Structural Equation Modelling Based Data Fusion for Technology Forecasting: A Generic Framework

Leon Staphorst[1], Leon Pretorius2, Tinus Pretorius2

[1]Council for Scientific and Industrial Research, Meiring Naude Road, Brummeria, Pretoria 0001, South Africa
[2]Graduate School of Technology Management, University of Pretoria,
Lynnwood Road, Pretoria, SA-0002, South Africa

Abstract - Technology Intelligence (TI) involves the process of capturing technology related data, converting this data into information by determining relational connections and refining information to produce knowledge that can guide strategic decision makers. Technology indicators are those sources of technology related data that allow for the direct characterisation and evaluation of technologies over their whole life cycle. Future-oriented Technology Analysis (FTA), which is a forward-looking approach in scrutinizing the information that has been distilled from a set of technology indicators, can potentially provide decision makers with useful Technology Forecasting (TF) knowledge. The paper postulates that TF can be viewed as an instance of Data Fusion (DF), which is a formal framework that defines tool, as well as the application of these tools, for the unification of data originating from different sources. Within the field of DF relational connections define context. Context sensitive DF techniques refine the generated knowledge based on the characteristics of exogenous context related variables. Structural Equation Modelling (SEM), which is a statistical technique capable of evaluating complex hierarchical dependencies between latent and observed problem and context variables, has been shown to be effective in implementing context sensitive DF. In the paper a generic framework is introduced for SEM based DF of technology indicators in order to produce TF output metrics. The paper also provides the research methodology that will be used in a future study to evaluate the validity of the generic framework for the case of National Research and Education Networks (NRENs).

Keywords - Technology Intelligence, Technology Indicators, Technology Forecasting, Data Fusion, Structural Equation Modelling

I. Introduction and Research Method

Technological advancement continues at an astounding rate, surprisingly following exponential growth models such as Moore's [1], Nielsen's Law [2] and Metcalfe's Law [3]. Driven not only by the invention, innovation and diffusion of new technologies, but also by the move to the inclusive mindsets of globalisation and open innovation [4], this has created highly competitive global markets for technology based products and services [5]. Hence, the survival, growth and profitability of firms that play in these markets depend highly on their ability to monitor current, as well as predict future technological changes in order to create a solid and sustainable technological base that can withstand, or adapt to rapidly changing market requirements [5]. Moreover, firms need to effectively and efficiently manage technological changes both internally and externally if they are to create sustainable competitive advantages in rapidly high-tech markets [6]. Technology Intelligence (TI), which is a core process within the discipline of technology management, involves the process of capturing technology related data, converting this data into information by determining relational connections and refining information to produce knowledge that can guide strategic decision makers during strategic planning [6][7]. Technology indicators, such as technology maturity and degree of innovation, are those measureable sources of technology related data that allow for the direct characterisation and evaluation of technologies over their whole life cycle [7]. Scrutinizing the information that has been distilled from a set of technology indicators in a forward-looking approach, commonly referred to as Future-oriented Technology Analysis (FTA), can potentially provide decision makers with Technology Forecasting (TF) knowledge, amongst others [8].

Buchroithner [9] and Wald [10] define Data Fusion (DF), which was developed in the military domain for the generation of quality tactical knowledge through the multi-layered processing of sensor data [11], as "… a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application." Within the discipline of DF, context can be viewed as synonymous with a situation, which in turn is defined as a set of relational connections (i.e. an instantiated relation) [12]. Context can be used in each level of the DF process in order to refine data alignment and association, as well as during situation state estimation [12]. Recently, context sensitive DF techniques have been explored which effectively refine the generated knowledge at each level of processing based on the characteristics of exogenous context-related variables [12].

Regression analysis constitutes a family of statistical techniques geared at modelling and analysing the relationship between dependent and independent variables from empirical data [13]. Moreover, regression analysis attempts to explain the variations in independent variables as functions (commonly referred to regression functions) of variations in dependent variables [13]. With this knowledge it is then possible to perform prediction and forecasting of the values that dependent variable will assume for specific independent variable values [13]. Classic regression techniques (such as multiple regression, discriminant analysis, logistic regression and analysis of variance) can be classified as first generation techniques, since these techniques explicitly assume independence between multiple dependent variables [13]. This, unfortunately, limits the ability of such techniques to comprehensively model complex interrelationships, such as the interplay between two or more output variables in a TF model. More specifically, classic first generation regression techniques are not able to model the potential mediating or moderating effect that output variables could have on one another. To overcome this limitation, Jöreskog [14] proposed covariance based Structural Equation Modelling (SEM) as a second generation technique, which allows for the simultaneous modelling of relationships among multiple dependent and independent constructs. A further inherent limitation of first generation regression techniques is their explicit assumption that all dependent and independent variables are directly observable [13]. This assumption implies that all variables' values can be directly obtained from real-world sampling experiments [13]. As such, any variables that cannot be directly observed need to be considered unobservable and have to be excluded from first generation regression models [13]. However, such unobservable variables, commonly referred to as latent constructs, are supported by SEM. Steinberg postulated that SEM is ideally suited to implement context sensitive DF [12][15]. Not only does SEM support the complex structural models used in situation state estimation (as is required in TF), it also allows for non-linear and non-Gaussian factors and cyclical dependencies amongst model variables that can be either latent or directly observable [12].

According to Sohn and Moon [16] most TF techniques rarely take into account the structural relationships amongst technology indicators and TF output metrics. SEM, however, provides an advantage over these limited TF techniques by allowing for the modelling of complex hierarchical relationships between technology indicators and TF outputs metrics. Sohn and Moon [16] have shown that SEM, which can be viewed as a generalization of factor and path analysis methods such as Bayesian Networks [12], can successfully implementing TF of the Technology Commercialization Success Index (TCSI) TF output metric. This paper builds on the work of Steinberg [12][15], as well as Sohn and Moon [16], by postulating a generic framework for the SEM based DF of technology indicators in order to produce TF output metrics.

The paper is structured as follows: Firstly, consideration is given to the basics of SEM and how SEM can be used to perform context sensitive DF. This is followed by an overview of technology indicators and forecasting output metrics, as well as their use in the generic framework for SEM based DF for TF that is proposed in the paper. Next, the generic framework contributed by the paper is described. Lastly the paper presents the research methodology that will be used in the future evaluation of the framework for TF for National Research and Education Networks (NRENs). This includes a discussion on the approach that will be used to test the validity and reliability of the generic framework. Note that the research method at this stage is literature based and exploratory in nature using inductive reasoning, where appropriate.

II. SEM for Context Sensitive DF

Within SEM theory distinction is made between exogenous and endogenous latent constructs, with the former being variables that are not explained by the internal interrelationships embodied by the model and therefore always act as independent variables [13]. Due to its generality, SEM terminology does not refer to regression analysis' dependent and independent variables, but rather only to exogenous constructs, which are independent variables that are not functions of any relationship in the model, as well as endogenous constructs, which are either dependent or independent variables that are explained by the relationships with other dependent and/or independent variables present in the model.

With reference to the indicators measured as proxies to represent latent constructs, such latent constructs can be further classified as follows [13]: A latent construct with reflective indicators is one in which all measured indicator proxies, also commonly referred to as factors, are expected to have high correlations to the latent construct, as well as other potential reflective indicators, Therefore it will have the ability to represent the variance in the unobserved variable sufficiently. In contrast, latent constructs with formative indicators are those that are represented by a weighted combination of indicators that are not highly correlated to either the latent construct itself, or the other formative indicators included in the weighted combination. The formative indicators

2

of a latent construct can therefore be seen as representing different dimensions of this construct. Fig. 1 shows a generic SEM path diagram, depicting all possible configurations of exogenous and endogenous constructs, the path coefficients of interconnections between these constructs, reflective and formative indicators, as well as the loadings of these indicators on constructs [17].
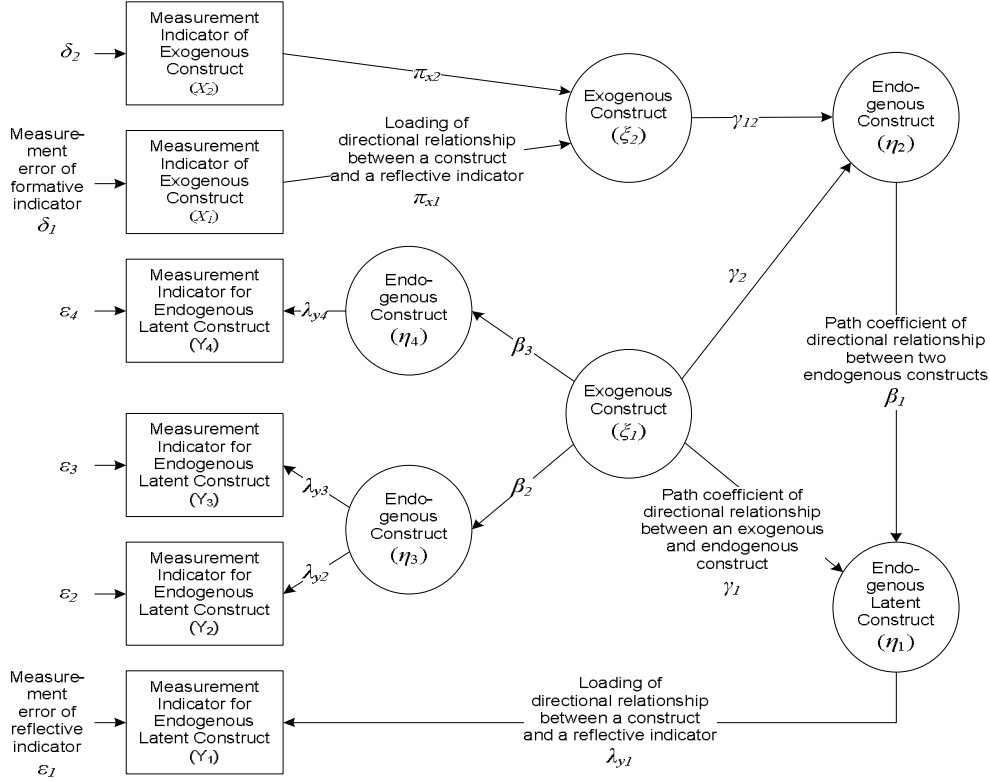

Figure 1: Generic SEM Path Diagram [17]

With reference to Fig. 1, the following popular symbol conventions apply in SEM [17]:

- ξ_n = The nth exogenous construct.
- η_m = The mth endogenous construct.
- X_i = The ith measurement indicator for the nth exogenous latent construct ξ_n.
- δ_i = The measurement error term associated with X_i. This term comprises of a random error part, as well as a systematic error part resulting from variance attributable to the measurement method itself, as opposed to the construct being measured.
- Y_j = The jth measurement indicator for the mth endogenous latent construct η_m.
- ε_j = The measurement error term associated with Y_j, also consisting of random and systematic error parts.
- λ_{xa} = The loading of a directional relation between the nth exogenous latent construct ξ_n and its ith reflective indicator X_i.
- λ_{yb} = The loading of a directional relation between the mth endogenous latent construct η_m and its jth reflective indicator Y_j.
- γ_c = The path coefficient of a directional relation between the mth endogenous latent construct η_m and the nth exogenous latent construct ξ_n.
- β_d = The path coefficient of a directional relation from the qth to the pth endogenous latent constructs, η_q and η_p.
- ζ_r = The rth disturbance term (or error term) in the rth endogenous construct η_r (not depicted in Fig. 1 due to space constraints). Hence, this term models the fact that the endogenous latent constructs are not perfectly explained by the independent variables.
- π_{xa} = The loading of a directional relation between the nth exogenous latent construct ξ_n and its ith formative indicator X_i.
- π_{yb} = The loading of a directional relation between the mth endogenous latent construct η_m and its jth formative indicator Y_j.

Using these symbol conventions it is now possible to create five sets of structural equations which fully represent the interrelationships embodied by a SEM model. Using matrix notation, the first set of equations relates exogenous latent constructs to their indicators and associated measurement errors [17]:

$$\mathbf{X} = \mathbf{\Lambda}_x \mathbf{\xi} + \mathbf{\delta}, \tag{1}$$

where the elements of matrices $\mathbf{X}$, $\mathbf{\Lambda}_x$, $\mathbf{\xi}$ and $\mathbf{\delta}$ are X_i, λ_{xa}, ξ_n and δ_i, respectively, for all applicable values of i, a and n [13].The second set of equations express endogenous latent constructs as functions of their reflective indicators and associated measurement errors [17]:

$$\mathbf{Y} = \mathbf{\Lambda}_y \mathbf{\eta} + \mathbf{\varepsilon}, \tag{2}$$

where the elements of matrices $\mathbf{Y}$, $\mathbf{\Lambda}_y$, $\mathbf{\eta}$ and $\mathbf{\varepsilon}$ are Y_j, λ_{yb}, η_m and ε_j, respectively, for all applicable values of j, b and m. The third set of equations considers the relationships between exogenous latent constructs and formative indicators, as well as measurement errors [17]:

$$\mathbf{\xi} = \mathbf{\Pi}_x \mathbf{X} + \mathbf{\delta}, \tag{3}$$

where the elements of matrices $\mathbf{\xi}$, $\mathbf{\Pi}_x$, $\mathbf{X}$ and $\mathbf{\delta}$ are ξ_n, π_{xa}, X_i and δ_i, respectively, for all applicable values of n, a and i. The fourth set of equations considers the relationships between endogenous latent constructs and formative indicators, as well as measurement errors [17]:

$$\mathbf{\eta} = \mathbf{\Pi}_y \mathbf{Y} + \mathbf{\varepsilon}, \tag{4}$$

where the elements of matrices $\mathbf{\eta}$, $\mathbf{\Pi}_y$, $\mathbf{Y}$ and $\mathbf{\varepsilon}$ are η_m, π_{yb}, Y_j and ε_j, respectively, for all applicable values of m, b and j. The last set of equations deals with the relationships between endogenous latent constructs and exogenous latent constructs, as well as the associated measurement errors [17]:

$$\mathbf{\eta} = \mathbf{B}\mathbf{\eta} + \mathbf{\Gamma}\mathbf{\xi} + \mathbf{\zeta}, \tag{5}$$

where the elements of matrices $\mathbf{\eta}$ ($\mathbf{\eta}$ is present on both sides of the equation since endogenous constructs can be dependent on one another), $\mathbf{B}$, $\mathbf{\Gamma}$, $\mathbf{\xi}$ and $\mathbf{\zeta}$ are η_m, β_d, γ_c, ξ_n and ζ_r respectively, for all applicable values of m, d, c, n and r [17].

Although Jöreskog in 1973 [14] originally proposed that the parameters of a SEM model be estimated using covariance based techniques, of which the LISREL program that was developed by Jöreskog in 1975 is arguably the most popular, variance based techniques, also commonly referred to as component based techniques, have also gained popularity [13]. Partial Least Squares (PLS), which was first introduced by Wold [18] as Non-linear Iterative Partial Least Squares (NIPALS), is one such variance based technique [13]. While covariance based techniques attempt to minimise the difference between the sample covariance values and those predicted by the regression model, which is equivalent to estimating the model parameters such that the covariance matrix of the observed measurements is reproduced, PLS regression, which is also sometimes referred to Projections to Latent Structures, focuses on maximising the variance of the dependent variables explained by the independent variables [13].

Recall that DF is essentially a framework for the multi-layered refinement of estimates of problem variables from multiple measurements, either directly or indirectly observable [12]. By noting that SEM is capable of the simultaneous modelling of relationships among multiple dependent and independent constructs, Steinberg [12] postulated that SEM is one potential statistical tool that lends itself naturally to implement DF. Moreover, based on the following argumentation Steinberg [12] showed that SEM allows for the inclusion of context sensitivity during the solving of DF inferencing problems: Firstly, Steinberg [12] defined a situation, or a context, as a set of relationships, where a relationship can be viewed as a specific instantiated relation. In general context is used in DF inferencing problems in order to refine ambiguous estimates, explain available data and constraint processing during data acquisition, cueing or fusion [12]. Next, Steinberg harmonized DF and SEM terminology by noting that DF problem variables are in fact SEM endogenous constructs, context variables can be viewed as SEM exogenous constructs and classic DF sensor measurements are the reflective and formative indicators present in SEM.

III. Technology Indicators and Forecasting Output Metrics

According to Porter and Cunningham [19] technology indicators employ empirical information to estimate technology characteristics that affect technological advance and successful commercialization. Watts and Porter [20] state that technology indicators are empirical measures stemming from general models of technological innovation and progression, such as the S-curve. Nyberg and Plamgren [4] expands on these definitions by describing technological indicators as those indices or statistical data that allow for the direct characterisation of characteristics of technology throughout their life cycles in order to allow decision makers to take strategic actions. Such indicators can in general be divided into three major categories based on their intended function: input indicators, byput indicators and output indicators [4][21]. Grupp [21] states that input indicators are variables related to drivers of technological progress, byput indicators are variables that are related to sub-phenomena of the technological progress and output indicators are variables related to the qualitative, quantitative or value-rated progress in process or product development [4]. A wide variety of sources exist that can be used to harvest technology indicators, ranging from patent databases and scientific publications [19], through to the rumour mill and financial market indicators [4]. In monitoring and mining these potential sources of technical indicators, Bibliometrics have emerged as one of the most popular set of quantitative techniques [4]. Bibliometrics uses counts of citations, publications or patents to produce indicators of technological progress in a specific domain [4].

Various frameworks have been proposed for the systematic categorisation of technology indicators. In [4] Nyberg and Palmgren presents a succinct summary of the frameworks proposed by Watts and Porter [20], Grupp [21] and Chang [22], which is repeated here: The Watts and Porter [20] framework consists of the following three categories:

- *Technology Life Cycle Status Indicators*: Based primarily on the S-curve, these metrics determine the level of progress of a technological development along its life cycle, as well as the growth rate of the technology [4].
- *Innovation Context Receptivity Indicators*: These indicators gauge the sufficiency of supporting technology, as well as the development of standards and regulations surrounding the technology under investigation [4].
- *Market Prospect Indicators*: The potential commercial payoffs of the technology are considered by this type of indicator. Of specific importance with these indicators are factors such as technology application areas, intellectual property and market competitiveness [4].

Although Grupp was the instigator of the general function based classification of technology indicators into input, byput and output indicators [21], he originally referred to these three types of indicators based on the stage in the technology's life cycle at which the measurement was performed:

- *Resource Indicators*: This input indicator type measures the various possible expenditures on research, development and innovation [4].
- *Research and Development (R&D) Results Indicators*: This is the output indicator type which measures qualitative, quantitative or value rated advances in production processes or products [4].
- *Progress Indicators*: Indicators of this type, for example the technometric indicator [21] that measures the number of features or product specifications, are byput metrics that measures sub-phenomena of the technological progress [4].

The Technology Indicator Ontology (TIO) proposed by Chang [22] divides technology indicators into the following two broad groupings, each with a number of sub-groups:

- *Technology Development Indicators*: This broad grouping includes measures that track the development, change, progress and trend of a technology from a technological perspective [4].
- *Market Development Indicators*: This broad grouping includes all indicators related to the market development and potential application areas of the technology, including sales, investment and industrial applications [4].

The proposed generic framework for SEM based DF for TF, as detailed in Section IV, allows for the use of any of the above stated types of technology indicators as latent or formative indicators for endogenous and exogenous constructs in the model. More specifically, input technology indicators will be used with exogenous constructs. Conversely, bypass and output indicators will be used for endogenous constructs. The TF output

metrics, which will eventually be used by decision makers to drive strategic action, will consist of bypass and output metrics related to endogenous constructs in the SEM model. External environment related indicators contributing to exogenous constructs that realise context sensitivity in the DF process, could also include technology indicators. For example, Sohn and Moon [16] used the Technology Commercialization Success Index (TCSI) metric, which is an example of a market prospect indicator in the framework proposed by Watts and Porter [20], as the primary TF output metric for their SEM model.

IV. Generic Framework for SEM Based DF for TF

Sohn and Moon showed [16] that SEM can be used as an effective regression technique to evaluate a multi-layered hierarchal model through progressive aggregations and refinements of input technology indicator data in order to produce a reliable statistical estimate of the TCSI TF output metric. Similarly, the Joint Directors of Laboratories' Data Fusion Group (JDL/DFG) recognized that, in a military environment, DF entails the progressive aggregation and refinement of sensor data in order to produce quality tactical knowledge. In an attempt to standardise the structure of the multi-layered DF process across all possible military applications and implementations, the JDL/DFG defined the following six levels of processing [12]:

- *Level 0:* Signal/Feature/Subject Assessment
- *Level 1:* Object Assessment
- *Level 2:* Situation Assessment
- *Level 3:* Impact Assessment
- *Level 4:* Process Refinement
- *Level 5:* User Refinement

While these JDL/DFG DF level definitions might not be appropriate for the use of DF in TF, the concept of multi-layered progressive aggregation and refinement of measurement indicator data is core to the proposed generic framework, which is shown in Fig. 2. In this generic framework reference is made to generic DF Levels 0 through N, where N is user selected. Note also that the use of bi-directional interconnections between indicators and constructs, as well as between multiple constructs, is based on SEM path diagram conventions [17] and illustrates that the relations represented by these interconnections can be either reflective or formative, as described in Section II.
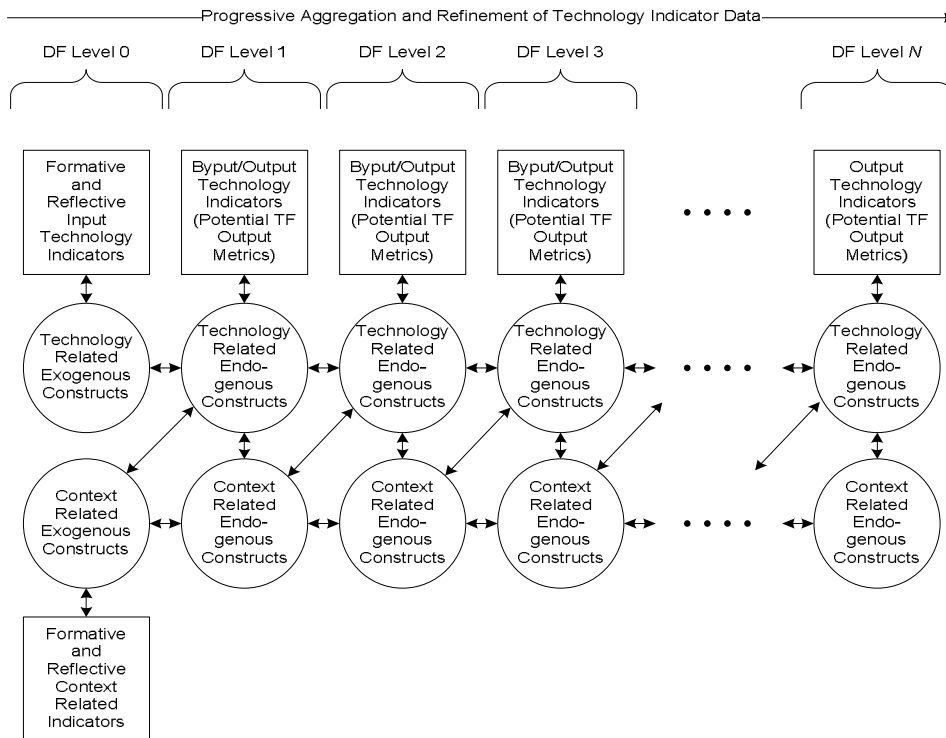
Figure 2: Proposed Generic Framework for SEM Based DF for TF

In this framework input technology indicators [4][21] and context related indicators [12] are used as inputs to technology and context related exogenous constructs, respectively. To gain an understanding of the multi-layered nature of this generic framework, consider the aggregation and refinement that occurs in progressing from DF Level 0 to DF Level 1: Regression analysis outputs generated for the technology related exogenous constructs at DF Level 0 contribute formatively of reflectively to technology related endogenous constructs at DF Level 1. Regression analysis outputs for the context related exogenous constructs of DF Level 0 contribute to context related and technology related endogenous constructs at DF Level 1. Note that the regression analysis results produced for DF Level 1 context related endogenous constructs can also contribute to technology related endogenous constructs at this same level. Technology indicators for the technology related constructs at DF Level 1 could potentially be selected as the TF output metrics, or could simply be byput technology [4][21] indicators if additional DF levels are required for further aggregation and refinement.

The aggregation and refinement achieved by moving from DF Level x to DF Level $x+1$, for $x = 1, 2, 3,…, N$-1, follows a similar interconnection structure as the progression from DF Level 0 to DF Level 1, with the exception that it is now endogenous constructs at DF Level x that contribute to endogenous constructs at DF Level $x+1$, not exogenous constructs. While classic DF based on the JDL/DFG model spans six levels of aggregation and refinement [12], the generic framework presented here allows for $N+1$ DF levels, where N would typically be user selected based on time and cost constraints, as well as potential diminishing returns resulting for additional levels of aggregation and refinement.

V. Future Evaluation of the Proposed Generic Framework

In order to evaluate the proposed generic framework a future study will be conducted to create and evaluate a SEM model instantiation for the case of the NREN technology domain. An NREN is defined as a specialised network service provider that exclusively supports a country's research and education communities [23]. Selection of this technology domain for the future study is motivated not only by the ease of access to information regarding relevant technology indicators and trends, but also by the fact that the domain itself is currently experiencing some rapid technology driven changes, resulting in evolving business models, innovative service offerings and increased international collaboration [23]. The future study will be conducted in two distinct phases: Phase 1 will be an exploratory study [24] and will attempt to construct a potential SEM model instantiation of the generic framework that can be used for TF in the NREN technology domain. Phase 2 will be confirmatory in nature [24] as it will attempt to determine the indicator loadings, path coefficients, validity and reliability of the SEM model determined during Phase 1.

Phase 1 will be performed qualitatively [24] and will attempt to identify applicable endogenous and exogenous model constructs, technology indicators and interactions between the various indicators and constructs in the SEM model being constructed. The unit of analysis [24] for this phase will be an NREN, while the population will be all NRENs in existence worldwide at the time of the study. Data collection will be accomplished through online surveys with open-ended questions as data collection instrument [24]. Respondents will be selected from the global community of NREN specialists through a snowball sampling approach [24]. Sufficiency of the sample size will be determined through the principle of data saturation [17]. Analysis of the collected qualitative data will firstly entail narrative inquiry by means of a process of theme extraction [17]. Thereafter frequency analysis will be performed on the extracted themes in order to produce a final set of importance ranked indicators, constructs and interconnections from which the SEM model will be constructed [17]. Testing the reliability and validity of the collected qualitative data will be accomplished by means of theory triangulation [17], as well as data triangulation [17] using as baseline published technology indicators from secondary data sources, such as TERENA's NREN Compendium [23].

Phase 2 will be performed quantitatively [24] and will, through PLS regression analysis, attempt to determine the applicable indicator loadings and path coefficients for the SEM model constructed during Phase 1. As with Phase 1 the population will be all NRENs in existence at that point in time, with the unit of analysis being a single NREN [24]. Quantitative online surveys, constructed using close-ended questions with Likert scaling, will be used as data collection instrument [24]. Senior managers at all of the NRENs in the population will be selected through a process of convenience sampling [24] as respondents for these surveys. In evaluating the reliability and validity of the regression analysis results produced by the SEM model, a popular approach will be followed the measurement portion of the model (also referred to as the outer model) is considered first, followed by the structural portion of the model (also referred to as the inner model) [25]. The logic behind this approach is that a lack in confidence in the accuracy and representivity of the measurement indicators in a SEM model

negates the need to continue testing the reliability and validity of the structural portion [25]. To evaluate the reliability and validity of the measurement portion (outer model) of the SEM model the following will be tested:

- *Indicator Reliability:* For a reflective indicator (denoted as X_i and Y_j for the indicators of exogenous and endogenous latent constructs, respectively) this reliability measure gives an indication of the level of variance in the measurement indicator that can be explained by its associated latent construct [25].
- *Construct Reliability:* The Indicator Reliability metric described above is designed to point to a given reflective indicator's inadequate measurement of a latent construct [25]. However, it is important to also consider whether the set of reflective indicators associated with a latent construct jointly measure it adequately [25]. To that end, Construct Reliability, also sometimes referred to as internal consistency, needs to be determined for each latent construct in a SEM model [25].
- *Convergent Validity:* The measurement of Convergent Validity considers the correlation between responses obtained by maximally different methods of measuring the same construct [25].
- *Discriminant Validity:* Discriminant Validity for the measurement portion considers the level of dissimilarity in the measurements obtained by the measurement tool for different constructs [25].

To evaluate the reliability and validity of the structural portion (outer model) of the SEM model the following will be tested:

- *Coefficients of Determination (R^2) for Endogenous Variables*: This metric reflects the share of an endogenous construct's variance explained by related endogenous or exogenous constructs [25].
- *Path Coefficient Significance:* Similar to covariance based multiple regression techniques, the quality of the structural portion of a SEM model can be investigated by means of a bootstrapping procedure [25] in order to determine the significance levels of the path coefficients γ_c and β_d, for all applicable indexes c and d.
- *Predictive Validity:* In order to determine the Predictive Validity of the SEM model the Stone-Geisser non-parametric test will be performed [24]. Based on a blindfolding procedure [25], this test requires two data sets: One set for SEM and the other for determining the SEM model's Predictive Validity. This test will be the most important assessment of the SEM model's TF capability.

VI. Conclusions

Applying inductive reasoning to the work of Sohn and Moon [16], as well as Steinberg [12][15], the paper derived a generic framework for SEM based DF for TF. Unlike most TF approaches [16] the proposed framework not only caters for complex and hierarchical structural relationships between technology indicators and TF output metrics, but also allows for non-linear and non-Gaussian factors and cyclical dependencies amongst model variables (which can be either latent or directly observable). The paper also presented the research methodology that will be used in a future study (in the context of the NREN technology domain) in order to construct and evaluate a SEM model (which will be an instantiation of the generic framework) for the TF of NREN related output technology indicators. Should this study be successful in proving the viability of the proposed generic framework as an effective tool for TF, additional studies could for example explore the optimal selection of SEM model instantiations for a given technology context [26].

References

[1]   Mack, C.A.; "Fifty Years of Moore's Law," *IEEE Transactions on Semiconductor Manufacturing*, vol. 24, no. 2, pp. 202-207, 2011.

[2]   Nielsen, J.; "Nielsen's Law of Internet Bandwidth, April 1988," Retrieved 8 January 2013 World Wide Web, http://www.nngroup.com/articles/nielsens-law-of-internet-bandwidth/.

[3]   Metcalfe, B., "Metcalfe's Law: A Network Becomes More Valuable as it Reaches More Users," *Infoworld*, vol. 17, no. 40, pp. 53-54, 1995.

[4]   Nyberg, A. and Plamgren, S.; "Using Indicators for Technology Monitoring: Steps Toward a Proposed Framework," *Chalmers University of Technology, Master of Science Thesis in the Master Degree Programme: Business Design*, 2011.

[5]   Porter, A.L.; "Tech Mining to Drive Open Innovation," in *Proceedings, First International Conference on Technology Innovation, Risk Management and Supply Chain Management*, pp. 1-13, 2007.

[6]   Lichtenthaler, E.; "Technological Change and the Technology Intelligence Process: A Case Study," *Journal of Engineering and Technology Management*, vol. 21, pp. 331-348, 2004.

[7]   Chang, H., Gausemeier, J., Ihmels, S. and Wenzelmann, C.; "Innovative Technology Management System with Bibliometrics in the Context of Technology Intelligence," *Trends in Intelligent Systems and Computer Engineering: Lecture Notes in Electrical Engineering*, vol. 6, pp. 349-361, 2008.

[8]   Porter, A.; "QTIP: Quick Technology Intelligence Processes," *Technological Forecasting and Social Change*, vol. 72, pp. 1070–1081, 2005.

[9]   Buchroithner, M.; "Geodata Interrelations: Inventory and Structuring Attempt of Taxonomic Diversity," in *Proceedings, 2nd Conference on Fusion of Earth Data: Merging Point Measurements, Raster Maps and Remotely Sensed Images*, pp. 11–15, 1998.

[10]   Wald, L.; "An Overview of Concepts in Fusion of Earth Data," in *Proceedings, EARSeL Symposium 1997, Future Trends in Remote Sensing*, pp. 385–390, 1998.

[11]   Wald, L.; "Some Terms of Reference in Data Fusion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 3, May 1999.

[12]   Steinberg, A.N.; "Context-Sensitive Data Fusion using Structural Equation Modeling," in *Proceedings, IEEE 12th International Conference on Information Fusion*, July 2009.

[13]   Haenlein, M. and Kaplan, A.M.; "A Beginner's Guide to Partial Least Squares Analysis," *Understanding Statistics*, vol. 3, no. 4, pp. 283-297, 2004.

[14]   Jöreskog, K.G.; "A General Method for Estimating a Linear Structural Equation System," in *Structural Equation Models in the Social Sciences*, eds.: A.S. Goldberger and O. D. Duncan, New York: Seminar, 1973.

[15]   Steinberg, A.N. and Rogova, G.; "Situation and Context in Data Fusion and Natural Language Understanding," in *Proceedings, IEEE 11th International Conference on Information*, 2008.

[16]   Sohn, S.Y. and Moon, T.H.; "Structural Equation Model for Predicting Technology Commercialization Success Index (TCSI)," *Technological Forecasting and Social Change*, vol. 70, no. 9, pp. 885-899, 2003.

[17]   Staphorst, L.; "The Impact of Intellectual Property Rights from Publicly Financed Research and Development on Governance Mode Decisions for Research Alliances," *Gordon Institute of Business Science, Master of Business Administration Thesis*, 2010.

[18]   Wold, H.,"Path Models with Latent Variables: The NIPALS Approach," in *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modelling*, eds.: H.M. Blalock, A. Aganbegian, F.M. Borodkin, R. Boudon and V. Capecchi (eds.), New York: Academic Press, pp. 307–357, 1975

[19]   Porter, A.L. and Cunningham, S.W.; *Tech Mining: Exploiting New Technologies for Competitive Advantage.* vol. 29, Wiley-Interscience, 2004.

[20]   Watts, R.J. and Porter, A.L.; "Innovation Forecasting," *Technological Forecasting and Social Change*, vol. 56, no. 1, pp. 25-47, 1997.

[21]   Grupp, H.; *Foundations of the Economics of Innovation: Theory, Measurement and Practice.* Cheltenham: Edward Elgar, 1998.

[22]   Chang, H.; *A Methodology for the Identification of Technology Indicators.* HNI-Verlagsschriftenreihe, vol. 233, 2007.

[23]   TERENA; "TERENA COMPENDIUM of National Research and Education Networks in Europe," Retrieved 12 January 2013 World Wide Web, http://www.terena.org/activities/compendium.

[24]   Zikmund, W.; *Business Research Methods, 7th Edition.* Ohio: South-Western Cengage Learning, 2003.

[25]   Vinzi, V.E., Chin, W.W., Henseler, J. and Wang, H.; *Handbook of Partial Least Squares: Concepts, Methods and Applications*. Berlin, Heidelberg: Springer, 2010.

[26]   Raftery, A.E.; "Bayesian Model Selection in Structural Equation Models," *Sage Focus Editions*, vol. 154, pp. 163-163, 1993.