific exogenous construct, *Government Influence over the NREN (ξ_1)*, is delineated. Three reflective indicators, each capable of independently representing the construct, operationalised this construct. The first indicator, *NREN Governance Mode (X_1)* with indicator loading λ_{x1}, assesses the governance style of the NREN, which may span from fully government-driven to completely independent governance, as reported in GÉANT (2022). The second, *Level of Government Funding (X_2)* with indicator loading λ_{x2}, measures the extent of financial support the government provides to the NREN. The third indicator, *Range of Institutions the NREN is Mandated to Connect (X_3)* with indicator loading λ_{x3}, evaluates the diversity of institutions that the NREN is obligated to connect, ranging from a singular type, like universities, to a broader spectrum encompassing research organisations, schools, and other institutions, as noted in GÉANT (2022).

A positive correlation is hypothesised between *Government Influence over the NREN (ξ_1)* and *NREN Infrastructure Capacity (η_1),* symbolised by the path coefficient γ_1. This relationship assumes that government involvement is often pivotal in various aspects of an NREN's operational chain, including infrastructure funding, policy formulation, and regulation. Such governmental participation is essential for the NREN's successful evolution in connectivity and advanced service provision (Greaves, 2009; Janz & Kutanov, 2012). This hypothesis underscores the integral role of governmental actions in shaping the infrastructure and service landscape of an NREN.

Incorporating additional context-related measurement indicators and constructs from various domains like political, economic, sociological, legal, and environmental spheres can significantly augment the predictive capacity of a model instantiation, such as the presented cross-sectional NREN model instantiation. These diverse indicators have the potential to refine and enhance the model's accuracy in forecasting TF output metrics by introducing a more multifaceted understanding of external influences. However, in the specific case of this NREN model instantiation, which underwent empirical testing with the data from the 2011 TERENA NREN Compendium (TERENA, 2011), the selection of context-related measurement indicators was confined to those relevant to the *Government Influence over the NREN (ξ_1)* construct. This limitation was dictated by the available data from the Compendium, highlighting the critical role that data accessibility and scope play in shaping the comprehensiveness of a model instantiation.

At Level 1 of the cross-sectional NREN model instantiation, which concentrates on service-related technology metrics, a singular exogenous technology-related construct, *NREN Advanced Services Capability (η_2)*, is defined. This construct encapsulates the NREN's ability to offer a comprehensive range of advanced services, as identified in TERENA (2011, 2012). These services include but are not limited to Authentication and Authorization Infrastructure (AAI) services, the provision and hosting of Identity Federation Services, and inter-federating with other NRENs. The construct is quantified using a reflective byput technology metric as its indicator (labelled *NREN Advanced Services Capability (Y_3)* with an indicator loading λ_{y3}). It assesses the breadth of the advanced services portfolio provided and hosted by the NREN.

A hypothesised positive link exists between *NREN Infrastructure Capability (η_1)* and *NREN Advanced Services Capability (η_2)*, indicated by the path coefficient β_1. This relationship is grounded in the premise, as Greaves (2009) suggested, that a robust and advanced infrastructure capability is a prerequisite for an NREN to deliver and manage a diverse array of advanced services effectively. This posited connection underscores the critical interdependence between an NREN's infrastructural foundations and its capacity to extend sophisticated service offerings.

At Level 1 within the cross-sectional NREN model instantiation, although no specific exogenous context-related construct is explicitly defined, a positive correlation is hypothesised between Level 0's *Government Influence over the NREN (ξ_1)* and the NREN's capacity to deliver advanced services. In Figure 8, the path coefficient γ_2 represents this relationship. The basis for this hypothesised relationship draws on insights from Greaves (2009) and Janz and Kutanov (2012), suggesting that government intervention plays a crucial role at various junctures within the NREN value chain. Such intervention is deemed essential for an NREN's successful development and maturation, particularly concerning its portfolio of advanced services. The underlying premise is that governmental support and regulatory frameworks can significantly influence and facilitate the expansion and enhancement of NREN's advanced service offerings.

Level 2 of the cross-sectional NREN model instantiation focuses on utilising the NREN. This aspect is crucial as it often serves as a proxy for assessing the impact of the NREN in its beneficiary communities, as noted in SANReN (n.d.), and for evaluating the Return of Investment (ROI) for the NREN's funders, as discussed by Bech (2011). A single context-related exogenous construct, *NREN Core Traffic Level (η_3)*, is introduced at this level. This construct reflects the bandwidth usage within the core network of the NREN, an aspect highlighted in TERENA

(2011, 2012), and is pivotal in representing the NREN's overall utilisation. The reflective measurement indicator *NREN Core Traffic Level (Y_4)* quantifies the *NREN Core Traffic Level (η_3)* construct and corresponds to the indicator loading λ_{y4}. This measurement indicator is essential for assessing the current usage of the NREN's core network and serves as the model's TF output metric. By focusing on core traffic levels, this construct provides an empirical basis for understanding the operational intensity of the NREN and its alignment with the requirements of its user base, thereby offering valuable insights into the network's efficiency and effectiveness.

In the cross-sectional NREN model instantiation, both *NREN Infrastructure Capability (η_1)* and *NREN Advanced Services Capability (η_2)* have positive relationships with *NREN Core Traffic Level (Y_4)*, as indicated by path coefficients β_2 and β_3, respectively. Findings in Savory (2012) support this correlation, emphasising the influence of robust infrastructure on network utilisation, particularly core network traffic. Similarly, Greaves (2009) and Janz & Kutanov (2012) suggest that a mature portfolio of advanced services is instrumental in driving the utilisation of broadband networks. Therefore, these relationships posit that the strength of the NREN's infrastructure and the richness of its service offerings are critical drivers in enhancing network usage, demonstrating the interconnected nature of these factors in influencing overall network performance and utilisation.

5.2.2.2 Cross-Sectional NREN Model Instantiation Research Propositions

In Figure 8's cross-sectional NREN model instantiation, a series of research propositions emerge from the hypothesised relationships between various constructs. Section 5.3.2 rigorously examines these propositions using PLS regression analysis. Figure 8 and Table 9 outline the alignment of these research propositions with specific paths in the NREN model instantiation.

- **Research Proposition H1:** There is a positive correlation between the NREN's infrastructure capability and the level of government influence over the NREN. As highlighted in Greaves (2009), this postulated correlation indicates that more significant government influence coincides with higher infrastructure capability within NRENs.
- **Research Proposition H2:** A positive correlation exists between NREN's advanced services capability and its infrastructure capability. According to Greaves (2009), these

two aspects frequently correlate, suggesting that NRENs with advanced infrastructure tend to have enhanced service capabilities.

- **Research Proposition H3:** The NREN's advanced services capability positively correlates with government influence over the NREN. Greaves (2009) and Janz and Kutanov (2012) suggest that these factors often align, indicating that government influence is typically associated with more developed NREN advanced service offerings.
- **Research Proposition H4:** Savory (2012) noted a positive correlation between the level of core network traffic in the NREN and its Infrastructure Capability. The assumed correlation implies that increased NREN network traffic levels accompany higher infrastructure capability.
- **Research Proposition H5:** A positive correlation exists between the level of core network traffic in the NREN and its advanced services capability. The proposition hypothesises that an extensive range of advanced services relates to higher network usage, as Greaves (2009) and Janz and Kutanov (2012) suggested.

5.2.2.3 Cross-Sectional NREN Model Instantiation Indicator Data

Like the indicator data used for the analysis of the autoregressive NREN model instantiation, the calculation of the indicator loadings and path coefficients for the cross-sectional NREN model instantiation, as illustrated in Figure 8, used secondary data from the 2011 TERENA NREN Compendium (TERENA, 2011). The study conducted this analysis using PLS regression, as described in Appendix A. The composition of the cross-sectional NREN model instantiation indicator data, as collated from the 2011 TERENA NREN Compendium, is summarised in Table 2. This data compilation used responses from 61 NRENs who participated in TERENA's 2011 Compendium survey, contributing a comprehensive dataset for the analysis. Section C.3 of Appendix C provides an excerpt of the data from TERENA (2012) used in Table 2 for illustrative purposes. Appendix D provides directions to access the Figshare collection containing this Compendium, part of the study's research data repository.

In this study, paralleling the approach taken with the autoregressive NREN model instantiation within the TFSEMDF framework, the SmartPLS software (Ringle et al., 2022) was utilised for the realisation of the cross-sectional NREN model instantiation, as illustrated in Figure 8. SmartPLS was employed to compute all indicator loadings and path coefficients via PLS regression, with the findings detailed in Section 5.3.2.1. An essential aspect of the SmartPLS setup was the normalisation of all indicator data to accommodate the diverse scaling methods and ranges used by TERENA in the initial data collection phase.

Table 2: Cross-Sectional NREN Model Instantiation Indicator Data Composition

Cross-Sectional NREN Model Instantiation Technology or Context-Related Indicator	**Indicator Data Composition**
NREN Governance Mode (X_1)	Extracted from the online profiles of the respondent NRENs of the 2011 Compendium (TERENA, 2011) using the following scaling: • The NREN is a government agency or part of a ministry = 3 • Government appoints at least half of the NREN's governing body = 2 • Indirect relationship between the NREN and government = 1 • No formal relationship between the NREN and government = 0
Level of Government Funding (X_2)	Level of government funding (as a percentage of total funding) received by respondent NRENs, as summarised in Graphs 6.4.2 and 6.4.3 in TERENA (2011)
Range of Institutions the NREN is Mandated to Connect (X_3)	The sum of the institution types supported by respondent NRENs, as shown in Table 2.2.1 of TERENA (2011)
Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)	The total length of dark fibre [in kilometres] owned by respondent NRENs as summarised in Table 3.6.3 of TERENA (2011)
Number of Managed Circuits Rented by the NREN (Y_2)	The total number of managed circuits rented by respondent NRENs as summarised in Table 3.3.2 of TERENA (2011)

NREN Advanced Services Capability (Y_3)	The total number of positive answers to the following questions in Table 5.3.1.1 in TERENA (2011): • Does the NREN provide AAI services? • Does the NREN provide Identity Federation services? • Does the NREN operate the Identity Federation services? • Does the NREN's Identity Federation services inter-federate with those offered by other NRENs?
NREN Core Traffic Level (Y_4)	Annual level (measured in TB per year) of traffic sent on to the backbone networks of respondent NRENs, as measured by T_1+T_4 in Graphs 4.2.1 and 4.2.2 in TERENA (2011)

Further, SmartPLS was instrumental in assessing the reliability and validity of the model, following the criteria established by Staphorst et al. (2013, 2017). Section 5.3.2.2 elaborates on the results of these assessments. A notable point in the data handling was that only 28 NRENs provided complete survey responses for the required calculations, as per Table 2. SmartPLS used a mean replacement algorithm to address the missing data issue, as suggested by Ringle et al. (2022). This approach ensured a thorough and rigorous analysis despite the constraints of incomplete data, thereby enhancing the reliability and validity of the model's outcomes.

5.3 RESULTS

The following two sections present a detailed analysis of the results obtained during this study phase. Section 5.3.1 provides the results from evaluating the autoregressive NREN model instantiation of the TFSEMDF framework, directly relating to Research Objective 2(a). Section 5.3.2 reports the results from analysing the cross-sectional NREN model instantiation of the TFSEMDF framework, which aligns with Research Objective 2(b).

5.3.1 AUTOREGRESSIVE NREN MODEL INSTANTIATION

The following sections present the PLS regression analysis, with reliability and validity assessment, performed for the autoregressive NREN model instantiation of Figure 7. As per the approach detailed in Table 1, indicator data for this analysis was composed of data from the 2011 and 2012 TERENA NREN Compendiums (TERENA, 2011, 2012).

As described in Section 5.2.1.3., the sample sizes of NRENs that responded to the 2011 and 2012 TERENA NREN Compendiums (TERENA, 2011, 2012) were 61 and 54, respectively. Furthermore, only 27 responding NRENs provided the complete data set required per the indicator data composition approach detailed in Table 1. However, as explained in Section 6.3.1.3, one of the strengths of PLS-SEM is its ability to accommodate small sample sizes. Moreover, a commonly accepted norm is that PLS regression for SEM necessitates a sample size that is at least tenfold the number of endogenous or exogenous formative indicators linked to the most complex latent construct within the model (Chin & Newsted, 1999; Goodhue et al., 2006). For the autoregressive NREN model instantiation depicted in Figure 7, the construct *NREN Infrastructure Capability* (η_1) has the most complex formative structure with two indicators, namely *Length of Dark Fibre Infrastructure Owned by the NREN* (Y_1) and *Number of Managed Circuits Rented by the NREN* (Y_2). Hence, PLS regression for this model instantiation was successful as the minimum calculated sample size exceeded 20.

5.3.1.1 Autoregressive NREN Model Instantiation SEM Regression Results

The following presentation of the PLS regression results for the autoregressive NREN model instantiation adheres to the reporting guidelines established by Vinzi et al. (2010) and refined in Staphorst (2010) and Staphorst et al. (2017). In alignment with these guidelines, this study categorises the results into two sections. The analysis first addresses the regression results of the measurement portion of the SEM path diagram of the TFSEMDF model instantiation. It encompasses the loadings of all measurement indicators within the autoregressive NREN model instantiation and offers a detailed view of each indicator's performance. Subsequently, the focus shifts to the structural portion of the SEM path diagram, reporting the path coefficients that explain the interrelationships between various constructs.

A. Measurement Portion SEM Regression Results

The measurement loadings for the autoregressive NREN model instantiation, as calculated using SmartPLS (Ringle et al., 2022), are tabulated in Table 3. While these loadings did not directly contribute to evaluating the research propositions outlined in Section 5.2.1.2, their thorough analysis was essential, aiming to identify any reflective indicators that failed to meet the minimum Indicator Reliability threshold of 0.4, as Section B.3.1 in Appendix B prescribes. The data presented in Table 3 reflects the final indicator loadings after excluding the *Access to IPv6 Addresses (Y_4)* indicator, identified as unreliable in an initial PLS regression SEM analysis. Its exclusion led to a notable enhancement in Construct Reliability for the relevant latent constructs (see Section B.3.1. in Appendix B).

Table 3: Autoregressive NREN Model Instantiation Indicator Loadings

Autoregressive NREN Model Instantiation Endogenous and Exogenous Constructs	**Construct Type**	**Indicator Type**	**Measurement Indicators**	**Indicator Loadings**
Government Influence of the NREN (ξ_1)	Latent	Reflective	*NREN Governance Mode (X_1)*	λ_{x1} = 0.7402
		Reflective	*Level of Government Funding (X_2)*	λ_{x2} = 0.7487
Scope of the NREN Mandate (ξ_2)	Observable	Reflective	*Range of Institutions the NREN is Mandated to Connect (X_3)*	λ_{x3} = 1.00
NREN Infrastructure Capability (η_1)	Latent	Formative	*Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)*	π_{y1} = 0.3888
		Formative	*Number of Managed Circuits Rented by the NREN (Y_2)*	π_{y2} = 0.9111
NREN Advanced Services Capability (η_2)	Latent	Reflective	*Level of Core Network Traffic (Y_3)*	λ_{y3} = 1.0
		Reflective	*Access to IPv6 Addresses (Y_4)*	Excluded as loading is less than 0.4

Current NREN Reach (η_3)	Observable	Reflective	*Current Number of Institutions Connected by the NREN (Y_5)*	$\lambda_{y5} = 1.0$
Forecasted NREN Reach (η_4)	Observable	Reflective	*Forecasted Number of Institutions Connected by the NREN (Y_6)*	$\lambda_{y6} = 1.0$

B. Structural Portion SEM Regression Results

The structural portion's path coefficients of the autoregressive NREN model instantiation, calculated utilising SmartPLS (Ringle et al., 2022), are catalogued in Table 4. The process of significance testing for these path coefficients, employing asymptotic *t*-statistics, is comprehensively detailed in Section 5.3.1.2.B. These calculated path coefficients, along with their corresponding significance test outcomes, serve as the foundational results for the evaluation of the research propositions enumerated in Section 5.2.1.2.

Table 4: Autoregressive NREN Model Instantiation Path Coefficients

Research Proposition: Autoregressive NREN Model Instantiation SEM Path	**Path Coefficient**
H1: Government Influence of the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	$\gamma_1 = 0.1659$
H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	$\beta_1 = 0.5097$
H3: Government Influence of the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	$\gamma_2 = -0.2073$
H4: NREN Advanced Services Capability (η_2) → Current NREN Reach (η_3)	$\beta_2 = 0.0704$
H5: NREN Advanced Services Capability (η_2) → Forecasted NREN Reach (η_4)	$\beta_3 = -0.0095$
H6: Scope of the NREN Mandate (ξ_2) → Current NREN Reach (η_3)	$\gamma_3 = 0.1586$
H7: Scope of the NREN Mandate (ξ_2) → Forecasted NREN Reach (η_4)	$\gamma_4 = -0.0229$
H8: Current NREN Reach (η_3) → Forecasted NREN Reach (η_4)	$\beta_4 = 0.9964$

5.3.1.2 Autoregressive NREN Model Instantiation Reliability and Validity Analysis Results

The following section presents the reliability and validity test results for the autoregressive NREN model instantiation adheres to the structured reporting guidelines for SEM advocated by Vinzi et al. (2010), refined in Staphorst (2010) and described in Appendix B. The analysis initially focuses on the measurement portion of the autoregressive NREN model instantiation, encompassing an assessment of Indicator Reliability, Construct Reliability, and Convergent Validity, as presented in Staphorst (2010) and described in Section B.3.1 in Appendix B. Subsequently, attention shifts to the structural portion of the TFSEMDF model instantiation. Here, a detailed evaluation of the Coefficients of Determination, the significance of Path Coefficients, and Predictive Validity is considered, as outlined in Staphorst (2010) and explained in Section B.3.2 in Appendix B.

A. Measurement Portion Reliability and Validity Analysis Results

This section considers the reliability and validity test outcomes for the measurement portion of the autoregressive NREN model instantiation of the TFSEMDF framework. This assessment was executed using the SmartPLS software (Ringle et al., 2022) and followed the criteria described by Staphorst (2010). The results, compiled in Table 5, encompass a detailed appraisal of Indicator Reliability alongside the evaluation of Construct Reliability and Convergent Validity. These findings are instrumental in validating the measurement model's robustness and ensuring the SEM constructs are theoretically and statistically representative of the phenomena under study. This level of scrutiny is crucial for affirming the model's measurement accuracy, laying a solid foundation for the subsequent structural analysis within the SEM framework.

The results from the Indicator Reliability tests, conducted as part of the initial PLS regression SEM analysis, indicated that the *Access to IPv6 Addresses (Y_4)* reflective indicator manifested loadings below the threshold of 0.4. Consequently, the study excluded this reflective indicator from all further SEM analyses due to its need for more reliability. Notably, the model maintained all formative indicators irrespective of their loadings, which aligns with the guidelines established by Staphorst et al. (2013, 2014). This decision aligns with the recognition that formative indicators contribute distinct facets to a construct's composition and, thus, their retention is critical for preserving the integrity and comprehensiveness of the model's conceptual structure.

Table 5: Autoregressive NREN Model Instantiation Reliability and Validity – Part A

Autoregressive NREN Model Instantiation Endogenous and Exogenous Constructs	**Measurement Indicators**	**Indicator Reliability Judgement**	**Construct Reliability**		**Convergent Validity**
			Cronbach's Alpha	**Composite Reliability**	
Government Influence of the NREN (ξ_1)	*NREN Governance Mode (X_1)*	Included	0.1958	$\rho_{\xi,1}$ = 0.7132	$AVE_{\xi,1}$ = 0.5542
	Level of Government Funding (X_2)	Included			
Scope of the NREN Mandate (ξ_2)	*Range of Institutions the NREN is Mandated to Connect (X_3)*	Included	Tests not applicable: This construct is directly observable (Staphorst et al., 2013, 2014)		
NREN Infrastructure Capability (η_1)	*Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)*	Included	Tests not applicable: This construct is directly observable and has formative indicators (Staphorst et al., 2013, 2014)		
	Number of Managed Circuits Rented by the NREN (Y_2)	Included			
NREN Advanced Services Capability (η_2)	*Level of Core Network Traffic (Y_3)*	Included	Tests not applicable: While the definition of the construct was latent, excluding Y_4 led to treating it as directly observable (Staphorst et al., 2013, 2014)		
	Access to IPv6 Addresses (Y_4)	Excluded			

Current NREN Reach (η_3)	*Current Number of Institutions Connected by the NREN (Y_5)*	Included	Tests not applicable: This construct is directly observable (Staphorst et al., 2013, 2014)
Forecasted NREN Reach (η_4)	*Forecasted Number of Institutions Connected by the NREN (Y_6)*	Included	Tests not applicable: This construct is directly observable (Staphorst et al., 2013, 2014)

The assessment of Construct Reliability within this study incorporated both the traditional Cronbach's Alpha metric and the more recent Composite Reliability measure, as outlined in the methodologies of Staphorst et al. (2013, 2014). The final determination regarding the adequacy of the reflective indicators in measuring their corresponding latent constructs was contingent upon the Composite Reliability measure surpassing a minimum threshold of 0.6, a criterion set forth by Vinzi et al. (2010). As detailed in Table 4, examination of the results reveals that post the exclusion of unreliable indicators, the sole latent construct with reflective indicators remaining in the model was the *Government Influence of the NREN (ξ_1)*. This construct successfully met the established requirement for Composite Reliability, indicating its robustness in reliably representing the construct within the autoregressive NREN model instantiation. This compliance with the Composite Reliability threshold underscores the construct's validity in capturing the essence of government influence on the NREN.

The evaluation of Convergent Validity in this study was conducted using the Average Variance Extracted (AVE) metric, as described by Staphorst et al. (2013, 2014). This metric quantifies the proportion of variance in the reflective indicators of each latent construct that is attributable to the construct itself relative to the total variance measured. Applying the study's predetermined threshold of 0.5 for the AVE metric, an analysis of the results presented in Table 5 indicates that the reflective indicators for the sole remaining latent construct, *Government Influence of the NREN (ξ_1)*, demonstrated an adequate level of AVE. This finding signifies that the construct variance, rather than measurement error, accounts for the predominant portion of the total variance observed in these indicators. Such a result confirms the high convergent validity of the

Government Influence of the NREN (ξ_1) construct, affirming that it effectively captures the variance of its indicators and thus reliably represents the intended concept in the model.

The assessment of Discriminant Validity within the autoregressive NREN model instantiation was deemed unnecessary as it hinges on the premise that the square root of the AVE for each latent construct should surpass its correlation with all other latent constructs, as illustrated by Staphorst et al. (2013, 2014). In the case of this model instantiation, the only remaining latent construct with reflective indicators post the removal of the unreliable *Access to IPv6 Addresses (Y_4)* indicator was the *Government Influence of the NREN (ξ_1)*. In this unique instance, the conventional criteria for evaluating Discriminant Validity became inapplicable because there were no other latent constructs against which to compare the *Government Influence of the NREN (ξ_1)* construct.

B. Structural Portion Reliability and Validity Test Results

This section presents the results of the reliability and validity tests for the structural portion of the autoregressive NREN model instantiation, aligning with the metrics outlined by Staphorst et al. (2013, 2014). The analysis executed using the SmartPLS software (Ringle et al., 2022) comprehensively evaluates the model's structural integrity and predictive efficacy. Table 6 outlines the results of the Path Coefficient tests, providing insights into the strength and significance of the relationships between various constructs within the model. In contrast, Table 7 showcases the outcomes of the Coefficients of Determination and Predictive Validity tests.

Table 6: Autoregressive NREN Model Instantiation Reliability and Validity – Part B

Research Proposition: Autoregressive NREN Model Instantiation SEM Path	**Asymptotic *t*-Statistic**	**Calculated *p*-Value**	**Significance Judgement for $\alpha = 0.10$**
H1: Government Influence of the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	1.2926	0.196	No
H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	3.3564	0.001	Yes
H3: Government Influence of the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	0.8783	0.380	No

H4: NREN Advanced Services Capability (η_2) → Current NREN Reach (η_3)	1.5468	0.122	No
H5: NREN Advanced Services Capability (η_2) → Forecasted NREN Reach (η_4)	1.3593	0.174	No
H6: Scope of the NREN Mandate (ξ_2) → Current NREN Reach (η_3)	2.3155	0.021	Yes
H7: Scope of the NREN Mandate (ξ_2) → Forecasted NREN Reach (η_4)	2.3489	0.019	Yes
H8: Current NREN Reach (η_3) → Forecasted NREN Reach (η_4)	39.0519	0.000	Yes

The Path Coefficient Significance test results from Table 5, derived using the bootstrapping function of SmartPLS (Ringle et al., 2022) with a resampling size of 1000, provide crucial insights. These results, calculated using the $t_{(999)}$ asymptotic t-statistic distribution, indicate that several paths did not achieve statistical significance at the defined maximum acceptable level of $\alpha = 0.10$, per the guidelines in Staphorst et al. (2013, 2014). The paths that yielded p-values exceeding this threshold and, thus, were considered statistically insignificant include:

- *Government Influence of the NREN (ξ_1)* to *NREN Infrastructure Capability (η_1)*.
- *Government Influence of the NREN (ξ_1)* to *NREN Advanced Services Capability (η_2)*.
- *NREN Advanced Services Capability (η_2)* to *Current NREN Reach (η_3)*.
- *NREN Advanced Services Capability (η_2)* to *Forecasted NREN Reach (η_4)*.

Table 7: Autoregressive NREN Model Instantiation Reliability and Validity – Part C

Autoregressive NREN Model Instantiation Technology or Context-Related Indicator	**Coefficients of Determination (R^2)**	**Predictive Validity (Q^2)**	
		Cross-validated Communality (H^2)	**Cross-validated Redundancy (F^2)**
Government Influence of the NREN (ξ_1)	Test not applicable: Exogenous variable	0.5542	0

Scope of the NREN Mandate (ξ_2)	Test not applicable: Exogenous variable	1.0	0
NREN Infrastructure Capability (η_1)	0.0275	0.4906	0.0137
NREN Advanced Services Capability (η_2)	0.2677	1.0	0.0079
Current NREN Reach (η_3)	0.0287	1.0	0.0035
Forecasted NREN Reach (η_4)	0.9852	1.0	-0.0011

The Coefficients of Determination test results, as detailed in Table 7 and aligned with the methodologies of Staphorst et al. (2013, 2014), reveal varied strengths in the explanatory power of the autoregressive NREN model instantiation's interrelationships. Specifically, the interactions involving *NREN Infrastructure Capability (η_1)* and *Current NREN Reach (η_3)* yielded explained variances below the minimum threshold of 10%, indicating limited explanatory power. In contrast, the interrelationships with *Forecasted NREN Reach (η_4)* demonstrated robustness, with an R^2 value exceeding 0.7, signifying high explanatory capacity. However, the connections with *NREN Advanced Services Capability (η_2)* appeared weak, as evidenced by an R^2 value below 0.3, suggesting a lower level of explained variance in this construct. These results provide a nuanced understanding of the model's structural efficacy, highlighting the varying degrees of influence among the endogenous latent constructs.

The Predictive Validity test results for the *Forecasted NREN Reach (η_4)* construct, as assessed through the TF output metric *Forecasted Number of Institutions Connected by the NREN (Y_6),* indicated that the Cross-validated Communality (H^2) was positive. At the same time, the Cross-validated Redundancy (F^2) was less than zero. The results suggest that although the measurement indicators of the model effectively capture the construct of *Forecasted NREN Reach (η_4)*, the structural relationships within the model do not robustly contribute to its predictive accuracy. Therefore, the model's measurement indicators are suitable for forecasting future NREN reach, but the structural relationships require further refinement to enhance their predictive effectiveness.

5.3.2 Cross-Sectional NREN Model Instantiation

The following section presents the PLS regression analysis, including reliability and validity assessment, for the cross-sectional NREN model instantiation. As explained in Table 1, this analysis employed indicator data from the 2011 TERENA NREN Compendium (TERENA, 2011).

Section 5.2.2.3 highlights that the sample size of NRENs participating in the 2011 TERENA NREN Compendium survey (TERENA, 2011) was 61, with only 28 responding NRENs providing the complete dataset required by the indicator data composition approach defined in Table 2. As explained above and investigated in Section 6.3.1.3, one of the advantages of PLS-SEM is its adaptability to accommodate modest sample sizes. Furthermore, recall the widely acknowledged convention that PLS regression for SEM mandates a sample size that is at least ten times the number of endogenous or exogenous formative indicators linked to the most intricate latent construct within the model (Chin & Newsted, 1999; Goodhue et al., 2006). For the cross-sectional NREN model instantiation portrayed in Figure 8, the construct denoted as *NREN Infrastructure Capability* (η_1) exhibits the most complex formative structure, encompassing two indicators: *Length of Dark Fibre Infrastructure Owned by the NREN* (Y_1) and *Number of Managed Circuits Rented by the NREN* (Y_2). Consequently, applying PLS regression for this model instantiation was successful as the constructed indicator data set surpassed the calculated minimum sample size requirement of 20.

5.3.2.1 Cross-Sectional NREN Model Instantiation SEM Regression Results

The presentation of the PLS regression results for the cross-sectional NREN model instantiation within the TFSEMDF framework, as detailed in the ensuing sections, adheres to the reporting guidelines established by Vinzi et al. (2010) and refined in Staphorst (2010). Following these guidelines, the results are organised systematically into two primary categories. The first category addresses the measurement component of the SEM path diagram, detailing the loadings of all measurement indicators within the cross-sectional NREN model instantiation. Next follows the category dedicated to the structural component of the SEM path diagram, which includes the path coefficients representing the interrelationships between the various constructs.

A. Measurement Portion SEM Regression Results

Table 8 systematically documents the indicator loadings for the measurement component of the cross-sectional NREN model instantiation of the TFSEMDF framework, as calculated using SmartPLS (Ringle et al., 2022). While these loadings did not directly influence the evaluation of the research propositions outlined in Section 5.2.2.2, their thorough examination was essential. This detailed analysis focused on identifying any reflective indicators that fell short of the minimum Indicator Reliability threshold of 0.4, a critical benchmark detailed in Section 5.3.2.2.A.

Table 8: Cross-Sectional NREN Model Instantiation Indicator Loadings

Cross-Sectional NREN Model Instantiation Constructs	Type	Measurement Indicators	Indicator Loadings
Government Influence over the NREN (ξ_1)	Reflective	*NREN Governance Mode (X_1)*	$\lambda_{x1} = 0.892$
	Reflective	*Level of Government Funding (X_2)*	$\lambda_{x2} = 0.805$
	Reflective	*Range of Institutions the NREN is Mandated to Connect (X_3)*	$\lambda_{x3} = 0.854$
NREN Infrastructure Capability (η_1)	Formative	*Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)*	$\pi_{y1} = 0.473$
	Formative	*Number of Managed Circuits Rented by the NREN (Y_2)*	$\pi_{y2} = 0.737$
NREN Advanced Services Capability (η_2)	Reflective	*NREN Advanced Services Capability (Y_3)*	$\lambda_{y3} = 1.0$
NREN Core Traffic Level (η_3)	Reflective	*NREN Core Traffic Level (Y_5)*	$\lambda_{y4} = 1.0$

B. Structural Portion SEM Regression Results

The path coefficients for the structural portion of the cross-sectional NREN model instantiation of the TFSEMDF framework, calculated using the SmartPLS software (Ringle et al., 2022), are comprehensively detailed in Table 9. The evaluation employed asymptotic t-statistics to assess the significance of these path coefficients, with an exposition of this testing process and its results available in Section 5.3.2.2.B. Subsequently, these path coefficients, along with their corresponding significance levels, played a crucial role in the assessment of the research propositions as delineated in Section 5.2.2.2. This assessment, elaborated in Section 5.4.2, involved a detailed analysis of the path coefficients and their significance, aligning the empirical findings with the theoretical hypotheses within the model.

Table 9: Cross-Sectional NREN Model Instantiation Path Coefficients

Research Proposition: Cross-Sectional NREN Model Instantiation SEM Path	**Path Coefficient**
H1: Government Influence over the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	$\gamma_1 = 0.599$
H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	$\beta_1 = 0.016$
H3: Government Influence over the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	$\gamma_2 = 0.855$
H4: NREN Infrastructure Capability (η_1) → NREN Core Traffic Level (η_3)	$\beta_2 = 0.289$
H5: NREN Advanced Services Capability (η_2) → NREN Core Traffic Level (η_3)	$\beta_3 = 0.187$

5.3.2.2 Cross-Sectional NREN Model Instantiation Reliability and Validity Analysis Results

Following the reporting guidelines suggested by Vinzi et al. (2010) for SEM, this study structures the presentation of reliability and validity test results for the cross-sectional NREN model instantiation of the TFSEMDF framework into two categories. Initially, the focus is on the measurement portion, encompassing assessments of Indicator Reliability, Construct Reliability, and Convergent Validity as recommended by Staphorst et al. (2013, 2017) and detailed in Appendix B. This initial effort ensures that the measurement indicators within the model accurately represent the constructs they intend to measure.

Subsequently, the analysis shifts to the structural portion of the model, including evaluations of Coefficients of Determination, Path Coefficient Significance, and Predictive Validity. The logic underpinning this approach, as highlighted by Vinzi et al. (2010), posits that the credibility of the structural portion of a model is heavily dependent on the accuracy and representativeness of the measurement indicators. In essence, if the measurement indicators are not accurate or representative, it undermines the purpose and necessity of assessing the structural portion's reliability and validity.

A. Measurement Portion Reliability and Validity Analysis Results

This section elaborates on the reliability and validity test results for the measurement portion of the cross-sectional NREN model instantiation of the TFSEMDF framework. Employing the SmartPLS software (Ringle et al., 2022), these tests were conducted in line with the metrics and guidelines defined by Vinzi et al. (2010) and further detailed by Staphorst (2010). The findings, encapsulated in Table 10, include a detailed evaluation of Indicator Reliability, Construct Reliability, and Convergent Validity, as outlined in Staphorst (2010).

Table 10: Cross-Sectional NREN Model Instantiation Reliability and Validity - Part A

Cross-Sectional NREN Model Instantiation Endogenous and Exogenous Constructs	**Measurement Indicators**	**Indicator Reliability Judgment**	**Construct Reliability**		**Convergent Validity**
			Cronbach's Alpha	**Composite Reliability**	
Government Influence over the NREN (ξ_1)	*NREN Governance Mode (X_1)*	Included	0.810	$\rho_{\xi,1} = 0.887$	$AVE_{\xi,1} = 0.724$
	Level of Government Funding (X_2)	Included			
	Range of Institutions the NREN is Mandated to Connect (X_3)	Included			
NREN Infrastructure Capability (η_1)	*Length of Dark Fibre Infrastructure Owned by the NREN (Y_1)*	Included	Tests not applicable: This construct has formative indicators (Staphorst, 2010)		

	Number of Managed Circuits Rented by the NREN (Y_2)	Included	
NREN Advanced Services Capability (η_2)	*NREN Advanced Services Capability (Y_3)*	Included	Tests not applicable: This construct is directly observable (Staphorst, 2010)
NREN Core Traffic Level (η_3)	*NREN Core Traffic Level (Y_4)*	Included	Tests not applicable: This construct is directly observable (Staphorst, 2010)

The Indicator Reliability test, which assesses the extent to which the variance in a measurement indicator is attributable to its associated latent construct, was conducted as part of the initial PLS regression analysis of the cross-section NREN model instantiation (Staphorst et al., 2016a, 2017). The results indicated that all reflective indicators had loadings exceeding the threshold of 0.4, signifying that they adequately represented their respective latent constructs. Consequently, there was no need to eliminate any reflective indicators, allowing all subsequent SEM analyses to proceed with the model instantiation as initially conceptualised.

Regarding the formative indicators, the model retained all regardless of their loadings. This decision aligns with the understanding that Indicator Reliability is not pertinent to formative indicators, as elucidated by Vinzi et al. (2010). In the context of formative indicators, low correlations with their associated latent constructs do not necessarily diminish their contribution. Instead, they can still significantly influence the overall variance of the constructs, underscoring the unique role and interpretation of formative indicators in SEM analysis. This approach ensures a holistic and appropriate treatment of different indicators within the model, enhancing its theoretical and empirical coherence.

Construct Reliability, a key metric in assessing the efficacy of reflective indicators in jointly measuring a latent construct, was evaluated using the traditional Cronbach's Alpha (Vinzi et al., 2010) and the more recent Composite Reliability measure (Fornell & Larcker, 1981). For the cross-sectional NREN model instantiation, the criterion for determining the adequacy of the reflective indicators in representing their respective latent constructs required that the Composite Reliability score surpass a minimum threshold of 0.6, as stipulated by Vinzi et al. (2010). An examination of the results in Table 10 reveals that the sole latent construct in the model with reflective indicators was *Government Influence over the NREN* (ξ_1). This construct satisfactorily met the specified Composite Reliability threshold, as per the guidelines by Fornell & Larcker (1981), affirming the construct's robustness in terms of reliability. This adherence to the Composite Reliability criterion underscores the construct's validity in capturing the essence of government influence on the NREN within the model's framework.

Convergent Validity, a concept that assesses the degree of correlation among different methods of measuring the same construct, is critical in validating a model's constructs (Vinzi et al., 2010). This validity is quantified using the AVE metric, as outlined (Staphorst, 2010; Fornell & Larcker, 1981). The AVE metric evaluates the proportion of variance in a latent construct's reflective indicators attributable to the construct itself, compared to the total variance observed. For the cross-sectional NREN model instantiation, the AVE metric's threshold was set at 0.5, as recommended by Staphorst (2010). Analysis of the results in Table 10 shows that the reflective indicators of the sole latent construct in the model, *Government Influence over the NREN* (ξ_1), achieved an AVE level that met this threshold, indicating that most of the variance captured by this construct's reflective indicators results from the construct's variance rather than measurement error.

Discriminant validity within the context of SEM pertains to the distinctiveness of measurements obtained by a measurement tool for different constructs (Vinzi et al., 2010). It primarily assesses whether a latent construct is sufficiently distinct from other constructs within the model. A fundamental criterion for achieving Discriminant Validity is that the shared variance between a latent construct and its indicators, calculated by taking the square root of its AVE, should be greater than the shared variance between that latent construct and any other latent constructs in the model. However, in the case of the cross-sectional NREN model instantiation within the TFSEMDF framework, the situation is unique because there is only one latent construct with reflective indicators, namely *Government Influence over the NREN* (ξ_1). Given this singular construct with reflective indicators, the requirement for conducting a Discriminant Validity test becomes redundant. Without multiple latent constructs to compare against, the criterion of shared variance exceeding that between different constructs cannot be applied. Therefore, the model's structure inherently satisfies the condition for Discriminant Validity by default, rendering a separate test for this validity measure unnecessary in this specific instantiation.

B. Structural Portion Reliability and Validity Analysis Results

This section details the results of the reliability and validity tests for the structural component of the cross-sectional NREN model instantiation of the TFSEMDF framework. These results were derived using the methodologies defined by Vinzi et al. (2010) and further detailed in Staphorst (2010), with the analyses performed via the SmartPLS software (Ringle et al., 2022). Table 11 presents the results of the Path Coefficient significance tests, providing an in-depth examination of the statistical significance of the relationships between the constructs in the model instantiation. Complementarily, Table 12 offers insights into the Coefficients of Determination and Predictive Validity test results.

Table 11: Cross-Sectional NREN Model Instantiation Reliability and Validity - Part B

Research Proposition: Cross-Sectional NREN Model Instantiation SEM Path	**Asymptotic *t*-Statistic**	**Calculated *p*-Value**	**Significance Judgment for $\alpha = 0.10$**
H1: Government Influence over the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	3.952	< 0.001	Yes
H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	0.124	0.901	No
H3: Government Influence over the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	4.840	< 0.001	Yes
H4: NREN Infrastructure Capability (η_1) → NREN Core Traffic Level (η_3)	2.015	0.044	Yes
H5: NREN Advanced Services Capability (η_1) → NREN Core Traffic Level (η_3)	1.662	0.097	Yes

In SEM, like covariance-based multiple regression techniques, the structural quality of a model instantiation is critically assessed through a bootstrapping procedure, as suggested by Vinzi et al. (2010). This approach is essential for determining the significance levels of the path coefficients, a key element in evaluating the model's structural integrity. Specifically, assessing the significance of these path coefficients (Goodness-of-Fit) uses asymptotic *t*-statistics. The Path

Coefficient Significance test results, as shown in Table 11, were obtained using the bootstrapping function in SmartPLS (Ringle et al., 2022), configured to generate 500 subsamples from the 61 cases in the original sample. According to these results, the only path with a *p*-value exceeding the maximum acceptable significance level of $\alpha = 0.10$ was the path from *NREN Infrastructure Capability (η_1)* to *NREN Advanced Services Capability (η_2)*.

Consequently, the analysis deemed this path statistically insignificant. This finding indicates that the hypothesised influence of *NREN Infrastructure Capability (η_1)* on *NREN Advanced Services Capability (η_2)* within the studied model did not meet the statistical threshold for a significant relationship, highlighting that the model's predicted dynamics do not align with the empirical data.

Table 12: Cross-Sectional NREN Model Instantiation Reliability and Validity - Part C

Cross-Sectional NREN Model Instantiation Technology or Context-Related Indicator	**Coefficients of Determination (R^2)**	**Predictive Validity (Q^2)**	
		Cross-validated Communality (H^2)	**Cross-validated Redundancy (F^2)**
Government Influence of the NREN (ξ_1)	Test not applicable: Exogenous variable	0.440	Test not applicable: Exogenous variable
NREN Infrastructure Capability (η_1)	0.359	0.054	0.191
NREN Advanced Services Capability (η_2)	0.749	1.0	0.727
NREN Core Traffic Level (η_3)	0.175	1.0	0.142

In the analysis of the Coefficients of Determination (R^2) for the structural portion of the cross-sectional NREN model instantiation, as per Vinzi et al. (2010), the results from Table 12 indicated that all endogenous latent constructs and their associated constructs in the model exceeded the minimum explanatory variance threshold of 0.1, as outlined by Staphorst (2010) and Staphorst et al. (2017). Specifically, the relationship with the *NREN Advanced Services Capability (η_2)* construct was notably strong, evidenced by an R^2 value exceeding 0.7, suggesting a high proportion of its variance explained by related constructs. In contrast, the relationship involving the *NREN Core Traffic Level* (η_3) construct was categorised as weaker, with an R^2 value lower than 0.3, indicating less explained variance.

To ascertain the Predictive Validity of a model instantiation, researchers perform the Stone-Geisser (referred to as Q^2) non-parametric test (Staphorst, 2010; Staphorst et al., 2017; Vinzi et al., 2010) based on a blindfolding procedure (Zikmund, 2009). The model instantiation has Predictive Validity if $Q^2 > 0$ (Vinzi et al., 2010). The Stone–Geisser test criterion can take on two distinct forms, depending on the type of prediction investigated: The first form, geared at determining the Predictive Validity of the measurement portion (although usually calculated during the structural portion's validity evaluation), is referred to as the Cross-validated Communality (Vinzi et al., 2010) and is denoted by H^2. Cross-validated Communality measures the capacity to predict the observable endogenous constructs from their latent construct scores (Vinzi et al., 2010). The second form, which evaluates the Predictive Validity of the structural portion, is referred to as Cross-validated Redundancy (Vinzi et al., 2010). This metric, denoted by F^2, measures the model instantiation's ability to predict the observable endogenous constructs using latent constructs that predict the data block (Vinzi et al., 2010).

The evaluation of Predictive Validity in a model instantiation, crucial for assessing its forecasting accuracy, is conducted using the Stone-Geisser (Q^2) non-parametric test, which employs a blindfolding procedure as outlined by Staphorst (2010), Staphorst et al. (2017), and Vinzi et al. (2010), with Zikmund (2009) providing the methodological context. For a model to be considered predictively valid, its Q^2 value must exceed 0. This test encompasses two key metrics: Cross-validated Communality (H^2) and Cross-validated Redundancy (F^2). The former evaluates the measurement portion of the model by measuring its ability to predict observable endogenous constructs from latent construct scores. The latter assesses the structural portion by gauging the model's effectiveness in predicting observable endogenous constructs using related latent constructs.

An examination of the Predictive Validity test results for the *NREN Core Traffic Level* (η_3) construct, measurable through the output forecasting technology metric *NREN Core Traffic Level* (Y_4), demonstrated positive outcomes for both Cross-validated Communality (H^2) and Cross-validated Redundancy (F^2). This finding indicates that the measurement indicators and the defined structural relationships within the cross-sectional NREN model instantiation are effectively aligned to accurately forecast an NREN's core network traffic level. The positive test results for H^2 and F^2 confirm that the model is proficient in capturing the variance within the construct (as indicated by H^2) and using the latent constructs to predict the observed values (as denoted by F^2). As Staphorst et al. (2016a) highlighted, this dual positive outcome underscores the model's robustness regarding predictive capabilities, particularly in forecasting an NREN's core network traffic levels.

5.4 DISCUSSION

The forthcoming sections discuss the results obtained during this study phase. Section 5.4.1 discusses the results of analysing the autoregressive NREN model instantiation in the TFSEMDF framework, pertinent to Research Objective 2(a). Conversely, Section 5.4.2 discusses the results from studying the cross-sectional NREN model instantiation in the TFSEMDF framework, aiming to fulfil Research Objective 2(b).

5.4.1 AUTOREGRESSIVE NREN MODEL INSTANTIATION

Utilising the path coefficients detailed in Table 4 and the results from the path coefficient significance tests found in Table 6, the research propositions outlined for the Autoregressive NREN model instantiation in Section 5.2.1.2 underwent a comprehensive evaluation. This process involved an analysis of the path coefficients' directionality and evaluating their significance levels, thereby determining the extent to which the empirical data supported each research proposition.

- **Research Proposition H1:** As proposed, the path coefficient $\gamma_1 = 0.1659$ suggests an alignment in the direction between the *Government Influence of the NREN* (ξ_1) construct and the *NREN Infrastructure Capability* (η_1) construct. However, its statistical insignificance at $\alpha = 0.10$ led to rejecting this hypothesised relationship.
- **Research Proposition H2:** The relationship between *NREN Infrastructure Capability* (η_1) and *NREN Advanced Services Capability* (η_2) was supported, with path coefficient $\beta_1 = 0.5097$, indicating both the proposed direction and statistical significance at $\alpha = 0.10$.
- **Research Proposition H3:** The path coefficient $\gamma_2 = -0.2073$ contradicts the hypothesised direction of the relationship between *Government Influence of the NREN* (ξ_1) and *NREN Advanced Services Capability* (η_2), resulting in the rejection of this proposition.

- **Research Proposition H4:** Despite aligning with the hypothesised direction, the path coefficient $\beta_2 = 0.0704$ for the relationship between *NREN Advanced Services Capability (η_2)* and *Current NREN Reach (η_3)* failed to reach statistical significance at $\alpha = 0.10$, leading to its rejection.
- **Research Proposition H5:** The negative path coefficient $\beta_3 = -0.0095$ contradicts the proposed direction between *NREN Advanced Services Capability (η_2)* and *Forecasted NREN Reach (η_4)*, thereby rejecting this research proposition.
- **Research Proposition H6:** The path coefficient $\gamma_3 = 0.1586$, significant at $\alpha = 0.10$, aligns with the hypothesised positive relationship between the *Scope of the NREN Mandate (ξ_2)* and *Current NREN Reach (η_3)*, leading to the acceptance of this proposition.
- **Research Proposition H7:** The analysis rejects the relationship between the *Scope of the NREN Mandate (ξ_2)* and *Forecasted NREN Reach (η_4)* because the path coefficient $\gamma_4 = -0.0229$ does not align with the direction of the proposed hypothesis.
- **Research Proposition H8:** The analysis upholds the proposition with a path coefficient $\beta_4 = 0.9964$, determined to be significant at $\alpha = 0.10$, which supports the hypothesised direction between *Current NREN Reach (η_3)* and *Forecasted NREN Reach (η_4)*.

Analysing the eight research propositions for the autoregressive NREN model instantiation yielded a blend of supported and unsupported hypotheses, reflecting the model's varied longitudinal predictive strength across different constructs. Most importantly, Proposition H8, exploring the relationship between current and forecasted NREN reach, was upheld, demonstrating the model's power in longitudinal forecasting this TF output metric. Secondary, but also necessary, Propositions H2 and H6, focusing on the relationships between NREN infrastructure capability and advanced services capability and between the scope of the NREN mandate and current NREN reach, respectively, were also supported, affirming the model's capacity in these aspects.

Conversely, Propositions H1, H3, H4, H5, and H7, which involved the influence of government on NREN orientation and services and the impact of advanced services on current and forecasted NREN reach, were not substantiated, indicating areas where the model's predictive capability was less effective. This finding likely stems from limitations in the TERENA (2011) indicator data used, although the foundational grounds for these relationships receive support in the literature (Greaves, 2009; Janz & Kutanov, 2012).

5.4.2 Cross-Sectional NREN Model Instantiation

Employing the path coefficients outlined in Table 9 and their corresponding significance test results presented in Table 11, the cross-sectional NREN model instantiation's research propositions, as detailed in Section 5.2.2.2, underwent a comprehensive evaluation. This process entailed examining the path coefficients' compliance with the required significance level and the polarity of the calculated path coefficients, thereby determining the empirical support for each research proposition within the cross-sectional NREN model instantiation of the TFSEMDF framework.

- **Research Proposition H1:** The path coefficient $\gamma_1 = 0.599$ aligns with the hypothesised correlation between *Government Influence over the NREN (ξ_1)* and *Infrastructure Capability (η_1)* and is significant at $\alpha = 0.10$. This result supports the association, as discussed by Greaves (2009) and Janz & Kutanov (2012), between government influence and developing an NREN's infrastructure capability, consistent with the typical government role in enhancing national research and educational infrastructure and services.
- **Research Proposition H2:** The hypothesised correlation between *NREN Infrastructure Capability (η_1)* and *NREN Advanced Services Capability (η_2)* was not supported because the path coefficient $\beta_1 = 0.016$ lacked significance at $\alpha = 0.10$, despite aligning with the proposed direction. This outcome diverges from the correlation suggested by Greaves (2009). It may reflect trends in the telecommunications industry, as described by Mbarika et al. (2000), where providers in developing countries often offer advanced services on less developed infrastructures, as seen for European NRENs supporting emerging NRENs in Africa (GÉANT, 2022).
- **Research Proposition H3:** The significant path coefficient $\gamma_2 = 0.855$, significant at $\alpha = 0.10$, corroborated the correlation between Government *Influence over the NREN (ξ_1)* and *NREN Advanced Services Capability (η_2)*. This finding aligns with the observations of Greaves (2009) and Janz and Kutanov (2012) regarding the relationship between government influence and the enhancement of an NREN's advanced services portfolio.
- **Research Proposition H4:** The relationship between *NREN Infrastructure Capability (η_1)* and *NREN Core Traffic Level (η_3)* was supported, with path coefficient $\beta_2 = 0.289$ significant at $\alpha = 0.10$. This finding aligns with Savory's (2012) observation of a positive correlation between an NREN's infrastructure capability and usage. It is consistent with

the notion that enhanced infrastructure leads to increased use, as GÉANT (2022) and Greaves (2009) discussed.

- **Research Proposition H5:** The path coefficient $\beta_3 = 0.187$, significant at $\alpha = 0.10$, supports the hypothesised correlation between *NREN Advanced Services Capability (η_2)* and *NREN Core Traffic Level (η_3)*, suggesting a positive relationship between an NREN's advanced services capability and its usage, resonating with insights from Greaves (2009) and Janz and Kutanov (2012), and the concept that a diverse range of advanced services can lead to increased utilisation of NREN infrastructure (GÉANT, 2022; Greaves, 2009).

The analysis using the 2011 TERENA NREN Compendium data essentially validated the hypothesised relationships within the cross-sectional NREN model instantiation of the TFSEMDF framework, except for the correlation between an NREN's advanced services capability and infrastructure capability. This finding aligns with technology leapfrogging (Mbarika et al., 2000), frequently observed in the NREN community (GÉANT, 2022).

Notably, the relationships involving the NREN core traffic level received robust support, demonstrating that the selected input, byput, and context-related metrics effectively forecast the output TF metric, which is critical for assessing the NREN's usage. The model instantiation's strength in transversally predicting the NREN core traffic level is a key finding, as it highlights the model's capacity to provide insightful cross-sectional predictions, essential for understanding current dynamics and making informed decisions within an NREN organisation.

5.5 Concluding Remarks

Chapter 5 was dedicated to the second phase of this study and delved into applying the TFSEMDF framework in the NREN technology domain. Its introduction links the framework's theoretical underpinnings, detailed in Chapter 4, to its practical application in this chapter, setting the stage for the empirical exploration of the TFSEMDF framework in the NREN technology domain.

The chapter dissected the research methodologies of the study's second phase into two distinct investigations. The initial inquiry delved into applying the TFSEMDF framework to the autoregressive NREN model instantiation proposed by Staphorst et al. (2013) and analysed in

Staphorst et al. (2014). The subsequent inquiry evaluated the cross-sectional NREN model instantiation developed and investigated by Staphorst et al. (2016a). This dual-pronged approach exhibits the TFSEMDF framework's ability to achieve transversal and longitudinal forecasting objectives.

The results section of the chapter conveyed the findings from the PLS-SEM regression analysis process, described in Appendix A, as applied to the two investigations constituting the study's second phase, as described above. PLS regression results for the autoregressive NREN model instantiation of the TFSEMDF framework are presented, including a detailed examination of the results' reliability and validity as verified by the criteria outlined in Appendix B. The analysis then transitioned to the PLS-SEM regression results for the cross-sectional NREN model instantiation, where the framework's reliability and validity were scrutinised and validated.

The discussion section examined the outcomes of the PLS-SEM regression analysis for the TFSEMDF framework as applied to the two separate NREN model instantiations. The chapter critiqued the implications of the regression results and the reliability and validity assessments detailed earlier. This critique provided a multifaceted view of the framework's analytical capacity, considering not only the statistical significance of the findings. Section 7.2.2 presents the research conclusions from the second phase of the study. This Section delineates the final judgments regarding the application and utility of the TFSEMDF framework in the NREN technology domain, as derived from the research outcomes documented in this chapter.

The cross-sectional NREN model instantiation of the TFSEMDF framework developed in this chapter is employed in Chapter 6 as baseline NREN model instantiation and altered to create a structurally disarranged NREN model instantiation. These model instantiations are used in the third research phase to evaluate the strengths and weaknesses of the TFSEMDF framework.

CHAPTER 6 – RESEARCH PHASE 3: TFSEMDF FRAMEWORK ASSESSMENT

6.1 INTRODUCTION

This chapter marks the transition into the third and final phase of the study, concentrating on the thorough assessment of the TFSEMDF framework. This evaluative stage addressed two pivotal research objectives: Research Objective 3(a) focused on evaluating the framework's strengths, while Research Objective 3(b) dedicated efforts to identifying and examining its potential weaknesses. This dual approach is pivotal in providing a holistic understanding of the framework's operational effectiveness and areas requiring further development.

The chapter begins by outlining the methodologies employed to evaluate the inherent strengths and weaknesses of the TFSEMDF framework, as undertaken in Research Objectives 3(a) and 3(b), respectively. After presenting these methodologies, the chapter moves into a detailed examination of the findings, methodically outlining the outcomes of these rigorous evaluations. The discussion then advances to thoroughly analysing these results, assessing the framework's effectiveness and identifying areas needing enhancement.

This chapter articulates methodologies, results, and ensuing discussions that expand and refine the concepts Staphorst et al. (2016b) initially presented. This publication laid a comprehensive foundation, which significantly informed and guided the evaluations conducted in this chapter. Its contribution was pivotal in shaping a robust and well-informed analysis, ensuring that the current chapter builds upon and extends the initial findings with added depth and clarity.

6.2 RESEARCH METHODOLOGY

The following sections present the research methodologies employed in this study phase. Section 6.2.1 examines the methods applied in assessing the strengths of the TFSEMDF framework, aligning with Research Objective 3(a). Conversely, Section 6.2.2 explores the methodology for evaluating the framework's weaknesses, addressing Research Objective 3(b).

6.2.1 Strengths of the TFSEMDF Framework

6.2.1.1 Methodology to Identify and Verify TFSEMDF Strengths

A structured two-step process accomplished Research Objective 3(a), focusing on the identification and empirical analysis of the inherent strengths within the TFSEMDF framework. The first step involved conducting an expansive literature review, as presented in Chapter 2, to explore the intrinsic strengths of SEM and context-sensitive DF. This literature review incorporated thematic analysis, following the method outlined by Braun & Clarke (2006), enabling a systematic identification and interpretation of themes and patterns indicative of the strengths inherent to SEM and context-sensitive DF (Ryan & Bernard, 2003).

The second step in this process utilised the cross-sectional NREN model instantiation, as constructed in Section 5.2.2.1, and then extensively analysed in Section 5.3.2.1. The study employed this model instantiation as a baseline NREN model instantiation for empirical validation. This verification aimed to ascertain if the inherent strengths identified through the literature review and thematic analysis of SEM and context-sensitive DF were also evident within the proposed TFSEMDF framework. As detailed in Section 6.3.1, the strengths identified and considered in this analysis included SEM-based DF's ability to incorporate context-related information, SEM's ability to model complex hierarchal interrelationships and PLS-SEM's ability to perform modelling with small sample sizes.

In the context of SEM model instantiations where covariance-based regression analysis is employed, a range of reliability and validity metrics, from Coefficients of Determination to Predictive Validity, are typically used for comparative purposes (Staphorst, 2010). However, Evermann and Tate (2010) noted that these metrics are not suitable for comparing model instantiations when employing PLS regression. Consequently, this study adopted the alternative approach Evermann and Tate (2010) suggested to assess the impact of including context-related information through SEM-based DF. This approach entailed evaluating the baseline NREN model instantiation and its variations, which differed in the levels of context-related information included, by analysing the presence or absence of SEM paths to understand the effects of context-related information integration.

This study did not examine emergent strengths from SEM and context-sensitive DF integration. Potential synergistic strengths, which might manifest uniquely in the confluence of SEM and context-sensitive DF, were beyond the scope of the current research.

6.2.1.2 NREN Model Instantiations for TFSEMDF Strengths Exploration

Section 6.3.1 investigated the SEM's ability to model complex hierarchical interrelationships (Danks et al., 2020), PLS-SEM's proficiency in modelling with small sample sizes (Chin & Newsted, 1999), and SEM-based DF's capacity to incorporate context-related information (Staphorst et al., 2013, 2014, 2016a, 2016b). The analysis used a baseline NREN model instantiation identical to the cross-sectional NREN model instantiation constructed in Section 5.2.2.1 and depicted in Figure 8.

The exploration assessed whether the TFSEMDF framework displayed SEM's capabilities in modelling complex interrelationships and accommodating small sample sizes, drawing directly from the PLS regression results presented in Section 5.3.2.2 for the cross-sectional NREN model instantiation. Additionally, the investigation into SEM-based DF's ability to utilise context-related information used the structure of the cross-sectional NREN model instantiation from Figure 8 while varying the inclusion level of reflective indicators for the context-related exogenous construct *Government Influence over the NREN* (ξ_1). The study compared the path coefficients and their significance results obtained by omitting combinations of the reflective indicators *NREN Governance Mode* (X_1), *Level of Government Funding* (X_2) and *Range of Institutions the NREN is Mandated to Connect* (X_3) against the results for the baseline NREN model instantiation.

6.2.1.3 TFSEMDF Strengths Exploration Research Propositions

Section 5.2.2.2 introduced five research propositions specifically crafted to define the hypothesised relationships (i.e. the SEM structural model paths) among the diverse constructs within the cross-sectional NREN model instantiation, as illustrated in Figure 8. The design of each proposition aimed to capture the nuanced interrelations within the model, providing a detailed theoretical framework for the subsequent PLS regression analyses. The study applied these propositions unchanged in the PLS regression analysis outlined in Section 6.3.1, which used the cross-sectional NREN model instantiation as the baseline for exploring TFSEMDF strengths, including the SEM-based DF's ability to integrate context-related information, SEM's capacity to model complex hierarchical interrelationships, and PLS-SEM's effectiveness in working with small sample sizes.

6.2.1.4 TFSEMDF Strengths Exploration Indicator Data

The PLS regression analysis of the baseline NREN model instantiation, conducted to evaluate the TFSEMDF framework strengths routed in SEM's ability to represent complex hierarchical interrelationships, PLS-SEM's ability to effectively model with small sample sizes and SEM-based DF's ability to incorporate context-related information, utilised TERENA's 2011 NREN Compendium data as the source for the technology and context-related indicators included in Figure 8. Adopting the indicator data sourcing strategy from Section 5.3.2 for the PLS regression analysis of the cross-sectional NREN model instantiation, Section 6.3.1 utilised the data composition approach outlined in Table 2 to extract technology and context-related indicator data from TERENA (2011) for the baseline NREN model instantiation. Specifically, the analysis probing SEM-based DF's capability to assimilate context-related information within the TFSEMDF framework applied various combinations of the data compiled for the reflective indicators *NREN Governance Mode (X_1)*, *Level of Government Funding (X_2)*, and the *Range of Institutions the NREN is Mandated to Connect (X_3)*.

6.2.2 WEAKNESSES OF THE TFSEMDF FRAMEWORK

6.2.2.1 Methodology to Identify and Verify TFSEMDF Weaknesses

The study implemented a structured two-step approach to achieve Research Objective 3(b), centred on identifying and empirically analysing inherent weaknesses in the TFSEMDF framework. The first step entailed an extensive literature review, as outlined in Chapter 2, dedicated to uncovering the intrinsic inadequacies of SEM and context-sensitive DF. Based on the methodology described by Braun & Clarke (2006), this review incorporated thematic analysis, facilitating a systematic identification and interpretation of themes related to weaknesses inherent in SEM and context-sensitive DF (Ryan & Bernard, 2003).

Subsequently, the second step involved using the cross-sectional NREN model instantiation developed in Section 5.2.2.1 and thoroughly analysed in Section 5.3.2.1 as a baseline for empirical verification. A structurally disarranged NREN model instantiation was also constructed and analysed. In this instance, the dual approach aimed to empirically validate whether the weaknesses identified from the literature review were present and impactful within the

TFSEMDF framework. Weaknesses considered in this study, as detailed in Section 6.3.2, included the implications of SEM model misspecification and PLS-SEM's tendency to overfit when a model, tailored too closely to a specific dataset, captures random noise instead of underlying relationships, leading to poor predictive performance on new data. (Danks et al., 2020).

As explained in Section 6.2.1.1's methodology for investigating context-related information inclusion during SEM-based DF as a strength of the TFSEMDF framework, standard reliability and validity metrics like Coefficients of Determination and Predictive Validity are unsuitable for comparing model instantiations employing PLS regression (Evermann & Tate, 2010). Therefore, comparing the baseline NREN model instantiation with the structurally disarranged NREN model instantiation developed in Section 6.2.2.2 to assess the impact of SEM model misspecification as a weakness of the TFSEMDF framework, the study adopted the alternative approach proposed by Evermann and Tate (2010). The analysis thoroughly examines the presence or absence of SEM paths in both the baseline and structurally disarranged NREN model instantiations, providing crucial insights into how model misspecification affects the outcomes.

As was the case with the exploration into strengths of the TFSEMDF framework, this study did not analyse potential emergent weaknesses that may arise from the integration of SEM and context-sensitive DF. Such weaknesses, potentially unique to the interplay between SEM and context-sensitive DF, were not within the research scope defined for this investigation.

6.2.2.2 NREN Model Instantiation for TFSEMDF Weaknesses Exploration

To examine possible weaknesses in the TFSEMDF framework, attributable to the inherent limitations of its SEM and context-sensitive DF building blocks, both baseline and structurally disarranged NREN model instantiations were employed. For the baseline NREN model instantiation, a replica of the cross-sectional NREN model instantiation, as constructed in Section 5.2.2.1 and depicted in Figure 8, was employed. Conversely, the study used a structurally altered version of the cross-sectional NREN model instantiation as the structurally disarranged NREN model instantiation.

Moreover, to investigate the ramifications of poorly defined structures in TFSEMDF framework model instantiations, the study scrutinised the disarranged baseline NREN model instantiation depicted in Figure 9, derived by varying the cross-sectional NREN model instantiation in Figure 8. This variant maintained identical measurement components to the baseline NREN model instantiation, encompassing the same technology and context-related measurement indicators and constructs, ensuring measurement consistency.

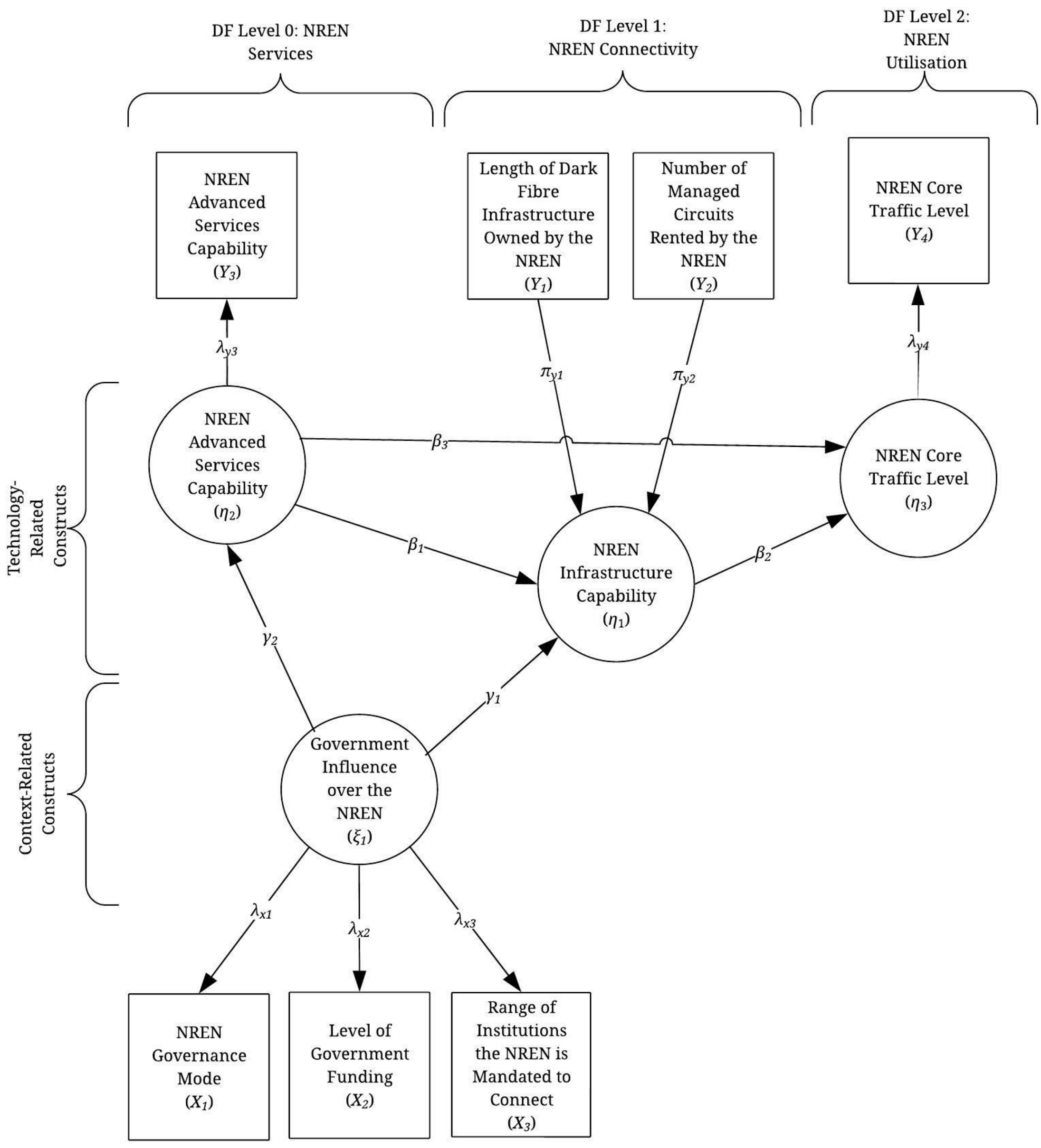


Figure 9: Disarranged NREN Model Instantiation of the TFSEMDF Framework

The structural portion in the disarranged NREN model instantiation of the TFSEMDF framework deliberately diverged from the theoretical foundation of the baseline NREN model instantiation. As a result, it contradicted both the OSI model (Zimmerman, 1980) and Greaves' (2009) NREN CMM, which advocate that infrastructure capability is foundational for advanced services capability. As depicted in Figure 9, this variant altered this established relationship by swapping the positions of DF Levels 0 and 1 from their arrangement in Figure 8, thereby creating a structural configuration that directly conflicts with the theoretical principles underpinning the baseline model.

6.2.2.3 TFSEMDF Weaknesses Exploration Research Propositions

The research propositions established for the baseline NREN model instantiation, as outlined in Section 5.2.2.2, broadly apply to the investigation focusing on the consequences of a poorly defined structural model. However, an exception exists for proposition H2, which necessitates alteration due to the directional change in the path between the *NREN Infrastructure Capability (η_1)* and *NREN Advanced Services Capability (η_2)* constructs in the variant model. Consequently, the comprehensive suite of research propositions adapted for the, as illustrated in Figure 9, is defined as follows:

- **Research Proposition H1:** The NREN's infrastructure capability positively correlates with government influence over the NREN. This hypothesis derives from the same conceptual principles that shaped H1 for the baseline NREN model instantiation.
- **Research Proposition H2:** The NREN's infrastructure capability is presumed to be positively related to its advanced services capability. This proposition, unique in its lack of a supporting conceptual basis, exemplifies the structural deficiencies highlighted in Figure 9.
- **Research Proposition H3:** A positive relationship exists between the NREN's advanced services capability and the level of government influence over the NREN. This hypothesis shares the same conceptual underpinnings as H3 for the baseline NREN model instantiation.
- **Research Proposition H4:** The level of core network traffic within the NREN is positively related to its infrastructure capability. This proposition draws on conceptual grounds like those of H4 in the baseline NREN model instantiation.
- **Research Proposition H5:** A positive association exists between the level of core network traffic in the NREN and its advanced services capability. This hypothesis relies on the conceptual foundations that informed H5 in the baseline NREN model instantiation.

6.2.2.4 TFSEMDF Weaknesses Exploration Indicator Data

The data composition scheme detailed in Table 2, which outlines the extraction and processing of TERENA's 2011 NREN Compendium data (TERENA, 2011), was utilised to produce technology and context-related indicator data for the PLS regression analysis of the baseline NREN model instantiation. Furthermore, the same scheme was employed to analyse the structurally disarranged NREN model instantiation, facilitating comparability and consistency in evaluating both model instantiations of the TFSEMDF framework.

6.3 RESULTS

The subsequent sections detail the results garnered during this research phase. Section 6.3.1 outlines the findings from evaluating the TFSEMDF framework's strengths, as undertaken in Research Objective 3(a). Conversely, Section 6.3.2 discusses the findings from assessing the framework's weaknesses, which correspond to Research Objective 3(b).

The study utilised the indicator data for the cross-sectional NREN model instantiation, as described in Section 5.2.2.3, in conducting the PLS regression analyses of the baseline and structurally disorganised NREN model instantiations to evaluate the strengths and weaknesses of the TFSEMDF framework. As such, the results presented in the following sections used the 2011 TERENA NREN Compendium (TERENA, 2011), with a sample consisting of 61 participating NRENs, but only 28 NRENs providing the complete set of data required to accomplish indicator data composition following the process defined in Table 2. As explained in Section 5.3.2, this was deemed sufficient per the "ten-times rule" sample size rule for PLS regression analysis in SEM (Chin & Newsted, 1999; Goodhue et al., 2006).

6.3.1 STRENGTHS OF THE TFSEMDF FRAMEWORK

6.3.1.1 TFSEMDF Strength: Context-Related Information Inclusion using SEM-based DF

As described in Section 6.2.1.2, exploring SEM-based DF's ability to include context-related information as a strength of the TFSEMDF framework, this study elected to examine the specific impact that varying levels of indicator data for the context-related construct *Government Influence over the NREN (ξ_1)* had on the overall structural validity of the model instantiation (Staphorst, 2010). This examination entailed comparing the path coefficients and their significance test results for the baseline NREN model instantiation, obtained through PLS regression analysis using varying combinations of indicator data extracted from TERENA (2011) for the reflective indicators *NREN Governance Mode* (X_1), *Level of Government Funding* (X_2) and *Range of Institutions the NREN is Mandated to Connect* (X_3).

Including the full suite of reflective indicators for the context-related construct *Government Influence over the NREN* (ξ_1) in the baseline NREN model instantiation produced the path coefficient results presented in Table 9 and their significance test results in Table 11. The path coefficient significance (Goodness-of-Fit) was tested via asymptotic *t*-statistics, resulting in associated *p*-values.

From Table 11's path coefficient significance test results, obtained using SmartPLS's bootstrapping function (Ringle et al., 2022) configured to generate 500 subsample sets from the 61 cases in the original sample, the only path that exhibited a *p*-value higher than $\alpha = 0.10$, the maximum acceptable significance level, was *NREN Infrastructure Capability* (η_1) $\rightarrow$ *NREN Advanced Services Capability* (η_2). Hence, this path was deemed insignificant in the baseline NREN model instantiation, including the full suite of reflective indicator data for the context-related construct *Government Influence over the NREN* (ξ_1).

Repeating this analysis for the baseline NREN model instantiation while limiting the level of reflective indicator data for the context-related construct *Government Influence over the NREN* (ξ_1) entailed including various combinations of the metrics X_1, X_2 and X_3 during the PLS regression analysis. This approach produced the path coefficient results in Table 13 and their significance test results in Table 14. Table 13 shows paths not supported due to negative or zero path coefficients in grey, while Table 14 shows paths judged as insignificant at $\alpha = 0.10$ in grey.

Table 13: TFSEMDF Strength: DF Context-Related Information Inclusion – Part A

Research Proposition: Baseline NREN Model Instantiation SEM Path	**Path Coefficient**	**Path Coefficients Given Context-Related Indicators Included**						
		None	X_1	X_2	X_3	X_1, X_2	X_1, X_3	X_2, X_3

H1: Government Influence over the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	γ_1	0	0.587	0.525	0.441	0.610	0.585	0.531
H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	β_1	0.523	0.313	0.193	0.170	0.144	0.045	0.070
H3: Government Influence over the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	γ_2	0	0.368	0.635	0.812	0.631	0.829	0.862
H4: NREN Infrastructure Capability (η_1) → NREN Core Traffic Level (η_3)	β_2	0.306	0.274	0.300	0.296	0.287	0.283	0.297
H5: NREN Advanced Services Capability (η_2) → NREN Core Traffic Level (η_3)	β_3	0.179	0.194	0.182	0.184	0.188	0.190	0.183

Table 14: TFSEMDF Strength: DF Context-Related Information Inclusion – Part B

Research Proposition: Baseline NREN Model Instantiation SEM Path	***p*-Values given Context-Related Indicators Included**						
	None	X_1	X_2	X_3	X_1, X_2	X_1, X_3	X_2, X_3
H1: Government Influence over the NREN (ξ_1) → NREN Infrastructure Capability (η_1)	1.000	0.001	0.001	0.013	0.001	0.001	0.001

H2: NREN Infrastructure Capability (η_1) → NREN Advanced Services Capability (η_2)	0.001	0.173	0.242	0.223	0.371	0.775	0.562
H3: Government Influence over the NREN (ξ_1) → NREN Advanced Services Capability (η_2)	1.000	0.284	0.001	0.001	0.005	0.001	0.001
H4: NREN Infrastructure Capability (η_1) → NREN Core Traffic Level (η_3)	0.225	0.375	0.337	0.355	0.255	0.443	0.216
H5: NREN Advanced Services Capability (η_2) → NREN Core Traffic Level (η_3)	0.462	0.520	0.547	0.561	0.449	0.602	0.430

6.3.1.2 TFSEMDF Strength: Modelling Complex Interrelationships using SEM

The primary strength of SEM within the TFSEMDF framework lies in its sophisticated ability to model intricate and hierarchical relationships among indicators and constructs. It extends beyond the capabilities of classic regression techniques by accommodating non-linear relationships, non-Gaussian distributions, and cyclical dependencies between model variables (Staphorst, 2010; Staphorst et al., 2016a; Vinzi et al., 2010). Techniques such as multiple regression, discriminant analysis, logistic regression, and analysis of variance are deemed first-generation due to their assumption of independence among numerous dependent variables.

First-generation techniques face constraints due to their inability to comprehensively capture complex interdependencies, particularly when addressing interactions among multiple output variables. Additionally, these methods often fail to effectively model the mediation effects that one construct may exert on another. Furthermore, these techniques rely on the assumption that all variables are directly observable, thereby restricting their application to variables that can be empirically represented (Staphorst, 2010; Staphorst et al., 2016a; Vinzi et al., 2010).

In contrast, SEM embraces observable and latent constructs, thus bridging the gap left by first-generation models and providing a more comprehensive framework for modelling complex systems (Staphorst, 2010; Staphorst et al., 2016a; Vinzi et al., 2010). As a result, SEM has become an indispensable technique for capturing the nuanced interplay of unobservable variables often encountered in advanced research contexts.

The PLS regression analysis of the baseline NREN model instantiation, as detailed in Section 5.3.21, along with the assessments of reliability and validity in Section 5.3.2.2, attest to the SEM's adeptness within the TFSEMDF framework for modelling sophisticated, multi-layered technology domains. This model instantiation effectively captured a comprehensive view of the NREN technology domain in three layers, each representing distinct yet interdependent facets: NREN connectivity, NREN services, and NREN utilisation.

Moreover, the baseline instantiation adeptly integrated a spectrum of constructs, encompassing technology-centric measures such as NREN core traffic level and context-related factors like governmental influence over the NREN. It also proficiently mapped the conjectured network of relationships among these constructs, affirming the framework's capacity to represent complex technology and context-related interrelations within NRENs.

6.3.1.3 TFSEMDF Strength: Modelling with Small Sample Sizes using PLS-SEM

A significant advantage of SEM, when implemented through PLS regression, is its ability to surmount challenges frequently encountered in survey-based business research that impede the effectiveness of traditional covariance-based regression methods (Staphorst, 2010). Such challenges include suboptimal response rates, incomplete item responses by participants, and the prevalence of highly interrelated survey items. Unlike classical covariance-based regression techniques, which tend to produce unreliable outcomes in the face of limited sample sizes and missing data, PLS regression maintains stability. It also mitigates the impact of multicollinearity on the standard errors of estimated regression coefficients, thereby preserving the inclusion of legitimate predictors within the regression model (Staphorst, 2010; Vinzi et al., 2010). Moreover, PLS regression is proficient in handling multiple correlated output variables, enhancing the analytical breadth of the SEM application (Staphorst, 2010; Vinzi et al., 2010).

The utilisation of PLS regression within SEM confers a distinct advantage due to the partial estimation approach of the PLS algorithm (referenced in Appendix A), which facilitates less restrictive sample size demands compared to traditional CB-SEM methodologies like LISREL

(Chin & Newsted, 1999). Herman Wold, the originator of the PLS regression technique, advocated for its suitability in preliminary model explorations. This approach aligns with the investigative nature of studies like the current one, which seeks to develop TFSEMDF model instantiations that accurately reflect the NREN technology domain rather than for the strict testing of hypotheses within established SEM frameworks (Wold, 1980).

Within the field of SEM, a prevalent view among scholars is that PLS regression necessitates a sample size that is at least tenfold the number of formative indicators linked to the most complex latent construct within the model, whether endogenous or exogenous (Chin & Newsted, 1999; Goodhue et al., 2006). The measure of relationship complexity, for this "10-times rule," is determined by the count of formative indicators contributing to the latent construct that possesses the most significant number of such indicators (Goodhue et al., 2006). This guideline offers a heuristic for researchers to estimate the minimum sample size needed for reliable analysis in PLS-SEM studies (Staphorst, 2010).

For the baseline NREN model instantiation, the "10-times-rule" was applied to determine the requisite sample size for PLS regression analysis. In this context, the construct *NREN Infrastructure Capability* (η_1), depicted in Figure 8, was identified as having the most complex formative structure with two indicators: *Length of Dark Fibre Infrastructure Owned by the NREN* (Y_1) and *Number of Managed Circuits Rented by the NREN* (Y_2). Consequently, based on this rule, the minimum required sample size for successful PLS regression analysis is 20. In the TERENA 2011 NREN Compendium data collection (TERENA, 2011), complete survey responses, as required to perform the data conditioning calculations in Table 2, were obtained from only 28 out of the 61 NRENs. However, this surpassed the minimum threshold established by the "10-times-rule." This sample size adequacy validated the collected data's suitability for the PLS regression analysis of the baseline NREN model instantiation.

6.3.2 Weaknesses of the TFSEMDF Framework

6.3.2.1 TFSEMDF Weakness: SEM Model Misspecification

Table 15 below summarises the path coefficients obtained for the structurally disarranged NREN model instantiation of Figure 9, while Table 16 details the path coefficient significance results.

Table 15: TFSEMDF Weakness: SEM Model Misspecification – Part A

Research Proposition: SEM Path for Structurally Disarranged NREN Model Instantiation	**Path Coefficient**
H1: *Government Influence over the NREN* (ξ_1) → *NREN Infrastructure Capability* (η_1)	$\gamma_1 = 0.557$
H2: *NREN Advanced Services Capability* (η_2) → *NREN Infrastructure Capability* (η_1)	$\beta_1 = 0.046$
H3: *Government Influence over the NREN* (ξ_1) → *NREN Advanced Services Capability* (η_2)	$\gamma_2 = 0.865$
H4: *NREN Infrastructure Capability* (η_1) → *NREN Core Traffic Level* (η_3)	$\beta_2 = 0.295$
H5: *NREN Advanced Services Capability* (η_2) → *NREN Core Traffic Level* (η_3)	$\beta_3 = 0.184$

Table 15 indicates that there are no unsupported paths due to path coefficients that are negative or zero. In contrast, Table 16 shows paths judged as insignificant at $\alpha = 0.10$ in grey. Table 16's results were obtained using SmartPLS's bootstrapping function (Ringle et al., 2022), configured to generate 500 subsample sets from the 61 cases in the original sample.

Table 16: TFSEMDF Weakness: SEM Model Misspecification – Part B

Research Proposition: SEM Path for Structurally Disarranged NREN Model Instantiation	***p*-Value**
H1: *Government Influence over the NREN* (ξ_1) → *NREN Infrastructure Capability* (η_1)	0.818
H2: *NREN Advanced Services Capability* (η_2) → *NREN Infrastructure Capability* (η_1)	0.985
H3: *Government Influence over the NREN* (ξ_1) → *NREN Advanced Services Capability* (η_2)	0.001
H4: *NREN Infrastructure Capability* (η_1) → *NREN Core Traffic Level* (η_3)	0.034
H5: *NREN Advanced Services Capability* (η_1) → *NREN Core Traffic Level* (η_3)	0.130

6.3.2.2 TFSEMDF Weakness: PLS-SEM Overfitting

Overfitting in PLS-SEM is a significant issue, particularly in complex models with numerous predictors. This problem arises when a model is overly attuned to the specific data it was trained on, including its outliers and anomalies, rather than capturing the underlying relationships. Such overfitted models exhibit high accuracy on their training datasets but perform poorly on new, unseen data due to their lack of generalisability (Danks et al., 2020).

In the context of TFSEMDF model instantiations for the NREN technology domain, consider a PLS-SEM model developed to predict network congestion based on a myriad of technology indicators such as user behaviour patterns, data usage, network signal strength, as well as context-related indicators like weather conditions. If this TFSEMDF model is overly complex, including excessive minute and interrelated predictors, it may become too tailored to the specific dataset used for its development. As a result, while the model might accurately predict congestion scenarios within the training dataset, its ability to forecast the TF output metric for network congestion under different conditions or in other network environments may be substantially compromised.

This example emphasises the importance of a balanced approach in TFSEMDF model instantiation development when using PLS regression analysis, particularly in fields such as telecommunications, where the dynamics are complex and evolving. Ensuring the model captures relevant patterns with sufficient detail without binding it so intricately to the training data that its broader applicability suffers is essential. To mitigate the risks of overfitting and enhance the model's utility in real-world applications, rigorous validation techniques, including cross-validation and testing the model on different datasets, are vital (Danks et al., 2020).

6.4 DISCUSSION

Subsequent sections offer a critical discussion of the results obtained in this study phase. Section 6.4.1 discusses the results of assessing the TFSEMDF framework's strengths pertinent to Research Objective 3(a). Section 6.4.2 discusses the results of evaluating the framework's weaknesses, addressing Research Objective 3(b).

6.4.1 Strengths of the TFSEMDF Framework

The results presented in Section 6.3.1.1 for the PLS regression analyses of the baseline NREN model instantiation, using various levels of context-related information inclusion, clearly demonstrate the significant contribution of this information to developing a valid and reliable TFSEMDF model structure. An examination of the results in Table 13 and Table 14 reveals that the omission of any combination of reflective indicators from the *Government Influence over the NREN* (ξ_1) construct, specifically *NREN Governance Mode* (X_1), *Level of Government Funding* (X_2), and *Range of Institutions the NREN is Mandated to Connect* (X_3), resulted in invalidating at least three of the research propositions in the model instantiation. Path coefficients diminished to zero for these instances, or their significance results exceeded $\alpha = 0.10$.

As indicated by the results, including a complete set of context-related measurement indicators in the baseline NREN model instantiation substantially enhanced the model's overall structural validity. This inference arises from an analysis where only a single path faced invalidation, suggesting that incorporating context-related indicators into the TFSEMDF framework's forecasting calculations for the NREN technology domain enhances the model's structural soundness (Staphorst, 2010; Vinzi et al., 2010). Consequently, this inclusion likely improves the model's capacity to accurately estimate the TF output metric, *NREN Core Traffic Level* (Y_4), which is pivotal in measuring the utilisation of an NREN.

During this exploration, a notable observation emerged: Research Proposition H2, which posits a relationship between *NREN Infrastructure Capability* (η_1) and *NREN Advanced Services Capability* (η_2), survived in scenarios excluding context-related information. Contrary to expectations, this was the only proposition dismissed in the baseline model instantiation when including all context-related information. The phenomenon of technology leapfrogging within the global NREN community (Mbarika et al., 2000) offers a plausible explanation for this outcome. However, caution is advised in interpreting this result in isolation, as excluding all context-related information led to rejecting all other research propositions. Thus, the singular validity of H^2 does not imply an SEM model's efficacy in estimating the TF output metric *NREN Core Traffic Level* (Y_4).

In Sections 6.3.1.2 and 6.3.1.3, the study scrutinises two key strengths of the TFSEMDF framework: SEM's capability to model complex hierarchical systems and PLS-SEM's adeptness in handling regression analysis with small sample sizes. The results from these sections robustly indicate the framework's proficiency in developing valid and reliable models for intricate technology domains, such as NRENs. This process encompasses the integration of both technology and context-related constructs and indicators. Notably, the framework demonstrated its effectiveness even with limited datasets, as evidenced by applying TERENA (2011) indicator data for this analysis.

6.4.2 Weaknesses of the TFSEMDF Framework

Section 6.3.2.1 in the study addresses the impact of inadequate structural design in the TFSEMDF framework, pinpointing it as a significant weakness. As detailed in Table 15 and Table 16, the analyses expose the effects of a poorly formulated structure for a TFSEMDF model instantiation. Specifically, the study examined a scenario that altered the path between *NREN Infrastructure Capability (η_1)* and *NREN Advanced Services Capability (η_2)*, deviating from its original theoretical foundation. The result was the rejection of three essential paths, including a crucial path linked to the *NREN Core Traffic Level (η_3)* construct, which is vital for the TF output metric. Consequently, this structurally disarranged model instantiation was ineffective in accurately forecasting the *NREN Core Traffic Level (Y_4)*, emphasising the critical importance of a theoretically sound structural design in model development within the TFSEMDF framework.

The distinction between the cross-sectional NREN model instantiation (Section 5.2.1.1) and the autoregressive NREN model instantiation (Section 5.2.2.1) exemplifies the importance of proper structure in TFSEMDF framework model instantiations. Based on Staphorst et al. (2016a) and used as the baseline NREN model instantiation in this chapter, the cross-sectional model contrasts with the autoregressive model, which needs more hypothesis integration from peer-reviewed literature as in Staphorst et al. (2016b). The latter's limited success, with only three valid paths out of eight, highlights the significance of grounding model structures in well-established theoretical hypotheses.

Section 6.3.2.2 addresses a fundamental weakness in the TFSEMDF framework related to the tendency of PLS-SEM to overfit, especially in complex models with an abundance of predictors. This issue, where the model becomes overly sensitive to the nuances of the training data, including outliers and anomalies, can detract from its ability to discern genuine underlying relationships. Future research directions for the TFSEMDF framework will include employing SEM post-processing techniques, using datasets different from those originally used to create model instantiations to calculate output TF metrics. This examination will focus on strategies to mitigate the occurrence of PLS-SEM overfitting within the TFSEMDF framework.

6.5 CONCLUDING REMARKS

This chapter delved into the study's third phase, focusing on assessing the inherent strengths and weaknesses of the TFSEMDF framework within the NREN technology domain. It commenced with an introduction that set the context for this phase. It built upon the insights derived from the earlier analyses of the improved cross-sectional NREN model instantiation and a structurally disarranged NREN model instantiation. The chapter's objective was to critically assess the efficacy and limitations of the TFSEMDF framework, leveraging the empirical findings from Chapter 5's analysis to inform a nuanced understanding of its performance as a TF tool in complex technology domains, such as the NREN ecosystem.

This chapter's section on the research methodology for the third phase of the study consists of two distinct yet interconnected parts. The first part explored the TFSEMDF framework's intrinsic strengths, while the second examined its inherent weaknesses. This approach utilised the initial simplex autoregressive and the improved cross-sectional NREN model instantiations, as analysed in Chapter 6. This dual-faceted methodology facilitated a holistic evaluation, allowing for a balanced assessment that underscored the framework's capabilities and identified areas for potential improvement.

The results section of this chapter presents findings following the two research objectives set for this phase: understanding the strengths and weaknesses of the TFSEMDF framework. The first part of the section delved into the framework's strengths, providing a thorough analysis that validated its efficacy in the NREN technology domain. The second part transitions to critically exploring the framework's weaknesses, shedding light on areas where it may face constraints or challenges.

The chapter's discussion section critically reviewed the findings from the investigation into the TFSEMDF framework's inherent strengths and weaknesses. It thoroughly explored the implications of these aspects, carefully evaluating their collective impact on the framework's overall effectiveness and practical use. This analysis aimed to place the findings within the larger context of TF in the NREN domain, providing insights on utilising the framework’s strengths and mitigating its weaknesses to improve performance.

Finally, Section 7.2.3 articulates conclusions about this research phase. This section provides an extensive discussion of the conclusions drawn from the chapter's results and discussions, delivering an objective assessment of the TFSEMDF framework's strengths and weaknesses in the context of the NREN technology domain.

CHAPTER 7 – CONCLUSIONS, REFLECTIONS AND FUTURE DIRECTIONS

7.1 INTRODUCTION

The chapter begins by articulating the conclusions from the study's three phases. Initially, it examines the first phase, detailing conclusions from the methodologies employed and findings gathered in developing the TFSEMDF framework. This stage encompassed the creation of generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF, alongside their integration into the TFSEMDF framework. The narrative then progresses to the second phase, presenting conclusions based on applying the TFSEMDF framework, specifically the construction and PLS regression analysis of the autoregressive and cross-sectional NREN model instantiations. Finally, the chapter concludes the third phase by evaluating the methodologies implemented and results obtained, particularly in assessing the TFSEMDF framework's strengths and shortcomings.

Next, this chapter reflects on the study's achievement of the stated research aims and objectives. It covers the development, application, and evaluation of the TFSEMDF framework in the NREN technology domain, highlighting significant accomplishments. Furthermore, it details the thesis sections that outline methodologies and results, providing a clear and comprehensive summary of the achievement of the research goals.

This chapter also reflects ongoing research efforts that enhance and extend the TFSEMDF framework and other independent studies focusing on TF within the NRENs. These studies employ diverse approaches, including participatory foresight and systems thinking, to address complex challenges in forecasting technology trends and service adoption in NRENs. Additionally, the growing trend of exploring technological forecasting from the perspective of complex systems receives attention. This trend highlights an increasing shift towards more integrative and holistic approaches in the field, underlining the importance of considering complex interdependencies within TF.

Lastly, the study focuses on the exploration of future research avenues. Future research will focus on adapting the TFSEMDF framework for use in technology domains beyond NRENs and efforts to strengthen and address identified weaknesses. Additionally, it considers future research could focus on addressing the shortcomings in the research methodologies and secondary data sources employed throughout this study, paving the way for more robust and comprehensive research approaches in the field of TF.

7.2 RESEARCH PHASE CONCLUSIONS

The subsequent sections detail the conclusions from each phase of the research. Section 7.2.1 addresses the outcomes of the first phase, which focused on developing the TFSEMDF framework. This phase involved creating and integrating generic frameworks for SEM, DF, and technology indicator relational mapping for TF, thereby building upon the preliminary conclusions from Staphorst et al. (2013). Section 7.2.2 examines the findings of the second phase, which involved applying the TFSEMDF framework to construct and analyse the autoregressive and cross-sectional NREN model instantiations. This section enhances the initial findings presented by Staphorst et al. (2014, 2016a). Section 7.2.3 explores the conclusions from the third phase, investigating the inherent strengths and weaknesses of the TFSEMDF framework within the NREN technology domain and extending the foundational work discussed in Staphorst et al. (2016b).

7.2.1 RESEARCH PHASE ONE: TFSEMDF FRAMEWORK DEVELOPMENT

The first phase of the study, focusing on developing the TFSEMDF framework, concludes with significant achievements in addressing Research Objectives 1(a) and 1(b). This foundational phase, captured in Chapter 4, led to the creation of generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF and their subsequent integration into the cohesive TFSEMDF framework. The results from this phase represent a synthesis of theoretical models into a practical and functional framework, highlighting the study's innovative approach to seamlessly merging these diverse methodologies.

The successful completion of Research Objective 1(a) entailed the development of generic frameworks for SEM, context-sensitive DF, and technology indicator relational mapping for TF. Achieving this required conducting a wide-ranging literature review and employing a systematic approach that combined comparative analysis (Glasser, 1965; Smith, 2015) with grounded theory methodology (Charmaz, 2006). This strategic methodology was essential in distilling each field's complex features and thematic patterns. As a result, the study has abstracted frameworks encapsulating the core principles of SEM, DF, and technology indicator relational mapping for TF, showing the potential for interplay and synergy among these domains. These frameworks, characterised by their structural overlap and scalability in layering complexity, are tailored to meet diverse user requirements and application contexts in TF.

The successful completion of Research Objective 1(b) is marked by the strategic application of the framework unification methodology, effectively integrating the generic frameworks for SEM, DF, and technology indicator relational mapping for TF into the cohesive TFSEMDF framework. The process incorporated comparative analysis (Glasser, 1965; Smith, 2015) and enhancement through cross-disciplinary integration (Johnson & Onwuegbuzie, 2004), seamlessly weaving unique characteristics of each domain into a unified model. The result of this concerted effort is a synthesis of diverse methodologies, presenting a framework that is not only complex in its structural design but also robust in its functional capabilities. This framework encapsulates the theoretical depth and practical versatility of SEM, the context-sensitive data processing sophistication of DF, and the forward-looking insights of TF, making it a groundbreaking tool in the field of TF.

The TFSEMDF framework is particularly notable for its ability to accurately represent complex hierarchical relationships between technology indicators and TF output metrics, even in scenarios of limited data availability. It accommodates various indicator data types and can effectively navigate complex construct interplays. The framework's versatility extends to seamlessly handling non-linear and non-Gaussian factors and addressing cyclical or mediating dependencies among latent and directly observable variables (Vinzi et al., 2010). Its scalable architecture allows it to adjust to diverse user requirements and contexts, seamlessly integrating context-related information into the forecasting process. Thus, the TFSEMDF framework presents a revolutionary approach to TF, uniquely tailored to complex technology domains with intricate structural and contextual dynamics.

7.2.2 Research Phase Two: TFSEMDF Framework Application

The second phase of the study, encompassing Research Objectives 2(a) and 2(b), was devoted to the realisations of the TFSEMDF framework by creating and analysing two distinct NREN model instantiations. Chapter 5 details this study phase's methodologies, results, and discussions. Objective 2(a) involved developing an autoregressive NREN model instantiation for longitudinal TF, applying PLS regression analysis to forecast technological trends over time. In contrast, Objective 2(b) focused on a cross-sectional NREN model instantiation, utilising PLS regression for transversal TF to capture a snapshot of technological status across various NRENs. These model instantiations were rigorously analysed using the 2011 TERENA NREN Compendium data. This phase was crucial in determining key model parameters, such as indicator loadings and path coefficients, and in validating the reliability and structural integrity of the model instantiations. It extended the foundational research of Staphorst et al. (2013, 2014, 2016a), demonstrating the TFSEMDF framework's capability to adapt to different modelling approaches within the NREN context.

The construction of the autoregressive NREN model instantiation in this study was grounded in deductive reasoning, drawing insights from the author's experiential learning while serving as the Director of SANReN and from the TERENA 2011 NREN Compendium. This model instantiation was a three-layer model instantiation of the TFSEMD framework, incorporating eight defined Research Propositions. The design of these propositions stemmed from hypothesised interrelationships among various technology and context-related constructs. By employing path coefficient and significance testing, the study rigorously tested these hypothesised relationships within the autoregressive NREN model, providing a methodological approach for validating the interconnectivity of the constructs defined.

The PLS regression results for the autoregressive NREN model instantiation offered comprehensive insights. Most technology-related indicators, barring *Access to IPv6 Addresses* (Y_4), effectively measured their designated constructs. Context-related indicators showed a similar pattern of adequacy in representing their constructs. The study found evidence supporting several hypothesised relationships within technology constructs (e.g., between *NREN Infrastructure Capability* (η_1) and *NREN Advanced Services Capability* (η_2)), as well as between technology and context constructs (e.g., between *Scope of the NREN Mandate* (ξ_2) and *Current NREN Reach* (η_3)), underpinned by secondary data from the TERENA NREN Compendium series (GÉANT, 2022; TERENA, 2011, 2012). However, the data did not corroborate certain postulated relationships, such as between the *Scope of the NREN Mandate* ($\xi 2$) and *Forecasted NREN Reach* (η_4) constructs. Furthermore, the reliability and validity evaluation of the NREN model instantiation revealed that while the measurement portion effectively contributed to the forecast of future NREN reach, the structural portion exhibited limitations.

In addressing the construction of the cross-sectional NREN model instantiation, the study applied learnings from the methodological approach used in the autoregressive model instantiation, aiming to refine its definition of research propositions to avoid excessively rejecting SEM paths. Drawing on a combination of action research from SANReN, TERENA's yearly NREN Compendium data (GÉANT, 2023; TERENA, 2011, 2012), and literature on hypothesised interrelationships between constructs, defining five research propositions for the cross-sectional model instantiation (which also consisted of three DF layers). These propositions aimed to represent the theorised interconnections between diverse technology and context-related constructs. Moreover, the cross-sectional NREN model instantiation, as conceived in this study, theorised that an NREN's infrastructure capability (defined at DF level 0) and its advanced services capability (defined at DF level 1) both have a positive relationship with the government's influence, categorised as a contextual factor.

Additionally, the cross-sectional NREN model instantiation hypothesised a positive correlation between the NREN's infrastructure and advanced services capabilities. These capabilities, in conjunction, were believed to positively influence the NREN's core network traffic (an indication of its utilisation), defined as the TF output metric for this specific model instantiation. The study used path coefficient analysis and significance testing to test these propositions within the cross-sectional NREN model instantiation.

Applying the 2011 TERENA NREN Compendium data for PLS regression in the cross-sectional NREN model instantiation provided significant insights. This analysis confirmed that all selected technology indicators effectively measured their corresponding constructs. However, the data did not support one hypothesised relationship, the positive correlation between an NREN's infrastructure and its advanced services capability. This anomaly might be attributed to the trend of technology leapfrogging prevalent in the global NREN community, where developing NRENs, often with support from more advanced counterparts, are rapidly adopting advanced services despite relatively nascent infrastructure capabilities (GÉANT, 2023; Mbarika et al., 2000; Melhem et al., 2021). The final analysis of the cross-sectional NREN model instantiation, using 2011 TERENA NREN Compendium data, demonstrated that its measurement and structural components significantly contributed to accurately forecasting the technology usage metric for NRENs.

The study's second phase underscores three vital insights about the TFSEMDF framework: First, its proficiency in modelling complex, multi-layered technology domains like NRENs, where contextual factors significantly influence technological evolution. Second, the framework demonstrates versatility in generating longitudinal and cross-sectional TF metrics, catering to varied analytical needs within the technology domain. Third, constructing TFSEMDF model instantiations necessitates a harmonious blend of specialist knowledge (sourced from action research), established theoretical underpinnings, and robustly collected technology and context-related data. This balanced approach is crucial for the framework's efficacy in delivering insightful and accurate TF outcomes.

7.2.3 Research Phase Three: TFSEMDF Framework Assessment

The third phase of this study, detailed in Chapter 6, delved into the TFSEMDF framework's inherent strengths and weaknesses derived from its constituent elements, PLS-SEM and context-sensitive DF. Expanding upon Staphorst et al.'s earlier work (2013, 2014, 2016a), this phase examined strengths like the SEM-based DF's integration of context-related information, SEM's modelling of complex hierarchical interrelationships, and PLS-SEM's effectiveness with small sample sizes in its efforts to address Research Objective 3(a). It also explored weaknesses, such as the ramifications of SEM model misspecification and the tendency of PLS-SEM to overfit in cases where models become too narrowly tailored to specific datasets, as part of Research Objective 3(b). This exploration used the cross-sectional NREN model as a baseline and a structurally disarranged derivative, utilising the 2011 TERENA NREN Compendium for data. Following Evermann and Tate's (2010) SEM model comparison method, the focus was on path coefficients and significance results to explain the impact of these strengths and weaknesses.

Integrating context-related information into the TFSEMDF framework significantly enhanced its structural validity, especially in forecasting the core traffic volumes of NRENs within the baseline model instantiation. This enhancement demonstrates the framework's profound strength in capturing the nuanced interplay of various technological and contextual factors within a technology domain. However, a potential vulnerability arises if the model specification does not consider all pertinent context-related information. Such an oversight can result in superficially structured models needing more depth to comprehensively encapsulate the full array of dynamics in the specific technology domain. Such underdeveloped models risk omitting critical aspects essential for a thorough and accurate representation of the technology domain, potentially limiting the effectiveness and applicability of the framework in real-world scenarios.

Within the TFSEMDF framework, using SEM to model the intricate and hierarchical relationships within NRENs represents a significant strength. SEM's ability to handle complex, nonlinear relationships and its proficiency in dealing with non-Gaussian distributions and cyclical dependencies far exceed the capabilities of traditional regression techniques. This sophistication is particularly evident in its integration of observable and latent constructs, providing a more holistic understanding of the NREN environment and capturing crucial aspects like connectivity, services, and utilisation. The application of SEM in this context, reinforced by the PLS regression analysis, illustrates its capacity to map out the multi-dimensional nature of NRENs and highlights its essential role in dissecting the intricate interplay of technology and context-related metrics.

Using PLS-SEM within the TFSEMDF framework for modelling with small sample sizes represents a substantial methodological strength. This approach successfully addresses common challenges in survey-based studies, such as suboptimal response rates and the complexity of interrelated survey items, which often impede traditional covariance-based methods. The stability of PLS regression in scenarios of limited sample sizes and its ability to mitigate the effects of multicollinearity enhances the reliability of the model outcomes. This aspect is crucial in the NREN context, where comprehensive data gathering can be challenging, making PLS-SEM a valuable tool for insightful analysis with constrained datasets. The practical adherence to the "ten-times rule" for sample size determination, as demonstrated in the baseline NREN model instantiation, not only exemplifies the applicability of this rule but also highlights the robustness of PLS-SEM in producing reliable insights from smaller, more focused datasets. This capability is especially pertinent in advancing NREN technology research, where the precision and accuracy of model predictions are paramount yet often limited by dataset size.

The study illuminated that model instantiations of the TFSEMDF framework, if not solidly based on theoretical underpinnings, face significant path validity issues. This misalignment often leads to the invalidation of paths, marked by zero or negative coefficients and low significance scores, and this effect can extend to even those theoretically grounded paths. On the other hand, the framework's core strength lies in its adeptness at modelling complex hierarchical structures between technology indicators and TF output metrics. It capably handles non-linear and non-Gaussian factors and cyclical dependencies among variables, showcasing its robustness in intricate technology forecasting contexts (Staphorst, 2010; Vinzi et al., 2010).

The issue of overfitting in PLS-SEM poses a significant challenge, particularly for models with many predictors. This challenge is crucial in telecommunications and similar fields, where the dynamics are intricate and constantly evolving. Overfitting manifests when a model, finely tuned to specific training data, fails to generalise and perform accurately on new, unseen datasets. This issue becomes especially problematic in complex model instantiations that predict phenomena like network congestion, which might rely on various technology and context-related indicators. The tendency of such models to fit closely to specific datasets, including outliers and anomalies, compromises their predictive validity in different or evolving network environments. A balanced approach in developing the TFSEMDF model instantiation is essential to counteract this. While the model should be detailed enough to capture the relevant intricacies within the NREN technology domain, it should avoid being overly specific to the point where its broader applicability is compromised. Incorporating rigorous validation techniques, including cross-validation and testing across diverse datasets, is fundamental. These practices are critical for mitigating overfitting risks and enhancing the model's practical utility and adaptability.

7.3 REFLECTIONS ON RESEARCH AIMS AND OBJECTIVES

As outlined in Section 1.3, this study aimed to achieve three primary research objectives within the NREN technology domain: the development, application, and assessment of the TFSEMDF framework. Following the structured roadmap depicted in Figure 2, the research unfolded across three distinct phases, each targeting specific research objectives detailed in Chapter 3. To briefly present this journey, Table 17 provides an overview of how the study methodically addressed each research aim and its associated objectives. The table includes references to the applicable methodology and results sections and a short description of the achievements for each research objective.

Table 17: Reflections on Research Aims and Objectives

Phase: Research Aim	**Research Objective**	**Method and Results Sections**	**Achievements**
Phase 1: TFSEMDF Development	1(a) Generic SEM, DF and technology indicator relational mapping for TF framework abstraction	4.2.1 and 4.3.1	Developed generic frameworks for SEM, context-sensitive DF and technology indicator relational mapping for TF using inductive reasoning, with thematic analysis and grounded theory
	1(b) TFSEMDF framework construction	4.2.2 and 4.3.2	Developed TFSEMDF framework by combining generic frameworks for SEM, context-sensitive DF and technology indicator relational mapping for TF using inductive reasoning, with comparative analysis and cross-disciplinary integration
Phase 2: TFSEMDF Application	2(a) Autoregressive NREN model instantiation development	5.2.1 and 5.3.1	Constructed autoregressive NREN model instantiation using action research and TERENA (2011, 2012), then performed PLS regression analysis using TERENA (2011, 2012) data to determine model parameters, as well as their reliability and validity

	2(b) Cross-sectional NREN model instantiation development	5.2.2 and 5.3.2	Constructed cross-sectional NREN model instantiation using action research, TERENA (2011) and published hypotheses, then performed PLS regression analysis using TERENA (2011) data to determine model parameters, as well as their reliability and validity
Phase 3: TFSEMDF Assessment	3(a) Explore TFSEMDF strengths	6.2.1 and 6.3.1	Used cross-sectional NREN model instantiation as a baseline to perform PLS regression using TERENA (2011) data to explore TFSEMDF strengths, including context-related information inclusion, complex interrelationship modelling and accommodating small sample sizes
	3(b) Explore TFSEMDF weaknesses	6.2.2 and 6.3.2	Used baseline NREN model instantiation and structurally disarranged derivative thereof to perform PLS regression using TERENA (2011) data to explore TFSEMDF weaknesses, including model misspecification and PLS-SEM overfitting

7.4 Reflections on Related Research

This study amalgamated SEM, context-sensitive DF, and TF indicator relational mapping, creating the innovative TFSEMDF framework. This framework has demonstrated its applicability and robustness within the NREN technology domain and holds immense potential in shaping the future of TF. The author's ongoing efforts to further develop this framework, as outlined in Section 7.5, aim to enhance its application methodologies, extend its utility to other complex technology domains, and address its identified weaknesses. It is worth noting that while this research is progressing, other scholars are also actively involved in enhancing TFSEMDF, and numerous others are conducting related research into NRENs, thereby advancing the broader field of TF.

In their comparative analysis, Dash and Paul (2021) explore CB-SEM and PLS-SEM, including its variant, Consistent PLS. Focusing on applying these methodologies within social sciences and TF, the researchers explicitly cite the work by Staphorst et al. (2016a) and reference the development of the TFSEMDF framework, illustrating the foundational influence of Staphorst et al. (2013, 2014, 2016a, 2016b) on their study, particularly in how to apply SEM methodologies to complex forecasting scenarios. Dash and Paul (2021) use empirical data from an international sample to evaluate the strengths and weaknesses of CB-SEM and PLS-SEM in accurately predicting and explaining structural relationships. While PLS-SEM tends to yield higher item loadings and better construct reliability and validity, CB-SEM offers superior model fit indices. This distinction underscores PLS-SEM's utility in environments where adaptive modelling and theoretical exploration are prioritised over conventional theoretical confirmation, making it highly applicable in the TFSEMDF framework.

The research by Yaver et al. (2016) is a significant contribution to the field, as it focuses on identifying future trends in advanced technology services within NRENs through a foresight methodology. Their approach, which involved consultations with international experts, primarily directors from various NRENs, yielded key findings with profound implications. The study suggests a growing trend towards providing dark fibre links to users, which enhances flexibility and agility in delivering new and improved services. It also highlights the potential for NRENs to collaborate with mobile service providers to offer academically innovative solutions. Another crucial outcome is the importance of extending connectivity to non-university institutions like hospitals, which could facilitate remote healthcare services. These insights are invaluable for any NREN seeking to expand or enhance its service offerings and could significantly influence global strategic planning and technological upgrades in educational and research networks. Yaver et al.'s (2016) findings align with this study's cross-sectional NREN model instantiations' proposed relationship between an NREN's infrastructure capability and the utilisation of the NREN.

Pillay et al. (2021) developed a comprehensive systems-thinking model to advance the adoption of value-added services within NRENs. This model, articulated through a rigorously designed causal diagram, systematically identifies and visualises the intricate interconnections and dependencies that influence service adoption decisions within NRENs. The approach adopted by Pillay et al. (2021) involved the application of design science research methodologies, which facilitated a robust validation process through iterative refinements based on feedback from a panel of international NREN experts. This validation process strengthened the model's reliability. It ensured its applicability across diverse NREN contexts, providing stakeholders with a powerful tool for strategic decision-making and effective communication regarding NREN service strategies. Furthermore, the model encapsulates a holistic view of the NREN service ecosystem, allowing for a deeper understanding of the dynamics and potential levers for enhancing service management and delivery. The model developed by Pillay et al. (2021) supports the hypothesised relationship in this study's cross-sectional NREN model instantiation between an NREN's advanced services capability and the utilisation of the NREN.

Exploring technological forecasting from the perspective of complex systems theory, as outlined in the detailed analysis by Feng et al. (2022), marks a crucial trend in the TF field. This research highlights the importance of employing robust modelling techniques and dynamic adaptations to enhance the precision and relevance of forecasting methodologies. The study advocates integrating diverse data sources, thereby enriching the forecasting models to accurately capture the intricate interdependencies within technological ecosystems. Feng et al. (2022) demonstrate how complex systems theory can profoundly impact TF by emphasising nonlinear interactions and emergent behaviours, offering insights into more resilient and adaptive forecasting models. These insights are crucial for developing models that are reactive to current technological trends and proactive in anticipating future developments. The relevance of Feng et al.'s (2022) approach to the TFSEMDF framework developed in this thesis is significant. It underscores the necessity of modelling complex hierarchical relationships and achieving comprehensive data integration through context-sensitive DF, which is core to the framework's ability to effectively navigate and forecast the intricate dynamics of technology development within context-sensitive technology domains, such as NRENs.

7.5 Future Research Directions

This study paves the way for future research endeavours involving the TFSEMDF framework. These prospective studies encompass refining application methodologies for the TFSEMDF framework in the NREN technology domain and expanding its use to other complex technology domains where contextual factors are significant. Further research could address the TFSEMDF framework's identified weaknesses, capitalising on its strengths, exploring emergent features, and conducting comparative analyses with other TF methodologies. The following list outlines the specifics of these future research directions, providing a roadmap for ongoing scholarly investigation in this field.

- The TERENA NREN Compendiums, limited in respondent scope from the global NREN community, restrict the sample size for PLS regression in NREN model instantiations. Additionally, the data collection and processing methodologies employed by TERENA are not transparent. While PLS-SEM can handle smaller datasets, broader responses would enhance model accuracy. Future research will thus involve a qualitative study to identify improved constructs and indicators, directly extracting quantitative data on technology and context-related indicators from the global NREN community.
- While this study successfully determined the TFSEMDF's usefulness in longitudinal and transversal TF for the NREN technology domain, it did not compare its capabilities against other popular TF techniques. Hence, future research could include a comparative study between TFSEMDF and other methods, most notably TFDEA, thereby critically comparing their capabilities, limitations and use cases.
- TFSEMDF's application in a complex, context-loaded technology domain, such as NRENs, showed its prowess at performing longitudinal and transversal TF in such technology domains. Exploring whether other fields that technology and contextual factors heavily influence, such as the socio-politically charged and technologically complex domain of AI, would also benefit from this approach would be interesting (Ashok et al., 2022; Chatterjee & Bhattacharjee, 2020; Crawford, 2021).
- The autoregressive NREN model instantiation exhibited capability in conducting longitudinal TF within the NREN technology domain. However, its overall structural performance suggests a need for more in-depth investigation into this type of application of the TFSEMDF framework. Exploring more sophisticated longitudinal SEM techniques like CLPM (Hamaker et al., 2015) and DSEM (Asparouhov et al., 2018) could enhance understanding and effectiveness in this domain. These advanced methodologies may offer improved insights and accuracy in modelling the temporal dynamics and causal relationships inherent in longitudinal TF studies.
- This study did not consider SEM post-processing, specifically the forecasting of TF output metrics with datasets distinct from those utilised in constructing and testing the

NREN model instantiations. Future studies could address this gap and explore strategies to mitigate PLS-SEM overfitting in TF applications characterised by many technology and context-related indicators and constructs. Such investigations would enhance the model's predictive accuracy and generalisability, particularly in complex, indicator- and construct-rich environments.

This study's development, application, and assessment of the TFSEMDF framework have opened new avenues in applying SEM for context-sensitive DF in TF. Despite the study's success in conceptualising, implementing, and evaluating this framework, there remains significant scope for further exploration and enhancement. The untapped potential of this approach, particularly the integration context sensitivity into SEM for TF, suggests vast opportunities for future research and advancements in this area.

"There are many methods for predicting the future. For example, you can read horoscopes, tea leaves, tarot cards, or crystal balls. Collectively, these methods are known as 'nutty methods.' Or you can put well-researched facts into sophisticated computer models, more commonly referred to as 'a complete waste of time.'"
(Source: Scott Adams, the "Dilbert" comic strip creator, n.d.)

APPENDIX A – PLS REGRESSION ANALYSIS FOR TFSEMDF

A.1 INTRODUCTION

PLS regression, conceived by Herman Wold in the 1960s, is positioned uniquely between supervised and unsupervised learning methodologies. Initially applied within scientific research, PLS regression has diversified its applications, extending its utility to chemometrics, bioinformatics, and environmental modelling (Wold et al., 2001). Its chief merit lies in its adeptness at handling scenarios involving multiple correlated predictors, a common characteristic of intricate datasets. PLS regression accomplishes this by creating predictor variables called components through linear combinations of the original predictors, effectively simplifying the data structure (Abdi, 2010; Rosipal & Krämer, 2006).

One distinguishing feature of PLS regression is its focus on components that exhibit correlations with the response variable. In contrast to principal component regression, which primarily seeks to account for variance among predictors, PLS prioritises forecasting outcomes over comprehending the underlying data structures (Wold, 1966). The methodology operates iteratively, incrementally extracting one component at a time to enhance the model's predictive capacity. This iterative process offers a systematic means of constructing models, with the number of components extracted determined by the model's predictive performance (Wold, 1966).

PLS regression has found a niche in fields like chemometrics due to its prowess in managing sets of predictors afflicted by multicollinearity, notably in the analysis of spectroscopic data (Wold & Sjöström, 1977). Furthermore, PLS regression holds a distinct advantage over conventional regression methods when confronted with scenarios characterised by a plethora of predictors and limited sample sizes. This adaptability is especially valuable in the domain of genomic data analysis, which aids in elucidating the relationships between genetic data and phenotypic traits (Boulesteix & Strimmer, 2007). Consequently, PLS regression is an indispensable tool for data scientists and statisticians grappling with high-dimensional datasets.

A.2 APPLICATION OF PLS REGRESSION IN SEM ANALYSIS

The application of PLS within the SEM domain constitutes a significant advancement, particularly in business research. SEM, originally conceptualised by Jöreskog in 1973, initially relied on covariance-based estimation techniques, with the LISREL program, developed by Jöreskog in 1975, gaining widespread acceptance (Haenlein & Kaplan, 2004; Jöreskog, 1975). However, the introduction of PLS, pioneered by Wold in 1975 under the moniker NIPALS, ushered in a paradigm shift towards variance-based or component-based methodologies. Unlike covariance-based approaches, which aim to minimise the disparity between observed sample covariances and those predicted by the model, PLS regression, also known as Projections to Latent Structures, prioritises the maximisation of explained variance among variables (Haenlein & Kaplan, 2004).

In business research, PLS regression confronts the unique challenges posed by survey data, elevating its utility compared to traditional covariance-based regression techniques. As articulated by Vinzi et al. (2010), business surveys often present issues such as low response rates, incomplete item responses, and highly correlated survey items, all of which can limit the effectiveness of conventional covariance-based methods.

These challenges become particularly pronounced when dealing with modest sample sizes and needing more data. Additionally, multicollinearity can exacerbate the margin of error associated with regression coefficients, potentially leading to the exclusion of pertinent predictors from the model (Haenlein & Kaplan, 2004). Conversely, PLS regression demonstrates both stability and precision in navigating the intricacies of such complex data scenarios.

Furthermore, PLS regression's inherent capability to model correlated variables bestows a distinct advantage within SEM, especially when grappling with intricate model structures like those encountered in the instantiation of the TFSEMDF. This attribute proves invaluable in the context of business research, where a frequent objective involves the examination of outcomes and their interdependencies. PLS's adaptability and resilience are exceptionally well-suited for data analysis characterised by multicollinearity partial or missing responses – common hurdles encountered in business and social science research (Vinzi et al., 2010).

PLS regression has garnered acclaim as a flexible and well-suited tool within SEM, adept at addressing the complexities inherent in business research. Its emphasis on maximising explained variance, capacity to handle correlated variables, and robustness when confronted with the challenges posed by survey data positions it as a valuable alternative to traditional covariance-based methodologies. The emergence of PLS by Wold and its subsequent integration into SEM underscores the evolving landscape of analytical techniques tailored to meet the demands of business research (Haenlein & Kaplan, 2004; Vinzi et al., 2010).

A.3 Suitability of PLS Regression in TFSEMDF

The selection of PLS regression over covariance-based techniques in evaluating TFSEMDF model instantiations emerges from a confluence of distinct yet interrelated considerations. Central to this choice is the regression analysis's primary objective: constructing a predictive TF model. Further, TFSEMDF model instantiations endeavour to synthesise a robust theoretical framework, amalgamating multiple theoretical perspectives. A critical aspect of these models lies in the varied relationships between latent constructs and their respective indicators, manifesting in formative and reflective modes, as Haenlein and Kaplan (2004) explained. Additionally, the unique capability of PLS regression to model the influence of first-order factors with reflective indicators on several second-order factors, as highlighted by Pateli (2009), significantly contributes to its preference. The methodological advantages of PLS regression are further underscored by its comparatively lenient sample requirements, in contrast to the stringent demands of classical covariance-based regression techniques (Pateli, 2009).

Implementing PLS regression in SEM offers a notable advantage: the algorithm's partial nature significantly eases the sample size requirements compared to the more demanding needs of traditional covariance-based SEM techniques, such as LISREL, as Chin and Newsted (1999) have documented. Wold (1980) argued convincingly for the superior suitability of PLS regression in facilitating exploratory model searches. This attribute aligns particularly well with the nature of typical TFSEMDF model instantiations, which primarily engage in exploratory analyses of the interplays between technology-related and context-related variables, diverging from the conventional hypothesis testing approach in accepted SEM models, a methodology extensively supported by Staphorst et al. (2013, 2014, 2016a, 2016b).

Dash and Paul (2021) conducted a comparative analysis of the efficacy of PLS regression in SEM for TF against CB-SEM, basing their work on the TFSEMDF framework developed in this study and presented in the prior works of Staphorst et al. (2013, 2014, 216a, 2016b). Dash and Paul's (2021) study scrutinised the disparities between the PLS and the Consistent PLS algorithms, drawing on data from various countries. Their findings revealed that PLS-SEM typically yielded higher item loadings than CB-SEM. Furthermore, the employment of a consistent PLS algorithm showed structural relationships akin to those observed in CB-SEM (Dash & Paul, 2021). A critical revelation of this study was the demonstration of notable values of AVE and Composite Reliability in PLS-SEM, underscoring the reliability and validity of its constructs. While CB-SEM exhibited superior model fit indices, with its evolving metrics, PLS-SEM emerged as more adept for composite-based models. This proficiency renders PLS-SEM particularly apt for TF endeavours within the TFSEMDF framework.

The choice to deploy PLS regression analysis for TFSEMDF model instantiations is based on its superior flexibility in handling factor structures and its aptness for composite-based models. This decision finds robust backing in empirical research, exemplified by studies such as Dash and Paul (2021), which collectively reinforce the efficacy of PLS in SEM within the domain of TF. PLS regression is particularly pertinent in this context, as it meets the need for adaptable methodologies capable of incorporating a spectrum of theoretical perspectives and managing complex model relationships.

A.4 PLS Regression Process

Vinzi et al. (2010) propose that PLS regression comprises two primary stages, each consisting of a sequence of distinct steps, as detailed below (Staphorst, 2010, 2015):

Stage 1 – Iterative Estimation of Weights and Latent Variable Scores:

This stage involves recursively refining a set of outer weights ($w_{\xi,i}$ and $w_{\eta,j}$) and inner weights ($e_{\xi,n}$ and $e_{\eta,p}$) used in the approximation of the values the exogenous and endogenous latent constructs, also frequently referred to as the scores for these constructs (Vinzi et al., 2010). The recursive process, detailed below, terminates when convergence is attained in the weight estimates:

1. **Initial Weight Assignment:** The initial outer weight $w_{\xi,i}$ for the exogenous latent construct ξ_n (for all applicable values of n), and $w_{\eta,j}$ for the endogenous latent construct η_m (for all relevant values of m), can be set to 1, calculated using principal component analysis or derived from factor analysis loadings (Vinzi et al., 2010).
2. **Outer Model Estimation:** Exterior estimates for each latent construct in the model are computed as weighted linear combinations of their respective measurement indicators. For exogenous latent construct ξ_n (for all applicable values of n), the calculation is performed using Equation (A.1), where $w_{\xi,i}$ is the outer weight associated with measurement indicator X_i:

$$\hat{\xi}_n = \sum_i w_{\xi,i} X_i \tag{A.1}$$

Equation (A.2) computes an estimate for endogenous latent construct η_m (for all applicable values of m), where $w_{\eta,j}$ is the outer weight associated with measurement indicator Y_j.:

$$\hat{\eta}_m = \sum_j w_{\eta,j} Y_j \tag{A.2}$$

3. **Weight Updating:** Next, the outer weights are subject to refinement governed by one of two distinct modes, each applicable to the nature of the indicators associated with the latent constructs. The selection of the mode is contingent upon the reflective or formative characteristic of the indicators and is delineated as follows:

   a. **Mode A for Reflective Indicators:** This mode employs simple linear regression to calculate weight $w_{\xi,i}$ (for all applicable values of i) for reflective indicators X_i. A similar approach calculates $w_{\eta,j}$ when Y_j is reflective.
   b. **Mode B for Formative Indicators:** Here, multiple linear regression is used to determine weight $w_{\xi,i}$ (for all applicable values of i) for formative indicators X_i. The same methodology applies for calculating $w_{\eta,j}$ when Y_j is formative.

4. **Inner Latent Variable Approximation:** The next step in Stage 1 involves calculating internal estimates for endogenous latent construct η_m (for all applicable values of m) using weighted linear combinations of the outer estimations, as per Equation (A.3). The inner weights $e_{\xi,n}$ and $e_{\eta,p}$ associated with ξ_n and η_p, respectively, are determined using centroid, factor, or path weighting schemes.

$$\hat{\eta}_m = \sum_n e_{\xi,n}\hat{\xi}_n + \sum_{p \neq m} e_{\eta,p}\hat{\eta}_p \tag{A.3}$$

5. **Test for Weight Calculation Convergence:** The recursive algorithm oscillates between the external and internal approximation procedures, optimising the variance the latent constructs explain until the delta of successively calculated outer weights satisfies a predetermined convergence criterion. Wold (1975) recommends setting this criterion to 10^{-5}, ensuring a balance between rapid convergence and minimal error in weight estimation.

Stage 2 – Calculation of Loadings and Path Coefficients:

Upon achieving convergence in the iterative procedures of Stage 1, Stage 2 commences with the determination of the SEM model's parameters. This stage is bifurcated into two integral components, focusing on the quantification of the measurement model loadings and the ascertainment of the structural model path coefficients:

1. **Loadings Calculation for the Measurement Part:** Loadings λ_{xi} for reflective indicators of exogenous latent constructs and λ_{yj} for endogenous latent constructs are calculated by ordinary least squares regressing the indicators on the final latent variable scores. In the case of formative indicators, the loadings π_{xi} and π_{yj} are equivalent to the respective outer weights determined during Stage 1.
2. **Path Coefficients Calculation for the Structural Part:** In the structural part, path coefficients γ_c and β_d (for all applicable values of c and d) are estimated by regressing the endogenous latent variable scores on their respective exogenous predictors (for γ_c) and on other endogenous variables (for β_d), using ordinary least squares regression.

A.5 PLS Regression for TFSEMDF using SmartPLS

This study utilised the SmartPLS freeware software (Ringle et al., 2022) to develop both the initial and refined SEM path diagrams of the NREN model instantiations proposed by Staphorst et al. (2013, 2014, 2016a, 2016b) and for the computation of all loadings and path coefficients using PLS regression (Staphorst et al., 2015). The configuration of SmartPLS was tailored to normalise (Ringle et al., 2022) all secondary data sourced from TERENA (2011, 2012), thereby accommodating the assorted scaling methods and ranges utilised in the data processing phase.

Furthermore, SmartPLS proved instrumental in aiding the evaluation of reliability and validity test criteria (Staphorst, 2010; Staphorst et al., 2017), as elaborated in Appendix B.

APPENDIX B – TFSEMDF RELIABILITY AND VALIDITY ANALYSIS

B.1 INTRODUCTION

Reliability primarily encompasses repeatability and internal consistency in quantitative research, as Zimund (2003) explains. Researchers must assess whether their results consistently represent the target population over time and whether their methodology can reliably reproduce these results, a point Golafshani (2003) emphasises. Additionally, Vinzi et al. (2010) argue that understanding measurement errors, which consist of random and systematic parts, is crucial for achieving complete reliability, defined as the absence of random errors. This understanding is vital for ensuring that research outcomes accurately reflect the investigated phenomena rather than being mere by-products of measurement processes.

Further, the relationship between reliability and validity in quantitative research is paramount. Golafshani (2003) and Zimund (2003) highlight that while reliability pertains to consistency, validity addresses the accuracy and truthfulness of the research instrument, ensuring it measures the intended constructs. Litwin (2017) underscores the necessity of reliability for validity, clarifying that while a reliable measure may not always be valid, validity cannot exist without reliability. George and Mallery (2019) recommend several strategies, including pilot testing and using established measurement tools to enhance reliability. These strategies not only refine research tools but also ensure the generation of consistent and replicable results, a cornerstone for robust and credible scientific inquiry.

B.2 OVERVIEW OF RELIABILITY AND VALIDITY ANALYSIS FOR PLS-SEM

The application of PLS regression in SEM allows for a comprehensive evaluation of reliability, spanning three fundamental domains: internal consistency, convergent reliability, and discriminant reliability. Hair et al. (2017) assert that researchers should assess internal consistency reliability using methods like Cronbach's alpha or Composite Reliability to ensure uniformity in responses across various construct items. Simultaneously, convergent reliability, measured by the AVE, quantifies the agreement level among items within a construct. Distinct from these, discriminant reliability, a concept developed by Fornell and Larcker (1981), involves determining the uniqueness of a construct in contrast to others by comparing its AVE with squared inter-construct correlations. Moreover, Henseler et al. (2015) draw attention to the unique approach required for formative constructs in PLS-SEM, which diverges from the conventional methods used for reflective constructs. This approach necessitates critically analysing the indicators' weights and loadings to verify their significance and relevance to the construct they represent. Such a detailed and context-specific approach to reliability in PLS regression-based TFSEMDF solidifies its standing as an effective and robust methodology for complex model evaluations across various research disciplines.

B.3 TFSEMDF Reliability and Validity Analysis Process

In the assessment of the reliability and validity of PLS regression analysis for SEM models, such as those created for TFSEMDF, researchers typically initiate the process by scrutinising the item measures within the research survey, often designated as the measurement or outer model (Vinzi et al., 2010). Subsequently, they evaluate the SEM's structural portion, or inner model (Vinzi et al., 2010). This sequential approach stems from the rationale that if the measurement indicators in the SEM model lack accuracy and representativeness, further examination of the reliability and validity of the structural portion is unnecessary.

B.3.1 Measurement Portion (Outer Model) Reliability and Validity Analysis

To thoroughly evaluate the reliability and validity of the measurement portion, also known as the outer model, of a TFSEMDF model instantiation, the study executed a comprehensive array of tests as proposed by Vinzi et al. (2010). These tests are essential for ascertaining the measurement model's accuracy and consistency, ensuring that the constructs are measured in a manner that is both reliable and valid (Staphorst, 2010; Staphorst et al., 2017):

- **Indicator Reliability:**
- The assessment of Indicator Reliability within the study's framework focuses particularly on reflective indicators, represented as X_i and Y_j, for exogenous and endogenous latent constructs, respectively. Following the conceptualisation by Vinzi et al. (2010), this reliability measure gauges the proportion of variance in a measurement indicator

attributable to its associated latent construct. Adhering to the widely accepted criterion that over 50% of a reflective indicator's variance should be answerable by its corresponding latent construct, the study necessitated that the loadings λ_{xi} and λ_{yj}, which link indicators X_i and Y_j to exogenous latent construct ξ_n and endogenous latent construct η_m respectively, should meet or exceed the threshold of √0.5, equivalent to 0.707. However, the study acknowledged the potential for tolerance of weaker reflective indicators, particularly those with loadings below this threshold, in contexts such as exploratory studies focused on TFSEMDF model instantiations. Nonetheless, it aligned with Vinzi et al.'s (2010) recommendation to exclude reflective indicators from the SEM model if their loadings fell below 0.4. It is crucial to note that the concept of Indicator Reliability does not extend to formative indicators, as these can demonstrate low correlations with their associated latent constructs yet still contribute substantially to the overall variance of the construct, as Vinzi et al. (2010) postulate. As such, the study refrained from assessing Indicator Reliability for formative indicators. Instead, it evaluated the relative contributions of each formative indicator within the indicator set corresponding to a specific latent construct. Moreover, the study selected to retain all formative indicators in the analysis, regardless of their low relative contributions to the associated latent construct, to ensure a comprehensive evaluation of the model's components.

- **Construct Reliability:** The Indicator Reliability metric aims to ascertain the sufficiency with which a reflective indicator measures a latent construct, as articulated by Vinzi et al. (2010). It is, however, also crucial to evaluate the collective efficacy of the reflective indicators associated with a latent construct to ensure joint measurement adequacy, necessitating the assessment of Construct Reliability, or internal consistency, for each latent construct within an SEM model (Vinzi et al., 2010). Construct Reliability is often verified within business research using Cronbach's Alpha to ensure a robust mutual association among reflective indicators assigned to the same latent construct. However, the Composite Reliability measure is becoming more prevalent in academic circles for its ability to address the limitations of Cronbach's Alpha, which neglects the differential impact of the reflective indicators' respective loadings (Vinzi et al., 2010). In terms of the SEM framework outlined in Chapter 3 for TFSEMDF, one calculates the Composite Reliability measure for the exogenous latent construct ξ_n as follows, according to Fornell and Larcker (1981):

$$\rho_{\xi,n} = \frac{(\sum_i \lambda_{xi})^2}{(\sum_i \lambda_{xi})^2 + \sum_i \mathrm{var}(\delta_i)} \tag{B.1}$$

For the endogenous latent construct η_m, the definition is as follows, also from Fornell and Larcker (1981):

$$\rho_{\eta,m} = \frac{\left(\sum_j \lambda_{yj}\right)^2}{\left(\sum_j \lambda_{yj}\right)^2 + \sum_j \mathrm{var}(\varepsilon_j)} \tag{B.2}$$

The analysis of TFSEMDF model instantiations harnessed the capabilities of the SmartPLS software package, which generates both Cronbach's Alpha and Composite Reliability measures (Ringle et al., 2022). The study, however, elected to adopt exclusively the Composite Reliability measures. It deemed a threshold level of 0.6 for Composite Reliability as an acceptable standard for evaluating Construct Reliability, aligning with the benchmarks established by Vinzi et al. (2010). The study acknowledged that testing for internal consistency is inapplicable for latent constructs with formative indicators, as these indicators inherently display low mutual association, a concept detailed by Hulland (1999).

- **Convergent Validity:** Convergent Validity measurement assesses the correlation among responses garnered through distinctly different methods that gauge the same construct (Vinzi et al., 2010). Evaluating Convergent Validity for reflective indicators of latent constructs in an SEM model requires an analysis of the AVE. For the reflective indicators of an exogenous latent construct ξ_n, one defines the AVE using the formula posited by Fornell and Larcker (1981):

$$AVE_{\xi,n} = \frac{\sum_i \lambda_{xi}^2}{\sum_i \lambda_{xi}^2 + \sum_i \mathrm{var}(\delta_i)} \tag{B.3}$$

Similarly, for an endogenous latent construct η_m, the AVE is defined as:

$$AVE_{\eta,m} = \frac{\sum_j \lambda_{yj}^2}{\sum_j \lambda_{yj}^2 + \sum_j \mathrm{var}(\varepsilon_j)} \tag{B.4}$$

- Equations (B.3) and (B.4) illustrate that AVE calculates the proportion of variance in a latent construct's indicators that the construct captures, as opposed to the total variance, which also encompasses measurement error. This study adopted a threshold of 0.5 for AVE, as Vinzi et al. (2010) recommended. An AVE falling below this threshold indicated a predominance of measurement error over the variance attributed to the indicators. Convergent validity does not apply to formative indicators, as these indicators do not have to be strongly interrelated (Vinzi et al., 2010).
- **Discriminant Validity:** Discriminant Validity assessment within the measurement portion scrutinises the extent of dissimilarity among measurements yielded by the measurement tool for distinct constructs (Vinzi et al., 2010). Achieving Discriminant Validity entails a condition whereby the shared variance between a latent construct and its indicators, which one derives by taking the square root of its AVE, surpasses the shared variance between the focal latent construct and any other latent constructs in the model. This criterion ensures that each construct is empirically unique and captures phenomena that are not merely reflections of other constructs within the model framework.

B.3.2 Structural Portion (Inner Model) Reliability and Validity Analysis

The study implemented a series of diagnostic tests to determine the reliability and validity of the structural portion, which details the interrelationships between constructs in the SEM path diagram of TFSEMDF model instantiations. As recommended by Vinzi et al. (2010), the following tests are instrumental in ascertaining the integrity of the model's hypothesised relationships, thereby ensuring that the conclusions regarding the dynamics among the constructs are empirically substantiated (Staphorst, 2010; Staphorst et al., 2017):

- **Coefficients of Determination for Endogenous Variables**: The study utilised this metric to evaluate the quality of the structural portion of the SEM of TFSEMDF model instantiations, focusing on the extent of variance in an endogenous construct that related endogenous or exogenous constructs can account for (Vinzi et al., 2010). Vinzi et al. (2010) assert that setting universal thresholds for R^2 values is untenable. However, Falk and Miller (1992) argue that R^2 values under 0.1 signify concerningly weak interrelationships between constructs. The study also embraced a classification system derived from traditional multiple regression techniques, which defines the strength of the relationship between an endogenous construct and associated constructs as weak, moderate, or vigorous, corresponding to less than 30%, between 30% and 70%, and over 70% of variance explained, respectively (Zikmund, 2003).

- **Path Coefficient Significance:** The evaluation of the structural portion of an SEM model mirrors covariance-based multiple regression techniques, employing a bootstrapping procedure to ascertain the significance levels of path coefficients, denoted as γ_c and β_d for all relevant indexes *c* and *d* (Chin, 1998a; Vinzi et al., 2010). Researchers tested the significance of path coefficients (Goodness-of-Fit) using asymptotic *t*-statistics through the bootstrapping functionality in SmartPLS, set to a resampling size of 1000 (Vinzi et al., 2010). The analysis considered various significance levels, interpreting *p*-values calculated from the $t_{(999)}$ distribution greater than $\alpha = 0.10$ as insignificant. Researchers regarded paths with insignificant coefficients or coefficients whose signs contradicted the hypothesised interrelations between constructs in TFSEMDF model instantiations as not substantiating the posited research propositions.
- **Predictive Validity:** This study implemented the Stone-Geisser non-parametric test (Geisser, 1975; Stone, 1975; Chin, 1998b; Vinzi et al., 2010) to judge the predictive validity of the TFSEMDF model instantiations. Researchers commonly use this test with a blindfolding procedure (Vinzi et al., 2010) that necessitates two distinct datasets: one for SEM and another for ascertaining the SEM model's predictive validity. This procedure, which requires specifying an Omission Distance *D* (Vinzi et al., 2010), involves the selective omission of data, whereby the SEM model parameters are estimated using an incomplete dataset. In this study, the omission distance was set to seven, aligning with the default setting in SmartPLS (Ringle et al., 2022). The next phase involves employing the SEM model derived from the incomplete dataset to reconstruct the omitted data. The reconstruction effectiveness, measured by the Stone–Geisser test criterion (Q^2), evaluates how precisely the empirically collected data can be replicated through PLS regression using the SEM model (Vinzi et al., 2010). The criterion is mathematically determined as follows:

$$Q^2 = 1 - \frac{\sum_D E_D}{\sum_D O_D} \tag{B.5}$$

In this formula, E_D denotes the squared prediction error, calculated as the variance between the omitted data during the blindfolding process and their predicted values (Vinzi et al., 2010). Conversely, O_D signifies the squared prediction error derived from the mean of the data that remained post-omission (Vinzi et al., 2010). A Q^2 value greater than zero suggests the SEM model's predictive validity (Chin, 1998a; Vinzi et al., 2010). The Stone–Geisser test criterion manifests in two distinct forms, contingent on the prediction type under investigation: Cross-validated Communality (H^2) and Cross-validated Redundancy (F^2). H^2, typically assessed during the structural portion's validity evaluation, gauges the SEM model's competency in forecasting observable endogenous constructs from their respective latent construct scores (Vinzi et al., 2010). F^2, on the other hand, evaluates the model's efficacy in predicting observable endogenous constructs using latent constructs pertinent to the data block in question (Vinzi et al., 2010). This investigation considered H^2 and F^2 metrics computed via SmartPLS (Ringle et al., 2022).

APPENDIX C – TERENA NREN COMPENDIUM EXCERPTS

C.1 INTRODUCTION

The GÉANT Compendium of NRENs in Europe, known as the TERENA NREN Compendium until 2015 when TERENA and DANTE merged their activities to form the GÉANT Association in 2015 (GÉANT, 2022), is an annual report providing a thorough overview of NRENs in Europe. As outlined in Section 2.5.3, this Compendium primarily focuses on data from the European NRENs connected through the GÉANT network. However, earlier editions also included data from NRENs in other regions, including Africa. The 2022 edition, for instance, was based on information gathered from 40 European NRENs. It assesses technological and contextual aspects of these networks, including organisational structure, network and traffic metrics, service portfolio, user base, European Commission-funded projects, growth trends, digital education services, and international connectivity. These metrics offer insights into the operational efficiency, service diversity, and strategic direction of NRENs, helping to identify trends and forecast future developments in academic and research networking.

A collaborative process involving an annual survey distributed to all European NRENs gathers the data for the compendium. The survey, crafted with guidance from subject experts, covers extensive details about each NREN's network infrastructure, services, and organisational aspects. Responses are complemented with supplementary data, including publicly available information and internal GÉANT data, particularly in trust, identity, and educational services. This data is then analysed by specialists and summarised in the Compendium report, which comprehensively portrays the NREN landscape. Additionally, past and current survey data are accessible online for further analysis. This rigorous data collection and analysis process ensures the Compendiums' accuracy and reliability, making it a valuable strategic resource for decision-making within the NREN community.

As outlined in Section 5.2.1.3, the PLS regression analysis applied to the autoregressive NREN model instantiation shown in Figure 7 utilised the 2011 and 2012 editions of the TERENA Compendiums (TERENA, 2011, 2012) for its indicator data. Conversely, Section 5.2.2.3, Section 6.2.1.4, and Section 6.2.2.4 explain the exclusive reliance on the 2011 TERENA NREN Compendium (TERENA, 2011) for the PLS regression analysis of the cross-sectional (also Chapter 6's baseline) and structurally disarranged NREN model instantiations, depicted in Figure 8 and Figure 9, respectively. The following appendix aims to show selected excerpts of the data available in the 2011 and 2012 TERENA NREN Compendiums (TERENA, 2011, 2012), highlighting the scope and fidelity of the data available in these resources.

C.2 2011 TERENA NREN Compendium Excerpts

The 2011 TERENA NREN Compendium was an indispensable data source for fulfilling Research Objectives 2(a) through 3(b), as outlined in Chapter 3 of this study. Table 1 and Table 2 detail the methodology employed for extracting and processing data from this Compendium, thereby generating the requisite indicator data for the PLS regression analysis pertinent to both the autoregressive, cross-sectional (also Chapter 6's baseline) and structurally disarranged NREN model instantiations. The subsequent sections present illustrative data from the 2011 TERENA NREN Compendium, focusing specifically on funding sources and the computation of core network traffic for the NRENs that contributed to the Compendium survey (TERENA, 2011).

C.2.1 NREN Funding Sources

Figure 10 and Figure 11 below show the data in Graph 4.6.2 and Graph 4.6.3 of TERENA (2011), respectively. Graph 4.6.2 represents the funding composition of European NRENs serviced by GÉANT, i.e. GÉANT partner NRENs, while Graph 4.6.3 provides similar information for the responding NRENs not serviced by GÉANT. The percentage of funding from national governments and public sources for each NREN was of specific importance in these graphs. As described in Table 1 and Table 2, respectively, this was used as measurement data for the reflective indicator *Level of Government Funding (X_2)* of the exogenous context-related construct *Government Influence over the NREN (ξ_1)* at DF Level 0 in Figure 7's autoregressive, Figure 8's cross-sectional (also Chapter 6's baseline model) and Figure 9's structurally disarranged NREN model instantiations of the TFSEMDF framework.

C.2.2 NREN Core Traffic

The PLS regression analyses of this study's autoregressive, cross-sectional (also Chapter 6's baseline) and structurally disarranged NREN model instantiations of the TFSEMDF framework, shown in Figure 7, Figure 8 and Figure 9, respectively, employed TERENA (2011) measurements data for the core traffic in each respondent NREN as reflective indicator data, albeit at differing DF levels and technology indicator related constructs. Ingress and egress traffic in NRENs are measured by GÉANT using the convention depicted in Figure 12. Of specific importance in this study was the combined ingress traffic from customer connections (T_1) and traffic from external networks and peering (T_4).

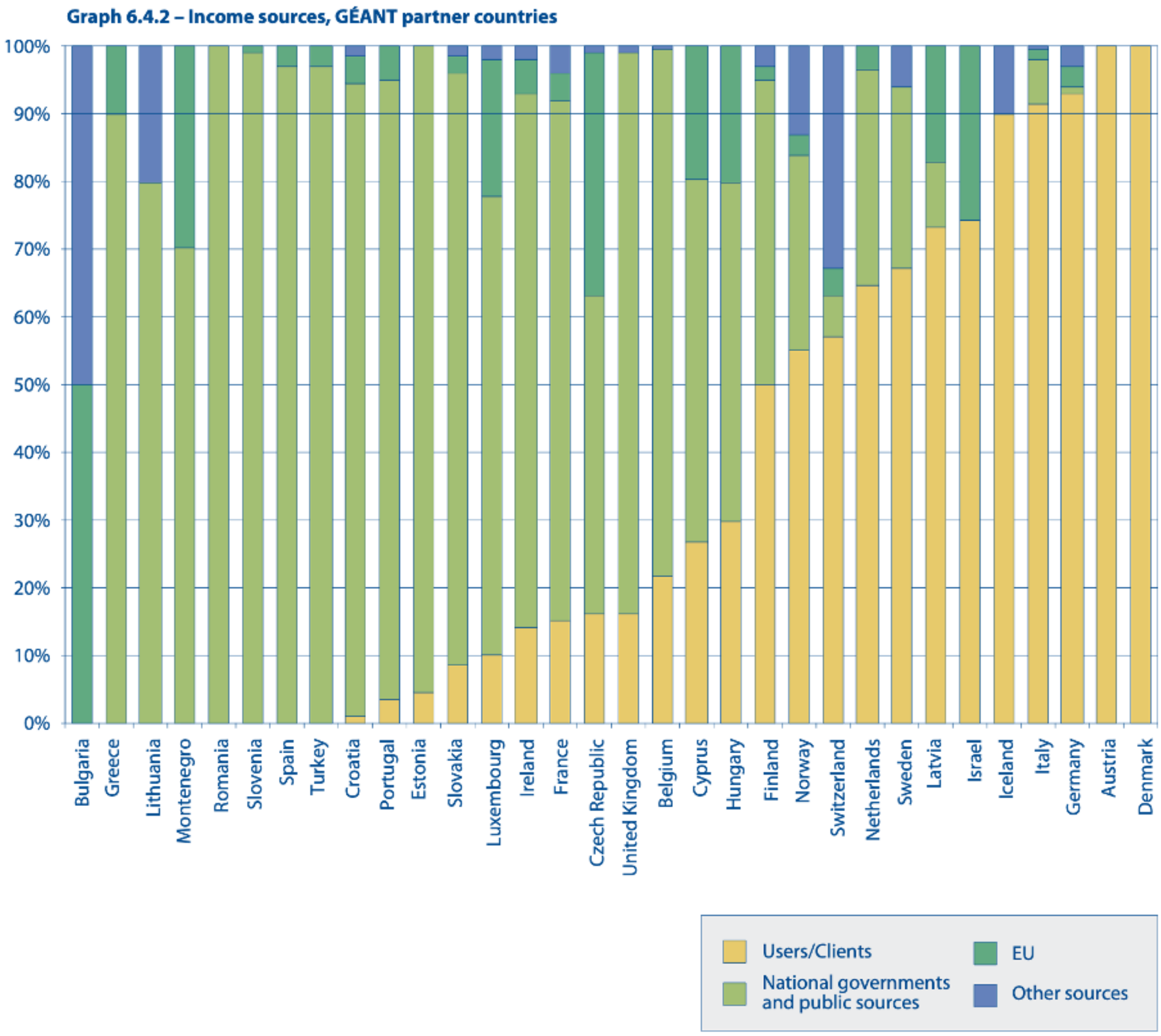


Figure 10: Funding Sources for GÉANT Partners, TERENA (2011) Graph 6.4.2

As described in Table 1, this measured data, which is given in Figure 13 for respondent NRENs with annual core traffic exceeding 3500 TB and in Figure 14 for NRENs with annual core traffic less than 3500 TB, was used as indicator data for the reflective indicator *Level of Core Network Traffic (Y_3)* of the technology-related construct *NREN Advanced Services Capability (η_2)* at DF Level 1 in the autoregressive NREN model instantiation of the TFSEMDF framework depicted in Figure 7. In the case of the cross-sectional (also Chapter 6's baseline) and structurally disarranged NREN model instantiations of Figure 8 and Figure 9, respectively, as explained in Table 2, this measurement data was used for the reflective indicator *NREN Core Traffic Level (Y_4)* for the technology-related construct *NREN Core Traffic Level (η_3)* at DF Level 2.

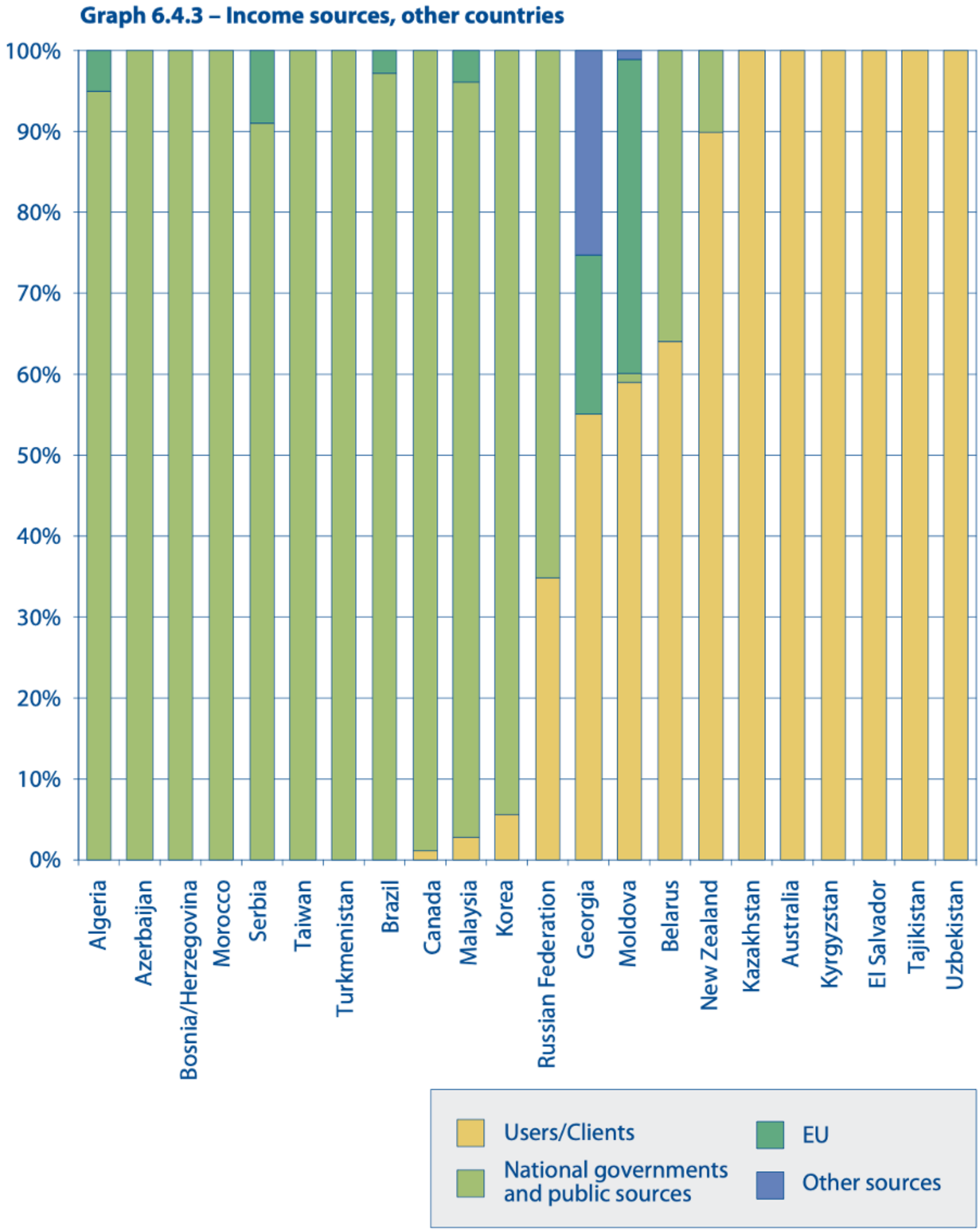
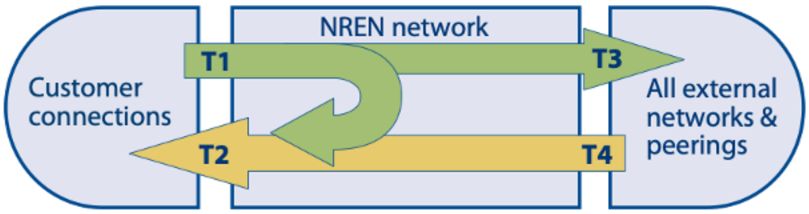


Figure 11: Funding Sources for non-GÉANT Partners, TERENA (2011) Graph 6.4.3

Diagram 4.0.1 – Types of traffic flow



T1	All IP traffic from customer sites and the NREN.
T2	All IP traffic to customer sites and the NREN.
T3	All IP traffic to external networks.
T4	All IP traffic to the NREN from external networks.

External traffic = all traffic to GÉANT, the commercial Internet, Internet Exchanges, etc. (i.e. composed of T3 and T4 above)

Figure 12: NREN Core Traffic Types, TERENA (2011) Diagram 4.0.1

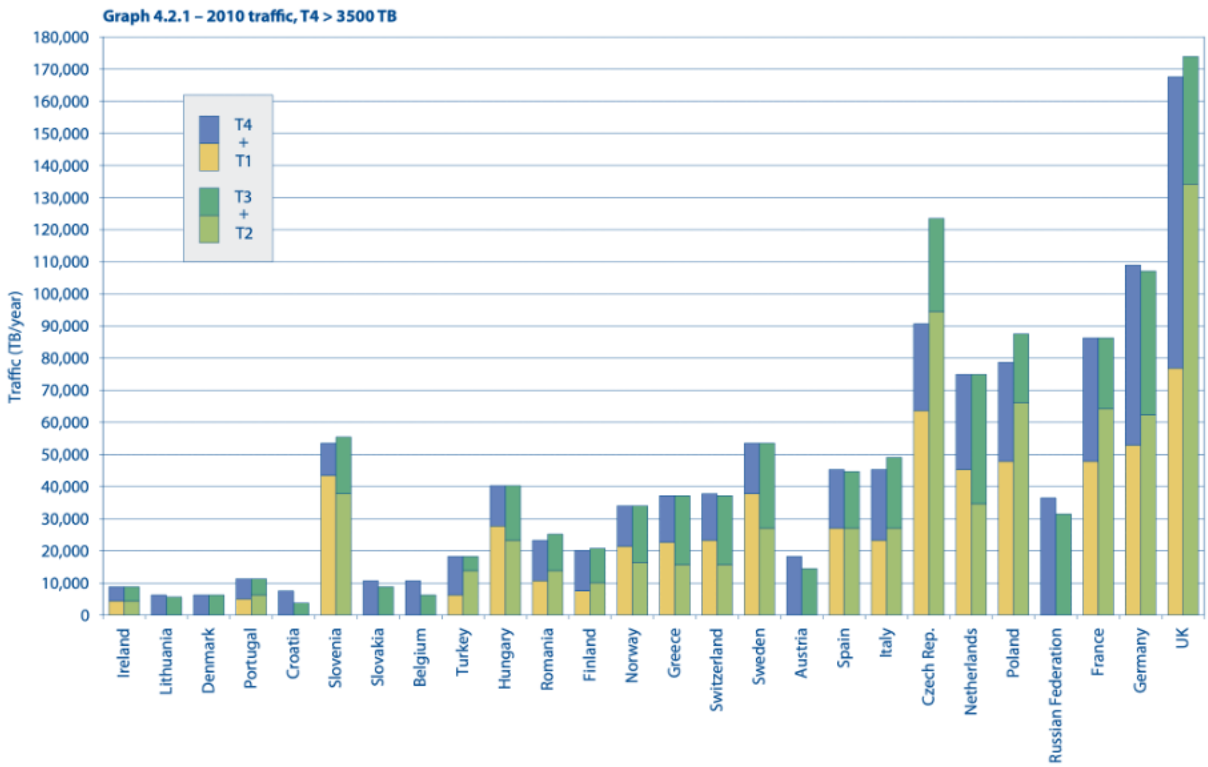


Figure 13: NRENs with Ingress Traffic > 3500TB, TERENA (2011) Graph 4.2.1

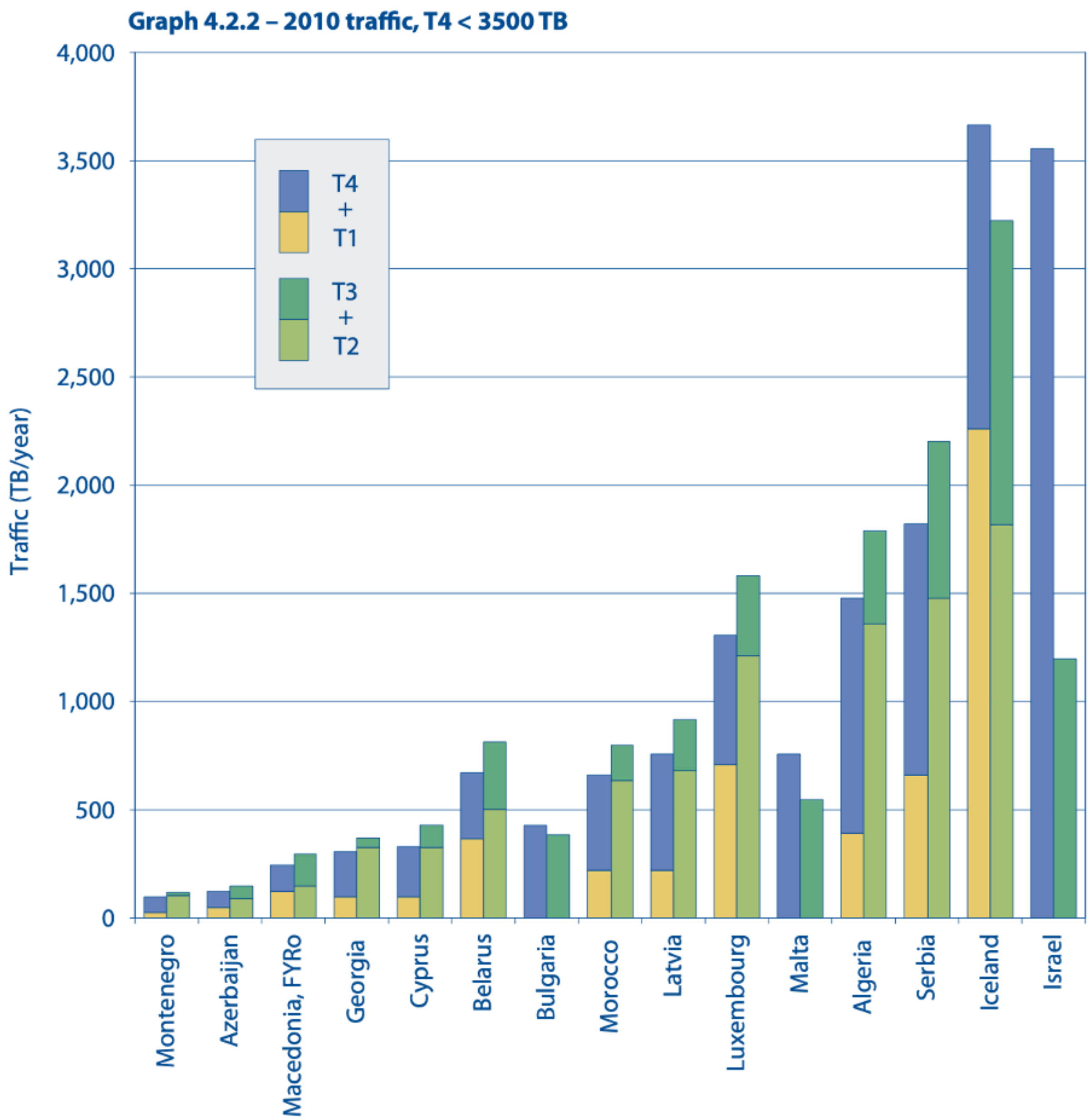


Figure 14: NRENs with Ingress Traffic < 3500TB, TERENA (2011) Graph 4.2.2

C.3 2012 TERENA NREN Compendium Excerpts

The 2012 TERENA NREN Compendium was the exclusive data source for accomplishing Research Objective 2(b), as detailed in Chapter 3 of this study. The methodology employed for extracting and processing data from the 2012 Compendium, as referenced in TERENA (2012), is comprehensively outlined in Table 1. This process was fundamental in producing the requisite indicator data essential for the PLS regression analysis for the autoregressive NREN model instantiation. The following section includes an excerpt from the 2012 TERENA NREN Compendium that provides the types and quantities of institutions connected to the NRENs that responded to the 2012 survey.

C.3.1 INSTITUTIONS CONNECTED TO NRENS

As explained in indicator data composition Table 1 for Figure 7's autoregressive NREN model instantiation of the TFSEMDF framework, the measurement data for the numbers of sites connected per type of institution provided in TERENA (2012) Table 2.2.1, and reflected in Table 7, Table 8 and Table 9 below, was used as indicator data for the reflective indicator *Forecasted Number of Institutions Connected by the NREN (Y_6)* for the technology-related construct *Forecasted NREN Reach (η_4)* at DF Level 2. The study took similar data from TERENA (2011) for the reflective indicator *Current Number of Institutions Connected by the NREN (Y_5)* of the technology-related construct *Current NREN Reach (η_3)* at DF Level 2.

Table 18: Institutions Connected to NRENs, TERENA (2012) Table 2.2.1 - Part A

Legend for Table 2.2.1

(dark green)	All or nearly all (more than 80%) of the institutions are connected by the NREN
(light green)	More than half (between 60 and 80%) of the institutions are connected by the NREN
(yellow)	About half (between 40 and 60%) of the institutions are connected by the NREN
(light orange)	Less than half (between 20 and 40%) of the institutions are connected by the NREN
(red)	None or very few (less than 20%) of the institutions are connected by the NREN
(grey)	Unknown or not applicable

Table 2.2.1 – Approximate market shares, number of connected institutions

Country	Universities	Research institutes	Institutes of further education	Secondary schools	Primary schools	Libraries, museums, archives, cultural institutions	Hospitals (other than university hospitals)	Government departments (national, regional, local)
GÉANT partner countries								
Austria	36	29	1			11	4	35
Belgium	65	37	6	5	0	11	16	50
Bulgaria	22	50	2	1500	500	15	5	
Croatia	141	36	46	426	904	10	15	9
Cyprus	6	2	3					
Czech Republic	26	25	10	105	14	50	40	39
Denmark	8	12	7	0	0	5	2	4
Estonia	21	21	16	126	21	84	1	3
Finland	51	12				6		8
France	445	368	335			12	6	20
Germany								
Greece	43	28	146	4390	9803	12	0	727
Hungary	25	73	28	34	0	210	50	4
Iceland	9	11	2	1		1		
Ireland	25	10	10	800	3200	0	0	8
Israel	12	5	0	0	0	0	0	0
Italy	139	203	0	5	1	35	53	5
Latvia	15	13	4	3	0	3	0	0
Lithuania	42	31	55	551	37	32	4	35

Table 19: Institutions Connected to NRENs, TERENA (2012) Table 2.2.1 - Part B

Table 2.2.1 – continued

Country	Universities	Research institutes	Institutes of further education	Secondary schools	Primary schools	Libraries, museums, archives, cultural institutions	Hospitals (other than university hospitals)	Government departments (national, regional, local)
GÉANT partner countries								
Luxembourg	6	22	1	59	200	12	1	14
Macedonia, FYRo	19	5	0	0	0	50	0	1
Malta	1	3	2					
Montenegro	19	2	1			2		1
Netherlands	14	32	64	0	0	19	12	0
Norway	8	81	56	4	4	17	0	0
Poland	180	194						
Portugal	42	12	0	0	0	3	0	13
Romania	50	55	10	350	140	40	10	30
Serbia	87	41	9	13	5	20	5	2
Slovakia	38	20	7	250	100	6		
Slovenia	4	54	20	156	528	205	0	18
Spain	90	170	0	0	0	25	50	75
Sweden	30	4	9			19		20
Switzerland	43	10	3	2	0	0	0	6
Turkey	899	18				2		6
United Kingdom	200	35	550			9	0	6
Other countries								
Algeria	57	25	34					4
Armenia	2	35				4	4	
Azerbaijan		30				4		
Belarus	10	57				17	5	8
Bosnia/Herzegovina	25	1	1					0
Georgia	10	5	18	3		4	3	2
Moldova	5	36	1	1	0	10	5	1
Morocco	15	8	80	0	0	2	0	2
Russian Federation	250	240						

Table 20: Institutions Connected to NRENs, TERENA (2012) Table 2.2.1 - Part C

Table 2.2.1 – continued

Country	Universities	Research institutes	Institutes of further education	Secondary schools	Primary schools	Libraries, museums, archives, cultural institutions	Hospitals (other than university hospitals)	Government departments (national, regional, local)
Other countries								
Ukraine	64	16				4		1
Australia	41	24	13	225	223	9		1
Canada	89	70	184	1300	700	26	62	50
Kazakhstan	51	7		5			1	2
Korea	51	61	0	1	0	3	1	18
Kyrgyzstan	27	17	0	34	0	3	3	1
New Zealand	8	13	13	81	93	5	1	2
Singapore	4	2						1
Taiwan	120	20	20	500	1000	5	5	20
Tanzania	4	1	4					

APPENDIX D – RESEARCH DATA REPOSITORY

D.1 INTRODUCTION

This study engaged in a systematic process of using and creating diverse research data sources. Central to this exploration was the utilisation of openly available data from the TERENA NREN Compendiums (TERENA, 2011, 2012), which provided a wealth of information on the technology and contextual trends of NRENs. These datasets were instrumental in generating indicator data for the PLS regression analyses (as described in Table 1 and Table 2), a key component in assessing the autoregressive, cross-sectional (also Chapter 6's baseline) and structurally disarranged NREN model instantiations of the TFSEMDF framework. Also crucial in these analyses were the SmartPLS SEM path diagrams constructed for each NREN model instantiation, as depicted in Figure 7, Figure 8 and Figure 9.

A significant outcome of the study was the creation of four core research publications and two supplementary publications. These works encapsulate the detailed findings and methodologies employed in each study phase. The core and supplementary publications from this study are summarised in Section 1.5.4 and Section 1.5.5, respectively.

The commitment of this study to the principles of open science data, especially in the context of the Research Data Alliance (RDA) guidelines, is exemplified in its approach to data dissemination. The RDA promotes the sharing of data to foster innovation and collaboration across various scientific disciplines (Research Data Alliance, n.d.). By adopting this ethos, the study endeavours to make its research data, including TERENA NREN Compendiums, SEM path diagrams for the TFSEMDF model instantiations, processed indicator data, and all associated publications, available to interested researchers and NREN stakeholders in a manner that will strictly comply with the University of Pretoria's policies and the publications' copyright requirements.

The study chose Figshare (Figshare, n.d.) as the repository platform for its data because it aligns with the principles of Findability, Accessibility, Interoperability, and Reusability (FAIR), which are foundational to successful open data practices (Wilkinson et al., 2016). Figshare effectively manages the transition from restricted to open access, ensuring the data remains secure and confidential during the examination phase and becomes widely available afterwards. This platform is particularly suited for the study's needs due to its strong emphasis on data discoverability and usability, capabilities for handling diverse data formats, and commitment to long-term data preservation. By utilising Figshare, the study adheres to best practices in data management and contributes to the ethos of open scientific research, facilitating future scholarly inquiries and advancements.

D.2 NAVIGATING THE RESEARCH DATA REPOSITORY

This study utilised TERENA NREN Compendiums (TERENA, 2011, 2012) as the primary data source. The second and third research phases used these Compendiums to generate indicator data sets to perform PLS regression analysis of the SEM path diagrams for the various NREN model instantiations, including the autoregressive, cross-sections (also Chapter 6's baseline), and structurally disarranged model instantiations. These TERENA NREN Compendiums, along with the indicator data sets, SEM path diagrams, and the core and supplementary publications produced by this study, were stored in the various Figshare (Figshare, n.d.) data collections, as described in Table 21, configured as a private repository.

Table 21: Research Data Repository Hosted on Figshare

Figshare Collection Name	Figshare Collection Contents
Staphorst PhD - TERENA NREN Compendiums	• TERENA (2011) • TERENA (2012)
Staphorst PhD - SmartPLS Path Diagrams	• Autoregressive NREN model instantiation SmartPLS path diagram • Cross-sectional NREN model instantiation SmartPLS path diagram • Baseline NREN model instantiation SmartPLS path diagram • Structurally disarranged NREN model instantiation SmartPLS path diagram

Staphorst PhD - SmartPLS Indicator Data	• Autoregressive NREN model instantiation SmartPLS indicator data • Cross-sectional NREN model instantiation SmartPLS indicator data • Baseline NREN model instantiation SmartPLS indicator data • Structurally disarranged NREN model instantiation SmartPLS indicator data
Staphorst PhD - Core Publications	• Staphorst et al. (2013) • Staphorst et al. (2014) • Staphorst et al. (2016a) • Staphorst et al. (2016b)
Staphorst PhD - Supplementary Publications	• Staphorst et al. (2015) • Staphorst et al. (2017)

Figshare stands at the forefront of research data management and dissemination, offering a sophisticated web-based platform tailored to the needs of the modern research community (Figshare, n.d.). It facilitates the storage of a wide array of scholarly outputs, including datasets, figures, and papers, in an organised manner through projects and collections. Projects in Figshare allow researchers to manage and collaborate on their data, maintaining coherence in large volumes of research outputs. At the same time, collections enable the grouping and showcasing related research items, enhancing discoverability and thematic categorisation.

Figshare significantly contributes to the open science paradigm by adhering to the FAIR principles (Figshare, n.d.). It ensures that digital assets are not only findable and accessible but also interoperable with different datasets and workflows and reusable across various research contexts. This alignment with FAIR principles underpins Figshare's commitment to transparency, reproducibility, and impact in scientific research.

Figshare has developed a tiered data access model, categorising it into three levels: private, embargoed, and public (Figshare, n.d.). Private access enables individual scholars or research collectives to handle data under stringent confidentiality, a provision particularly indispensable in the formative stages of research or when the data are sensitive. Embargoed access serves as a strategic tool, orchestrating the public dissemination of data to align meticulously with the timelines of scholarly publications or the directives of funding authorities. Public access, a cornerstone of the open science framework, ensures that data are openly available to the extensive research community and the public, thus democratising access to information and cultivating a collaborative and inclusive environment.

Through these features, Figshare serves as a repository and potent catalyst for collaborative research, ensuring that the scientific community's collective efforts are more integrated, accessible, and impactful (Figshare, n.d.). Its role in promoting open science is pivotal, propelling the global research community towards more significant innovation, transparency, and accelerated discovery.

References

Abdi, H. (2010). Partial Least Squares Regression and Projection on Latent Structure Regression (PLS-Regression). *Wiley Interdisciplinary Reviews: Computational Statistics*, 2, 97-106.

Aldowah, H., Rehman, S. U., Ghazal, S., & Umar, I. N. (2017). Internet of Things in Higher Education: A Study on Future Learning. In *Journal of Physics: Conference Series* (Vol. 892, No. 1, p. 012017). IOP Publishing.

Al-Emran, M., Malik, S. I., & Al-Kabi, M. N. (2020). A Survey of Internet of Things (IoT) in Education: Opportunities and Challenges. In *Toward Social Internet of Things (SIoT): Enabling Technologies, Architectures and Applications: Emerging Technologies for Connected and Smart Social Objects* (pp. 197-209).

Anderson, T. R., Daim, T. U., & Kim, J. (2008). Technology Forecasting for Wireless Communication. *Technovation*, 28(9), 602-614

Anderson, T., Färe, R., Grosskopf, S., Inman, L., & Song, X. (2002). Further Examination of Moore's Law with Data Envelopment Analysis. *Technological Forecasting and Social Change*, 69(5), 465-477.

Anderson, T. R., Hollingsworth, K., & Inman, L. (2001). Assessing the Rate of Change in the Enterprise Database System Market Over Time Using DEA. In *Proceedings of PICMET2001* (Vol. 1, pp. 203).

Arthur, W. B. (1996). Increasing Returns and the New World of Business. *Harvard Business Review*, 74(4), 100-109.

ASEM (n.d.-a). ASEM - American Society for Engineering Management. Retrieved November 30, 2023, from https://asem.org/.

ASEM (n.d.-b). EMBOK - A Guide to the Engineering Management Body of Knowledge, 5th Edition. Retrieved November 30, 2023, from https://asem.org/EMBoK.

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical Framework for Artificial Intelligence and Digital Technologies. *International Journal of Information Management*, 62, 102433.

Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 359-388.

Bauer, B., & Patrick, A. S. (2004). A Human Factors Extension to the Seven-Layer OSI Reference Model. Retrieved December 2, 2023, from http://www.andrewpatrick.ca/OSI/10layer.html.

Bech, M. (2011). Calculating the Return on Investment of an NREN. In Papers Presented at Trans-European Research and Education Network Association Networking Conference 2011.

Bera, S., Misra, S., & Vasilakos, A. V. (2017). Software-Defined Networking for Internet of Things: A Survey. *IEEE Internet of Things Journal*, 4(6), 1994-2008.

Bogetoft, P., & Otto, L. (2011). Benchmarking with DEA, SFA, and R. *International Series in Operations Research & Management Science*, 157. Springer.

Bollen, K. A., & Curran, P. J. (2004). Autoregressive Latent Trajectory (ALT) Models: A Synthesis of Two Traditions. *Sociological Methods & Research*, 32(3), 336-383.

Boulesteix, A. L., & Strimmer, K. (2007). Partial Least Squares: A Versatile Tool for the Analysis of High-Dimensional Genomic Data. *Briefings in Bioinformatics*, 8(1), 32-44.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). Time Series Analysis: Forecasting and Control. John Wiley & Sons.

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*, 3(2), 77-101.

Buchroithner, M. (1998). Geodata Interrelations: Inventory and Structuring Attempt of Taxonomic Diversity. In *Proceedings of the 2nd Conference on Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images* (pp. 11–15).

Burant, C. J. (2022). A Methodological Note: An Introduction to Autoregressive Models. *The International Journal of Aging and Human Development*, 95(4), 516-522.

Castanedo, F. (2013). A Review of Data Fusion Techniques. *The Scientific World Journal*, 2013.

Cepeda-Carrion, G., Cegarra-Navarro, J. G., & Cillo, V. (2019). Tips to Use Partial Least Squares Structural Equation Modelling (PLS-SEM) in Knowledge Management. *Journal of Knowledge Management*, 23(1), 67-89.

Charmaz, K. (2006). Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis. Sage.

Chang, H. (2007). A Methodology for the Identification of Technology Indicators. *HNI-Verlagsschriftenreihe*, 233.

Chang, H., Gausemeier, J., Ihmels, S., & Wenzelmann, C. (2008). Innovative Technology Management System with Bibliometrics in the Context of Technology Intelligence. In *Trends in Intelligent Systems and Computer Engineering: Lecture Notes in Electrical Engineering*, 6, 349-361.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the Efficiency of Decision-Making Units. *European Journal of Operational Research*, 2(6),429-444.

Chatterjee, S., & Bhattacharjee, K. K. (2020). Adoption of Artificial Intelligence in Higher Education: A Quantitative Analysis Using Structural Equation Modelling. *Education and Information Technologies*, 25, 3443-3463.

Chin, W. W. (1998a). The Partial Least Squares Approach to Structural Equation Modeling. In G. A. Marcoulides (Ed.), *Modern Methods for Business Research* (pp. 295–336). Mahwah, NJ: Lawrence Erlbaum.

Chin, W. W. (1998b). Issues and Opinion on Structural Equation Modeling. *Management Information Systems Quarterly*, 22(1), 7–16.

Chin, W. W., & Newsted, P. R. (1999). Structural Equation Modelling Analysis with Small Samples Using Partial Least Squares. In R. Hoyle (Ed.), *Statistical Strategies for Small Sample Research* (pp.307-341). London: Sage Publications.

Cho, Y. (2013). Investigating the Merge of Exploratory and Normative Technology Forecasting Methods. In *Proceedings of PICMET2013*.

Cho, Y., & Daim, T. (2013). Technology Forecasting Methods. In *Research and Technology Management in the Electricity Industry: Methods, Tools and Case Studies* (pp. 67-112). London: Springer London.

Chopra, G., Madan, P., Jaisingh, P., & Bhaskar, P. (2019). Effectiveness of E-Learning Portal from Students' Perspective: A Structural Equation Model (SEM) Approach. *Interactive Technology and Smart Education*, 16(2), 94-116.

Coates, V., Farooque, M., Klavans, R., Lapid, K., Linstone, H. A., Pistorius, C., & Porter, A. L. (2001). On the Future of Technological Forecasting. *Technological Forecasting and Social Change*, 67(1), 1-17.

Cooper, R., Hutton, J. S., & Smith, I. L. (1991). From JANET to SuperJANET. *Computer Networks and ISDN Systems*, 21(4), 347-351.

Cover, T. M. (1999). Elements of Information Theory. John Wiley & Sons.

Crawford, K. (2021). The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence. Yale University Press.

Dalkey, N. C., Brown, B. B., & Cochran, S. (1969). The Delphi Method: An Experimental Study of Group Opinion (Vol. 3, p. 107). Santa Monica, CA: Rand Corporation.

Danks, N. P., Sharma, P. N., & Sarstedt, M. (2020). Model Selection Uncertainty and Multi-model Inference in Partial Least Squares Structural Equation Modeling (PLS-SEM). *Journal of Business Research*, 113, 13-24.

Das, K. (2019). The Role and Impact of ICT in Improving the Quality of Education: An Overview. *International Journal of Innovative Studies in Sociology and Humanities*, 4(6), 97-103.

Dash, G., & Paul, J. (2021). CB-SEM vs PLS-SEM Methods for Research in Social Sciences and Technology Forecasting. *Technological Forecasting and Social Change*, 121092.

Dator, J. (2009). Alternative Futures at the Manoa School. *Journal of Futures Studies*, 14(2), 1-18.

Davis, P. K., & Bigelow, J. H. (1998). Experiments in Multiresolution Modeling (MRM). Santa Monica, CA: RAND Corporation.

De Paola, A., Ferraro, P., Gaglio, S., & Lo Re, G. (2016). Context-Awareness for Multi-Sensor Data Fusion in Smart Environments. In *AI*IA 2016 Advances in Artificial Intelligence: XVth International Conference of the Italian Association for Artificial Intelligence, Genova, Italy, November 29–December 1, 2016, Proceedings XV* (pp. 377-391). Springer International Publishing.

Deering, S., & Hinden, R. (1998). Internet Protocol, Version 6 (IPv6) Specification. RFC 2460. Defense Technical Information Center. (n.d.). Revisions to the JDL Data Fusion Model. Retrieved from https://apps.dtic.mil/sti/pdfs/ADA391479.pdf.

Diez-Olivan, A., Del Ser, J., Galar, D., & Sierra, B. (2019). Data Fusion and Machine Learning for Industrial Prognosis: Trends and Perspectives Towards Industry 4.0. *Information Fusion*, 50, 92-111.

Durmuşoğlu, A., & Dereli, T. (2011). On the Technology Forecasting Using Data Envelopment Analysis (TFDEA). In *Proceedings of PICMET2011: Technology Management in the Energy Smart World* (pp. 1-6). Portland, OR, USA.

Evermann, J., & Tate, M. (2010). Testing Models or Fitting Models? Identifying Model Mis-specification in PLS. In *Papers presented at the International Conference on Information Systems 2010*.

Feng, L., Wang, Q., Wang, J., & Lin, K. Y. (2022). A Review of Technological Forecasting from the Perspective of Complex Systems. *Entropy*, 24(6), 787.

Fey, V., & Rivin, E. I. (2005). Innovation on Demand: New Product Development Using TRIZ. Cambridge, UK: Cambridge University Press.

Figshare. (n.d.). Home page. Retrieved January 17, 2024, from https://figshare.com

Fisher, J. C., & Pry, R. H. (1971). A Simple Substitution Model of Technological Change. *Technological Forecasting and Social Change*, 3(1), 75-88.

Foley, M. (2016). The Role and Status of National Research and Education Networks in Africa. SABER-ICT Technical Paper Series. World Bank. Retrieved December 10, 2023, from http://hdl.handle.net/10986/26258.

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Errors. *Journal of Marketing Research*, 18, 39–50.

Gagolewski, M. (2022). Data Fusion: Theory, Methods, and Applications. arXiv preprint arXiv:2208.01644.

GÉANT. (2022). 2022 GÉANT Compendium of National Research and Education Networks in Europe. Retrieved December 10, 2023, from https://resources.geant.org/geant-compendia/.

George, D., & Mallery, P. (2019). IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference (16th ed.). New York, NY: Routledge.

Georghiou, L., Cassingena Harper, J., Keenan, M., Miles, I., & Popper, R. (2008). The Handbook of Technology Foresight: Concepts and Practice. Cheltenham, UK: Edward Elgar Publishing.

Glasser, B. (1965). The Constant Comparative Method of Qualitative Analysis. *Social Problems*, 12(4), California: University of California Press.

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, 8(4), 597-607.

Goodhue, D., Lewis, W., & Thompson, R. (2006). PLS, Small Sample Size, and Statistical Power in MIS Research. In *Proceedings of the 39th Hawaii International Conference on System Sciences*.

Gordon, T. J., & Helmer, O. (1964). Report on a Long-Range Forecasting Study. Santa Monica, CA: RAND Corporation.

Greaves, D. (2009). An NREN Capability Maturity Model. Retrieved December 2, 2023, from https://casefornrens.org/wp-content/uploads/2021/05/NREN-Capability-Maturity-Model-CMM.pdf.

Green, K. C., & Armstrong, J. S. (2007). Structured Analogies for Forecasting. *International Journal of Forecasting*, 23(3), 365-376.

Grupp, H. (1998). Foundations of the Economics of Innovation: Theory, Measurement and Practice. Cheltenham: Edward Elgar.

Gustavsen, B. (2008). Action Research, Practical Challenges and the Formation of Theory. *Action Research*, 6(4), 421-437.

Haenlein, M., & Kaplan, A. M. (2004). A Beginner's Guide to Partial Least Squares Analysis. *Understanding Statistics*, 3(4), 283-297.

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (2nd ed.). Sage Publications.

Hall, D. L., & Llinas, J. (1997). An Introduction to Multisensor Data Fusion. *Proceedings of the IEEE*, 85(1), 6-23.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. (2015). A Critique of the Cross-Lagged Panel Model. *Psychological Methods*, 20(1), 102-116.

Henseler, J., Ringle, C. M., & Sarstedt, M. (2015). A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling. *Journal of the Academy of Marketing Science*, 43(1), 115-135.

Howard, R. A., & Matheson, J. E. (2005). Influence Diagrams. *Decision Analysis*, 2(3), 127-143.

Inman, L. (2004). Technology Forecasting Using Data Envelopment Analysis. Department of Engineering and Technology Management, Portland State University.

Inman, L., Anderson, T. R., & Harmon, R. (2005). Improving Time to Market Forecasts: A Comparison of Two Technology Forecasting Techniques for Predicting U.S. Fighter Jet Introductions from 1944 to 1982. In *Technology Management: A Unifying Discipline for Melting the Boundaries* (pp. 533-537). Portland, OR, USA.

Janz, R., & Kutanov, A. (2012). The Case of NRENs in Central Asia. *Central Asian Journal of Global Health*, 1(1).

Johnson, L., Smith, R., Willis, H., Levine, A., & Haywood, K. (2011). 2011 Horizon Report. New Media Consortium.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed Methods Research: A Research Paradigm Whose Time Has Come. *Educational Researcher*, 33(7), 14-26.

Jolliffe, I. T. (2002). Principal Component Analysis for Special Types of Data (pp. 338-372). Springer, New York.

Jöreskog, K. G. (1973). A General Method for Estimating a Linear Structural Equation System. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences* (pp. 85-112). New York, NY: Seminar.

Jöreskog, K. (1975). Structural Analysis of Covariance and Correlation Matrices. *Psychometrika*, 40(4), 443-477.

Kahn, H. (1960). On Thermonuclear War. Princeton, NJ: Princeton University Press.

Kahn, H., & Aron, R. (1962). Thinking about the Unthinkable (Vol. 254). New York: Horizon Press.

Kalem, G., Vayvay, O., Sennaroglu, B., & Tozan, H. (2021). Technology Forecasting in the Mobile Telecommunication Industry: A Case Study Towards the 5G Era. *Engineering Management Journal*, 33(1), 15-29.

Kao, C., & Hwang, S. N. (2008). Efficiency Decomposition in Two-Stage Data Envelopment Analysis: An Application to Non-Life Insurance Companies in Taiwan. *European Journal of Operational Research*, 185(1), 418-429.

Kashefi, A., Taylor, S. J. E., Abbott, P. Y., Anagnostou, A., Tessa, O. M., Oaiya, O., Barry, B., & Alline, D. (2019). User Requirements for National Research and Education Networks for Research in West and Central Africa. *Journal of Network and Computer Applications*, 142, 1-10.

Khaleghi, B., Khamis, A., & Nahavandi, S. (2013). Multisensor Data Fusion: A Review of the State-of-the-Art. *Information Fusion*, 17, 28-44.

Koch, W. (2021). On Digital Ethics for Artificial Intelligence and Information Fusion in the Defense Domain. *IEEE Aerospace and Electronic Systems Magazine*, 36(7), 94-111.

Kompella, K., Swallow, G., Pignataro, C., Kumar, N., Aldrin, S., & Chen, M. (2017). Detecting Multiprotocol Label Switched (MPLS) Data-Plane Failures (No. rfc8029).

Kucharavy, D., & De Guio, R. (2005). Problems of Forecast. In *Proceedings of the ETRIA TRIZ Future 2005 Conference* (pp. 219-235).

Lahat, D., Adali, T., & Jutten, C. (2015). Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects. *Proceedings of the IEEE*, 103(9), 1449-1477.

Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., & Wolff, S. (2009). A Brief History of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22-31.

Li, J., Hong, D., Gao, L., Yao, J., Zheng, K., Zhang, B., & Chanussot, J. (2022). Deep Learning in Multimodal Remote Sensing Data Fusion: A Comprehensive Review. *International Journal of Applied Earth Observation and Geoinformation*, 112, 102926.

Lichtenthaler, E. (2004). Technological Change and the Technology Intelligence Process: A Case Study. *Journal of Engineering and Technology Management*, 21, 331-348.

Litwin, M. S. (2017). How to Measure Survey Reliability and Validity. Sage Publications.

Mack, C. A. (2011). Fifty Years of Moore's Law. *IEEE Transactions on Semiconductor Manufacturing*, 24(2), 202-207.

Manyika, J., & Durrant-Whyte, H. (1995). Data Fusion and Sensor Management: A Decentralized Information-Theoretic Approach. Prentice Hall PTR.

Martino, J. P. (1983). Technological Forecasting for Decision Making. Amsterdam, The Netherlands: North-Holland Publishing.

Mbarika, V. W., Byrd, T. A., Raymond, J. E., & McMullen, P. (2000). Investments in Telecommunications Infrastructure Are Not the Panacea for Least Developed Countries Leapfrogging Growth of Teledensity. *International Journal on Media Management*, 2(3/4), 133–142.

McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., & Turner, J. (2008). OpenFlow: Enabling Innovation in Campus Networks. *ACM SIGCOMM Computer Communication Review*, 38(2), 69-74.

McNeish, D., & Hamaker, E. L. (2020). A Primer on Two-Level Dynamic Structural Equation Models for Intensive Longitudinal Data in Mplus. *Psychological Methods*, 25(5), 610.

Melhem, S., Kelly, T., Park, L. M., Hurpy, C. P. M., & Bashir, S. (2021). Feasibility Study to Connect all African Higher Education Institutions to High-Speed Internet. World Bank Group. Retrieved December 10, 2023, from http://documents.worldbank.org/curated/en/320071626434508906/Feasibility-Study-to-Connect-all-African-Higher-Education-Institutions-to-High-speed-Internet.

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A Survey on Machine Learning for Data Fusion. *Information Fusion*, 57, 115-129.

Metcalfe, B. (1995). Metcalfe's Law: A Network Becomes More Valuable as it Reaches More Users. *Infoworld*, 17(40), 53-54.

Miles, M. B., & Huberman, A. M. (1994). Qualitative Data Analysis: An Expanded Sourcebook. Sage Publications.

Mingers, J., & Brocklesby, J. (1997). Multimethodology: Towards a Framework for Mixing Methodologies. *Omega*, 25(5), 489-509.

Moore, G. E. (1965). Cramming More Components onto Integrated Circuits. *Electronics Magazine*, 38(8).

National Research Council. (2010). Persistent Forecasting of Disruptive Technologies. Washington, DC: The National Academies Press.

National Science Foundation. (2020). Research on Emerging Technologies for Teaching and Learning (RETTL). Retrieved December 2, 2023, from https://www.nsf.gov/pubs/2020/nsf20612/nsf20612.htm.

Nielsen, J. (1988, April). Nielsen's Law of Internet Bandwidth. Retrieved January 8, 2013, from http://www.nngroup.com/articles/nielsens-law-of-internet-bandwidth/.

Nokia. (n.d.). Research and Education Networks. Retrieved December 2, 2023, from https://www.nokia.com/industries/research-and-education-networks/.

Nyberg, A., & Palmgren, S. (2011). Using Indicators for Technology Monitoring: Steps Toward a Proposed Framework. (Master of Science Thesis, Chalmers University of Technology).

Pateli, A. G. (2009). Decision Making on Governance of Strategic Technology Alliances. *Management Decision*, 47(2), 246-270.

Pillay, K., Erasmus, L., & Pretorius, J. H. C. (2021). A Systems Thinking Approach to Value-Added Services Adoption in National Research and Education Networks. In *2021 IEEE AFRICON* (pp. 1-6). IEEE.

Porter, A. (2005). QTIP: Quick Technology Intelligence Processes. *Technological Forecasting and Social Change*, 72, 1070–1081.

Porter, A. L., & Cunningham, S. W. (2004). Tech Mining: Exploiting New Technologies for Competitive Advantage. Wiley-Interscience.

Priyadarsini, M., & Bera, P. (2021). Software Defined Networking Architecture, Traffic Management, Security, and Placement: A Survey. *Computer Networks*, 192, 108047.

Research Data Alliance. (n.d.). About the RDA. Retrieved January 17, 2024, from https://www.rd-alliance.org/about-rda

Ringle, C. M., Wende, S., & Becker, J.-M. (2022). SmartPLS 4. Oststeinbek: SmartPLS. Retrieved December 2, 2023, from https://www.smartpls.com.

Robinson, J. B. (1990). Futures under Glass: A Recipe for People Who Hate to Predict. *Futures*, 22(8), 820-842.

Rohrbeck, R., & Gemünden, H. G. (2011). Corporate Foresight: Its Three Roles in Enhancing the Innovation Capacity of a Firm. *Technological Forecasting and Social Change*, 78(2), 231-243.

Rosen, E., Viswanathan, A., & Callon, R. (2001). Multiprotocol Label Switching Architecture. RFC 3031.

Rosipal, R., & Krämer, N. (2006). Overview and Recent Advances in Partial Least Squares. In *Subspace, Latent Structure and Feature Selection Techniques* (pp. 34-51). Springer, Berlin, Heidelberg.

Rowe, G., & Wright, G. (1999). The Delphi Technique as a Forecasting Tool: Issues and Analysis. *International Journal of Forecasting*, 15(4), 353-375.

Ryan, G., & Bernard, H. R. (2003). Techniques to Identify Themes. *Field Methods*, 15(1), 85–109.

SANReN. (n.d.). SANReN - The South African National Research Network. Retrieved November 30, 2023, from http://www.sanren.ac.za/.

Saritas, O., & Aylen, J. (2010). Using Scenarios for Roadmapping: The Case of Clean Production. *Technological Forecasting and Social Change*, 77(7), 1061-1075.

Savory, B. (2012). Build vs. Buy: The NREN Network Connectivity Dilemma: About Building Research & Education Fibre Networks. In *Proceedings and Report of the 5th UbuntuNet Alliance Annual Conference* (pp. 160-173).

Selig, J. P., & Preacher, K. J. (2009). Mediation Models for Longitudinal Data in Developmental Research. *Research in Human Development*, 6(2-3), 144-164.

Shah, H., & Nowocin, W. (2015). Yesterday, Today and Future of the Engineering Management Body of Knowledge. *Frontiers of Engineering Management*, 2(1), 60-63.

Smith, J. A. (2015). Qualitative Psychology: A Practical Guide to Research Methods. Sage.

Sohn, S. Y., & Moon, T. H. (2003). Structural Equation Model for Predicting Technology Commercialization Success Index (TCSI). *Technological Forecasting and Social Change*, 70(9), 885-899.

Stahlschmidt, S. R., Ulfenborg, B., & Synnergren, J. (2022). Multimodal Deep Learning for Biomedical Data Fusion: A Review. *Briefings in Bioinformatics*, 23(2).

Staphorst, L. (2010). The Impact of Intellectual Property Rights from Publicly Financed Research and Development on Governance Mode Decisions for Research Alliances. (Master of Business Administration Thesis, Gordon Institute of Business Science).

Staphorst, L., Holland, M., Pretorius, L., & Pretorius, M. W. (2015). Impact of Intellectual Property Rights from Publicly Financed Research and Development on Research Alliance Governance Mode Decisions. In *Proceedings of the International Association for Management of Technology (IAMOT) 2015 Conference*.

Staphorst, L., Pretorius, L., & Pretorius, T. (2013). Structural Equation Modelling Based Data Fusion for Technology Forecasting: A Generic Framework. In *Proceedings of PICMET2013* (pp. 2163-2170). San Jose, CA, USA.

Staphorst, L., Pretorius, L., & Pretorius, M. W. (2014). Structural Equation Modelling Based Data Fusion for Technology Forecasting: A National Research and Education Network Example. In *Proceedings of PICMET2014* (pp. 2908-2917). Kanazawa, Japan.

Staphorst, L., Pretorius, L., & Pretorius, M. W. (2016a). Technology Forecasting in the National Research and Education Network Technology Domain using Context Sensitive Data Fusion. *Technology Forecasting and Social Change*, 111, 110-123.

Staphorst, L., Pretorius, L., & Pretorius, M. W. (2016b). Technology Forecasting Using Structural Equation Modeling Based Data Fusion: Analysis of Strengths and Weaknesses Using a National Research and Education Network Example. In *Proceedings of PICMET2016*. Honolulu, Hawaii.

Staphorst, L., Pretorius, L., & Pretorius, M. W. (2017). Impact of Intellectual Property Rights on the Governance Mode Decisions of Engineering Managers during the Establishment of Research Alliances with Publicly Funded Entities. *Engineering Management Journal*, 29(1), 26-44.

Steinberg, A. N. (2009). Context-Sensitive Data Fusion using Structural Equation Modeling. In *Proceedings of the 12th IEEE International Conference on Information Fusion*.

Steinberg, A. N., & Bowman, C. L. (2017). Revisions to the JDL Data Fusion Model. In *Handbook of Multisensor Data Fusion* (pp. 65-88). CRC Press.

Steinberg, A. N., & Rogova, G. (2008). Situation and Context in Data Fusion and Natural Language Understanding. In *Proceedings of the 11th IEEE International Conference on Information*.

Stewart, T. R. (1987). The Delphi Technique and Judgmental Forecasting. *Climatic Change*, 11(1-2), 97-113.

TERENA. (2011). 2011 TERENA Compendium of National Research and Education Networks in Europe. Retrieved December 10, 2023, from https://resources.geant.org/geant-compendia/.

TERENA. (2012). 2012 TERENA Compendium of National Research and Education Networks in Europe. Retrieved December 10, 2023, from https://resources.geant.org/geant-compendia/.

Twinomugisha, A. (2006). Understanding NRENs and Key Considerations for Setting Them Up. Retrieved December 10, 2023, from https://casefornrens.org/wp-content/uploads/2021/05/Understanding-NRENs-and-key-considerations-for-setting-them-up.pdf.

Vanston, J. H. (2003). Better Forecasts, Better Plans, Better Results. *Research-Technology Management*, 46(1), 47-58.

Vanston, J. H., & Vanston, L. K. (2004). Testing the Tea Leaves: Evaluating the Validity of Forecasts. *Research-Technology Management*, 47(5), 33-39.

Vinzi, V. E., Chin, W. W., Henseler, J., & Wang, H. (2010). Handbook of Partial Least Squares: Concepts, Methods and Applications. Berlin, Heidelberg: Springer.

von Karman, T. (1945). Toward New Horizons. Washington, DC: The National Academies Press.

Wald, L. (1997). An Overview of Concepts in Fusion of Earth Data. In *Proceedings of the EARSeL Symposium 1997, Future Trends in Remote Sensing* (pp. 385–390).

Wald, L. (1999). Some Terms of Reference in Data Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3).

Waltz, E. L., & Llinas, J. (1990). Multisensor Data Fusion. Artech House.

Wang, C., Liu, X., & Xu, D. (1999). Chaos Theory in Technology Forecasting. In *Proceedings of PICMET1999*.

Watts, R. J., & Porter, A. L. (1997). Innovation Forecasting. *Technological Forecasting and Social Change*, 56(1), 25-47.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The FAIR Guiding Principles for Scientific Data Management and Stewardship. Scientific Data, 3, 160018.

Wold, H. (1966). Estimation of Principal Components and Related Models by Iterative Least Squares. In *Multivariate Analysis* (pp. 391-420). Academic Press.

Wold, H. (1975). Path Models with Latent Variables: The NIPALS Approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling* (pp. 307–357). New York, NY: Academic Press.

Wold, S., & Sjöström, M. (1977). SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. *Chemometrics: Theory and Application*, 52, 243-282.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-Regression: A Basic Tool of Chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109-130.

Wolfers, J., & Zitzewitz, E. (2004). Prediction Markets. *Journal of Economic Perspectives*, 18(2), 107-126.

Yaver, A., Sánchez-Torres, J. M., Amórtegui, M. A., & Giraldo-Ríos, L. (2016). A Participatory Foresight for National Research and Education Networks. *European Journal of Futures Research, 4*, 1-11.

Zikmund, W. (2003). Business Research Methods (7th ed.). Ohio: South-Western Cengage Learning.

Zimmerman, H. (1980). OSI Reference Model - The ISO Model of Architecture for Open Systems Interconnection. *IEEE Transactions on Communications*, 28(4), 425-432.