System Thinking: Crafting Scenarios for Prescriptive Analytics

Jens Weber
Heinz Nixdorf Institute (HNI)
University of Paderborn
Fuerstenallee 11
D-33102 Paderborn, Germany
jens.weber@hni.upb.de

Minhee Cho
Korea Institute of Science and Technology Information (KISTI)
245 Daehak-ro, Yuseong-gu, Daejeon, 305-806, Korea
mini@kisti.re.kr

Mikyoung Lee
Korea Institute of Science and Technology Information (KISTI)
245 Daehak-ro, Yuseong-gu, Daejeon, 305-806, Korea
jerryis@kisti.re.kr

Sa-kwang Song
Korea Institute of Science and Technology Information (KISTI)
245 Daehak-ro, Yuseong-gu, Daejeon, 305-806, Korea
esmallj@kisti.re.kr

Michaela Geierhos
Heinz Nixdorf Institute (HNI)
University of Paderborn
Fuerstenallee 11
D-33102 Paderborn, Germany
geierhos@hni.upb.de

Hanmin Jung
Korea Institute of Science and Technology Information (KISTI)
245 Daehak-ro, Yuseong-gu, Daejeon, 305-806, Korea
jhm@kisti.re.kr

ABSTRACT
This paper focuses on the first step in combining prescriptive analytics with scenario techniques in order to provide strategic development after the use of *InSciTe,* a data prescriptive analytics application. *InSciTe* supports the improvement of researchers 'individual performance by recommending new research directions. Standardized influential factors are presented as a foundation for automated scenario modelling such as the prototypical report generation function of *InSciTe*. Additionally, a use-case is shown which validates the potential of the standardized influential factors for raw scenario development.

Categories and Subject Descriptors
K.6.1 [**Management of computing and information systems**]: Project and People Management – *strategic information systems planning, systems analysis and design.*

General Terms
Design, Management, Measurement, Verification.

Keywords
Standardized Influential Factors, Prescriptive Analytics, Role Model Group, Scenario Technique.

Published at Ceur-ws.org
Proceedings of the First International Workshop on Patent Mining and Its Applications (IPAMIN) 2014. Hildesheim. Oct. 7th. 2014.
At KONVENS'14, October 8–10, 2014, Hildesheim, Germany.

1. INTRODUCTION
The software application *InSciTe* developed by the Korea Institute of Science and Technology Information (KISTI) uses prescriptive analytic methods in order to develop strategies and provide recommendations in order to improve research performance. For example, it calculates measures intended to increase the actual number of academic contributions or recommend in extreme cases a change of research topic. *InSciTe* is a mobile and web based application that uses text mining techniques and methods for Big Data analysis to support researchers in their activities [6]. The output of *InSciTe* describes, amongst other things, measurements to achieve a defined number of published academic contributions or describes which future collaborations the researcher should pursue. The overall goal for applying prescriptive analytics is to achieve continuous improvement in research performance. An additional benefit is the extension of methods of forecasting and identification of opportunities to detect future trends in Research and Development (R&D) [6]. However, the generated recommendations at the end of the analysis process are quite static and the opportunities for developing the suggested measures during the post-analysis process are only partially examined. The *InSciTe* report [5] may describe metrics for increasing the number of publications, recommend a potential strategic cooperation, or an increase in the current number of conference visits. The possible failures that have to be considered are, for example, the personal and individual difficulties in joining collaborations which occur between the researchers or the failure of a cooperation as well as an unreached number of published contributions. Furthermore, there is no analysis of the consequences if research performance decreases or stagnates during a cooperation. For that reason, further research in the context of forecasting and strategy development is required.

This paper describes an initial approach for using the output of *InSciTe* for scenario planning and scenario techniques limited especially to scenario field analysis [4]. In this way, possibilities for the derivation and formulation of strategies for future scenarios based on the *InSciTe* results are provided. This approach should finally support researchers in identifying invalid or infeasible results in the list of *InSciTe* recommendations and suggest alternative courses of action.

2. BACKGROUND

Literature provides a variety of information on different scenario techniques as well as other methods of forecasting for strategy development including prescriptive analytics.

2.1 Prescriptive analytics

Prescriptive analytics is "a set of mathematical techniques that computationally determine a set of high-value alternative actions or decisions given a complex set of objectives, requirements, and constraints, with the goal of improving business performance" [9]. From the business perspective, a pioneer in the field of prescriptive analytics is the enterprise *Ayata* (USA), founded in 2003. This company offers software solutions which allow the usage of hybrid data. Model synergies, data and rules are applied and mathematical models are then combined with hybrid data and business process rules. In this manner, problems in the field of operational research, optimization, decision support and Big Data, can be solved with the support of prescriptive analytics. [1]

2.1.1 InSciTe

InSciTe stands for "Intelligence in Science and Technology" and has been in development by the Korea Institute of Science and Technology Information (KISTI) since 2010. It is a software solution for areas pertaining to "Technology Intelligence Services", "Intelligent Decision Support Services", "Intelligent Technology Analysis Services" and "Prescriptive Analytics for Researchers". It contains semantic text mining techniques, a reporting function for technologies and organizations, representation of technology trends, roadmaps, role model recommendations and prescriptive analytics based on 5W1H[1] [6, 7]. The current status of the software tool in 2013 was *InSciTe Advisory* and the goal in 2014 is to adapt it to an improved system supporting prescriptive analytics. The overall goal is to extend the intelligence of *InSciTe* further. A partial goal is to provide the basic knowledge acquired over the course of this project as part of a useful and applicable business intelligence system. [6]

The described system does not initially support the solution from *Ayata* described in Section 2.1, which provides the analysis and improvement of business processes and future decisions, but determines instead the current position of research progress and performance within a chosen field, comparing existing researchers as well as deriving measures that enhance research capacity in a direction that the identified role models have demonstrated in order to attempt to generally improve and even exceed the performance of a given role model researcher. The relevant role model researchers are grouped together and with the support of these groups, measurements can be derived into a quantifiable form in order to strengthen overall research capacity and performance. The role model researcher could be one or more individual researchers or a research organization pertaining to one or more research fields. [2, 6]

[1] 5W1H: KAIZEN-technique to improve organization by the question-answering method of what, when, where, who, why and how [7]

Currently, the application *InSciTe* can be summarized in the following four steps (see [11]).

(1) **Step 1**: Measuring research performance

(2) **Step 2:** Finding role model researcher or group

(3) **Step 3:** Planning research activities

(4) **Step 4:** Evaluating and applying feedback and reports

2.2 Towards scenario techniques

Scenario techniques have been proven in the field as a method for forward thinking in the areas of changing markets, business fields, and technological development as well as in research and development [13]. This has been shown by the successful usage of these techniques by companies such as UNITY AG as well as Sinus GmbH (Germany). [10, 12]

The usage of scenario techniques is based on two principles. It promotes lateral and cross-functional thinking, which means that linked influential factors must be considered. It also furthers understanding of the considered system within the context of its surrounding environment and helps to make these kinds of systems both recognizable as well as manageable. Scenario techniques are also based on multiple potential futures, in which focus needs to remain on more than one influential factor [4].

The scenario developed on these mentioned principles is known as a generally comprehensible description of a possible future, which arises from a complex network of influential factors. [4] The representation of a development which could lead from the present circumstances to this future situation could be also described. [4]

Scenario techniques may be generally divided into several steps. The first step is the preparation of a scenario in which the target is identified and a general project goal setting is defined. Then the second step, called scenario analysis, starts, which determines the influential factors. The identification of certain influential factors as well as their relevance requires the use of a variety of analytical methods in this step. The third step derives multiple prognoses based on the key factors detected during the scenario analysis. Each key factor enables the identification of several projections, each representing a different development direction. The projections are then described in precise and understandable terms and result in a so-called projection catalog. [4]

These projections are then examined and compared in pairs for consistency within the scenario building context and the result is a collection of characteristics of influential factors that determine a similar consistency level. These bundled projections are clustered in order to provide the basis for creating raw scenarios. The raw scenarios can be tested and finally formulated as detailed, verbally expressed, future scenarios. Opportunities and threats can then be analyzed during the scenario transfer process. A proposed general strategic direction arises from the analyzed result. [4]

Some literature such as [3, 13, 14] refers to extensive explanations of how scenario techniques can be generated. Although those approaches differ, the results, processes, and goals of the scenario techniques are identical to a great extent.

Figure 1 illustrates the basic phases and milestones in brief.

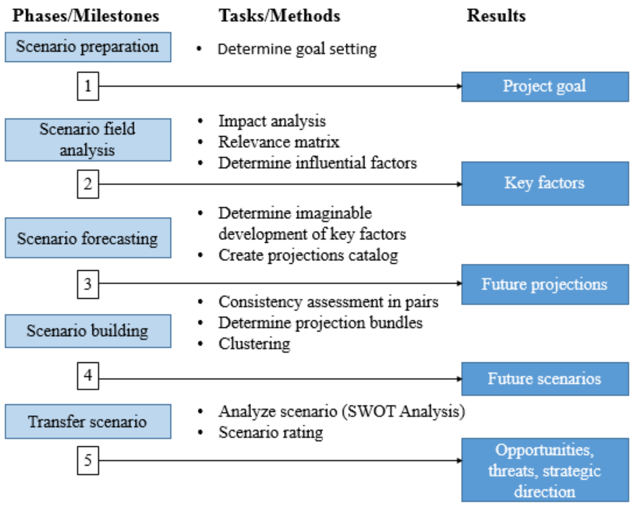

Figure 1. Development process of scenarios [4]

3. IDENTIFICATION OF INFLUENTIAL FACTORS BASED ON INSCITE REPORT

The results of prescriptive analytics via *InSciTe* are described in detail in an automatically generated report for recommendations. These recommendations should improve the research performance of the target person and include, for example, which cooperation should be joined, in which journals the researcher should publish papers, or how the research field should be organized [5]. In practice, difficulties can occur for all the recommendations due to limited resources, failure to establish the suggested cooperation or inability to publish papers due to scheduling restrictions. In order to promote static recommendations as well as to support strategic development, the idea came up to expand prescriptive analytics via *InSciTe* using scenario techniques. It then seems useful to structure the scenario technique according to the automatically generated *InSciTe* report.

3.1 Feasibility of scenario methods

It will generally be necessary to first check whether a potential scenario analysis is technically feasible. Furthermore, it must be proven whether a scenario field analysis will result in the generation of standardized influential factors which could be individually implemented for other *InSciTe* reports. Interfaces between *InSciTe* and scenario techniques allow researchers to create raw scenarios from the generated report. The influential factors predict different developmental directions. The standardization requirement results from a high amount of data, based on the *InSciTe* application and used for the report generation process. It would be too much work to manually create individual local influential factors and they could be incompatible with the predefined process.

3.2 Standardized influential factors for raw scenarios

The automatically generated report by *InSciTe* is always structured identically and contains identical topics so that the derived factors can be adapted to each analysis step. Consequently, the described process for identifying key factors [4] can be stored and the analysis process is more streamlined. The adaptable (standardized) influential factors were identified by detailed structural and text analysis as well as within discussions with the *InSciTe* development team (expert interviews). Table 1 illustrates these influential factors.

Table 1. Influential factors for the scenario analysis based on the InSciTe report

Influential factors			
1	*Role Model Group (RMG)*	7	*H-index*
2	*Research Power Index (RPI)*	8	*Cooperation*
3	*Number of research fields*	9	*Scholar activities* [5]
4	*Currentness of the research fields*	10	*Career activity* [5]
5	*Consumption of resources*	11	*Industrial activity* [5]
6	*Expansion of the research field*		

In addition to the influential factors' adaptability on the analysis results and reports there is the possibility of extension, in particular all possible development directions can be adjusted after more detailed application tests are performed. The following list presents the influential factors and the current developmental directions.

(1) **Role Model Group (RMG):** The RMG includes several researchers or organizations that – due to analysis by *InSciTe* – have certain similarities to the target researcher and therefore recommend the next steps, activities, or cooperation [5, 6].
   - Limited overlap of research fields: There is little intersection in the research focus in an RMG compared to the analyzed researcher and his research fields.
   - High-level overlap of research fields: There is a large intersection in the research areas of the researchers in the RMG and the analyzed researcher and his research topics.
   - No overlap of research fields: The researchers in the RMG have totally different research areas than the target researcher.

(2) **Research Power Index (RPI):** The RPI is a compilation of nine evaluation indicators called "Scholarity", "Influentiality", "Diversity", "Durability", "Emergability", "Partner Trend", "Market Share", "Supply Demand", and "Commerciality". It indicates the strength of the research performance for the analyzed researcher. The merits and demerits of a researcher are evaluated by the RPI. [5, 6]
   - The RPI for RMG members and the analyzed researcher is on the same level.
   - The RPI for RMG members is lower than the analyzed researcher's RPI.
   - The RPI for RMG members is higher than the analyzed researcher's RPI.

(3) **Number of research fields:** Number of research fields which the analyzed researcher focuses on.
  - Many (more than 5)
  - Standard/average (3 to 5)
  - Few (1 to 3)

(4) **Currentness of the research fields:** The factor includes currentness or popularity as well as rarity of his or her research fields.
  - Very current: Often presented in the media
  - Current, but timeless research fields
  - Normal: Mostly basic research or non-popular research fields

(5) **Consumption of resources in the research fields:** What level of human resources and technical equipment are necessary for conducting research in a special field.
  - High level of resource deployment: High expense in the research field – difficult to change the research field or difficult to find and maintain cooperation
  - Low level of resource deployment: low expense in the research field – easy to change the research field as well as easy to find and maintain cooperation

(6) **Expansion of the research field:** Willingness to enter into an additional research field to follow the RMG or to start cooperating.
  - Enter a new additional research field
  - No new additional research field

(7) **H-Index:** H-index is used to measure the impact and quantity of the research performance of an individual researcher [9].
  - Increase
  - Remain constant

(8) **Cooperation:** With whom (out of the RMG members) should the researcher cooperate in order to increase his or her research performance?
  - Cooperation with all members of the role model group
  - Cooperation with several members of the role model group
  - Cooperation with none of the role model group
  - Cooperation with researchers/organizations outside of the role model group

(9) ***Scholar activity*[2]*:*** Scholar activity is related to scientific research actions such as publishing papers, books, etc. [5].
  - There should be more conference and journal paper publications and conference visits, etc.
  - Number of conference and journal paper publications, conference visits, etc. should stay constant
  - The number of conference and journal paper publications, conference visits, etc. should be reduced.
  - No publishing of papers and other research results.

(10) ***Career activity*[2]*:*** Career activity is related to human actions such as receiving awards, building careers, obtaining degrees etc. [5].
  - Increase activities in total
  - Total activities stay constant
  - Fewer activities in total compared with other periods
  - No activities

(11) ***Industrial activity*[2]*:*** Industrial activity is related to commercial actions such as publishing patents, etc. [5].
  - Increase activities in total
  - Activities in total stay constant
  - Fewer activities in total compared with other periods
  - No activities

[2] The influential factors no. (9) to (11) are addressed as specific "activities" by the author; the names in the *InSciTe* report differ.

4. FIRST USE-CASE

The use-case should show which characteristics of the raw scenarios are possible in general. Moreover, the use-case shows how it is possible to prepare the development of scenarios by using standardized influential factors. We focus on the impact and consistency analysis. The consistency analysis guarantees that the raw scenarios contain only influential factors on a high consistency level for further development directions. Here, the impact analysis is necessary to select some important influential factors, because the usage of all eleven influential factors requires high computational effort. Furthermore, a handful of influential factors are generally sufficient for the presentation of the use-case. The influential factor analysis is performed according to [4, 10]. The IT scenario software *Szeno-Plan* developed by Sinus GmbH (Germany) supported the implementation of the use-case.

4.1 Impact analysis for the use-case

The (individual) influential factors are evaluated with regard to their mutual influence on each other. Indirect influential factors were also considered in [4]. The rating scale was from 0 (no influence) to 4 (very high influence). The results of the analysis are presented in the matrix in Figure 2. The results are normalized and plotted as a percentage. The quadrants of the matrix are divided into four sections: "Critical factors", "Driving factors", "Buffering factors" and "Driven factors".

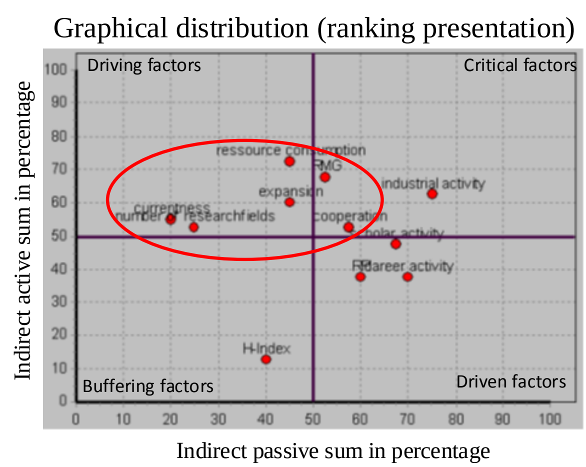

Figure 2. Graphical Distribution of indirect impact analysis

For the use-case, the influential factors in the quadrant "driving factors" were selected (see red marked zone in Figure 2), because these factors consist of a relatively high active sum and small passive sum. However, the factor "cooperation" from the quadrant "critical factors" was also selected because the *InSciTe* reports often recommended it explicitly. As a result, six influential factors were taken. The description of these factors and direction developments can be seen in Section 3.2. Table 2 shows the selected influential factors.

Table 2. Selected influential factors

	Selected Influential factors
1	*Role Model Group (RMG)*
3	*Number of research fields*
4	*Currentness of the research fields*
5	*Consumption of resources*
6	*Expansion of the research field*
8	*Cooperation*

4.2 Consistency analysis

The consistency analysis offers the possibility to identify which of the development directions from the influential factors occur in the several raw scenarios. From each influential factor only one development direction in one raw scenario is represented. The results were discussed with the *InSciTe* developers at KISTI. The highest consistency level is defined by the value 26 and the lowest consistency level is 0. The six influential factors offer 319 different raw scenarios in total. The allocation of the number of raw scenarios to the various consistency levels is shown in Figure 3.

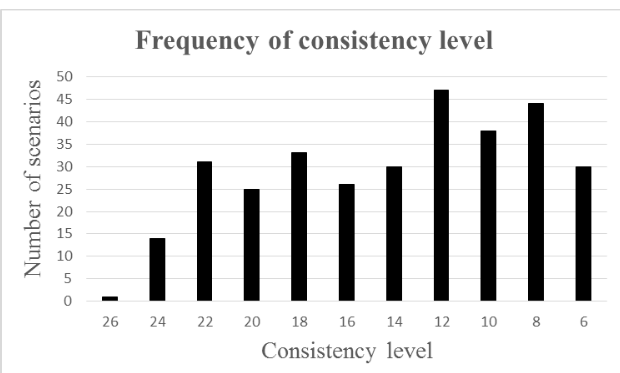

Figure 3. Frequency of consistency level of the several scenarios

Most of the scenarios have a consistency level value ranging from 8 to 12. The fewest scenarios show the highest consistency level values. This is an advantage for the analysis process because the fewer the number of scenarios determined with a high consistency level the lower the analysis effort. A total of 14 raw scenarios were determined in order to provide a consistency level value of 24 (near to the maximum value) and one raw scenario which represents the maximum value (26). Figure 4 shows an overview of the described situation.

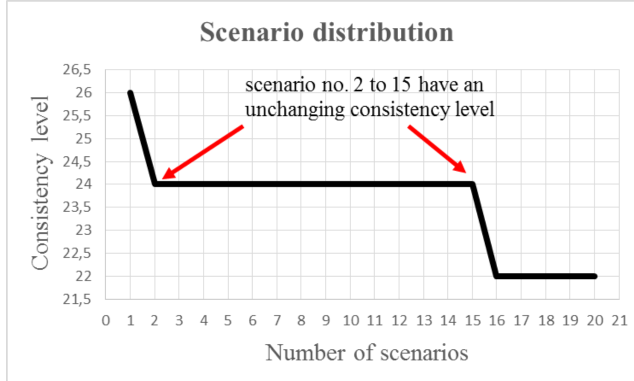

Figure 4. Scenario distribution and consistency level

As seen in Figure 4, the raw scenarios no. 2 to no. 15 seem to be interesting and are selected for further analysis analogous to [4, 10]. Raw scenario no. 1 was excluded; despite having the highest consistency value there were no other alternative characteristics at the same consistency level. There is only one developable scenario and consequently no possible alternative strategy development. Table 3 therefore illustrates briefly four possible configurations of the raw scenarios (see Section 3.2 for any details). Three of them show a consistency level of 24 and the excluded scenario which has the consistency value 26 is also illustrated.

Table 3. Selected influential factors

Scenario No.	**consistency measure**	**Scenario No.**	**consistency measure**
1	26	2	24
Influential factors	**Development directions**	**Influential factors**	**Development directions**
RMG:	limited overlap	RMG:	high overlap
Currentness of research fields:	current	Currentness of research fields:	very current
Consumption of resources:	low	Consumption of resources:	high
Expansion:	enter new	Expansion:	no new
Cooperation:	outside	Cooperation:	with all
number of researchfields:	standard	number of researchfields:	many
Scenario No.	**consistency measure**	**Scenario No.**	**consistency measure**
5	24	7	24
Influential factors	**Development directions**	**Influential factors**	**Development directions**
RMG:	high overlap	RMG:	high overlap
Currentness of research fields:	normal	Currentness of research fields:	current
Consumption of resources:	low	Consumption of resources:	high
Expansion:	enter new	Expansion:	no new
Cooperation:	with several	Cooperation:	with several
number of researchfields:	many	number of researchfields:	many

From the 14 scenarios several were selected which match the profile of the analyzed researcher. These raw scenarios will be transformed using statistical metrics into full scenarios (see [3, 4, 13]).

5. CONCLUSIONS AND OUTLOOK

We showed that automatically generated influential factors, which can be individually applied to the *InSciTe* system as well as to the use-case, can be used for the basic development of raw scenarios according to a standardized procedure. The usage of six selected influential factors leads to 14 raw scenarios as output. It can be estimated that this number may increase with the number of influential factors. The pool of developmental directions for the standardized influential factors is expandable. As another result, the key factor identification from the scenario process (see [4]) can be avoided or optionally enabled due to the usage of standardized influential factors. The potential number of raw scenarios could increase so that the focus is on the automatic transformation of raw scenarios into full scenarios based on the model-based approaches for fully-automated report generation and analysis by *InSciTe*. These described steps will be improved after testing the influence of standard key factors on the *InSciTe* report generation.

6. ACKNOWLEDGMENTS

This work was supported by the IT R&D program of MSIP/KEIT. [2014-044-024-002, Developing On-line Open Platform to Provide Local-business Strategy Analysis and User-targeting Visual Advertisement Materials for Micro-enterprise Managers]

7. REFERENCES

[1] Ayata: Prescriptive Analytics. USA, Austin, Texas. http://ayata.com/ [accessed 2014-06-11]

[2] Cho, M., Hwang, M.-N., Lee, S., Kim, D. J., Hwang, M., Gim, J., Song, S.-K., Jeong, D.-H. and Jung, H. 2013. Towards Prescriptive Analytics for Increasing R&D Competitiveness. In *Proceedings of the CENTRIC 2013. The Sixth International Conference on Advances in Human oriented and Personalized Mechanisms, Technologies, and Services.* (Venice, Italy, October 27 - November 01, 2013). Paper editors from Korea Institute of Science and Technology Information (KISTI). Republic of Korea, Daejeon.

[3] Dönitz, E. J. 2009. Effizientere Szenariotechnik durch teilautomatische Generierung von Konsistenzmatrizen. Empirie, Konzeption, Fuzzy- und Neuro-Fuzzy-Ansätze. 1st Edition. ISBN 978-3-8349-1668-6. Wiesbaden: Gabler, GWV Fachverlage GmbH.

[4] Gausemeier, J. and Plass C.. 2014. Zukunftsorientierte Unternehmensgestaltung – Strategie, Geschäftsprozesse und IT-Systeme für die Produktion von morgen. 2nd Edition. ISBN: 978-3-446-43631-2. München: Carl Hanser Fachbuchverlag.

[5] KISTI: InSciTe Advisory Report. 2014. Korea Institute of Science and Technology Information (KISTI). Republic of Korea, Daejeon. http://inscite-advisory.kisti.re.kr/recommendation?r_id=RE09048535&t_id=all [accessed 2014-06-09].

[6] KISTI: InSciTe. Intelligence in Science and Technology. 2014. Korea Institute of Science and Technology Information (KISTI). Republic of Korea, Daejeon. http://inscite.kisti.re.kr/ [accessed 2014-06-06].

[7] KAIZEN Institute: http://de.kaizen.com/home.html and http://cz.kaizen.com/slovnik/kaizen-slovnik.html [accessed 2014-06-13].

[8] KIT: Karlsruhe Institute of Technology. KIT Library. https://www.bibliothek.kit.edu/cms/h-index.php [accessed 2014-06-01].

[9] Lustig, I., Dietrich, B., Johnson, C. and Dziekan, C.. 2010. The Analytics Journey. In *Analytics Driving better Business Decisions*. USA. November/December 2010. http://www.analytics-magazine.org/november-december-2010/54-the-analytics-journey [accessed 2014-06-13].

[10] Sinus GmbH: Software and Consulting. http://www.sinus-online.com/ [accessed 2014-06-09].

[11] Song, S.-K., Jeong, D.-H., Kim, J., Hwang, M., Gim, J. and Jung, H. 2013. Research Advising System based on Prescriptive Analytics. In *Park, J. J., Pan, Y., Kim, C.-S., Yang, Y. Future Information Technology. Lecture Notes in electrical Engineering, Vol. 309.* S. 569 ff. LNEE. ISBN 978-3-642-55038-6. Heidelberg New York Dordrecht London: Springer. Paper editors from Korea Institute of Science and Technology Information (KISTI). Republic of Korea, Deajeon.

[12] UNITY AG: Management Consultancy for Future-oriented Corporate Management. Germany, Paderborn. http://www.unity.de/szenario-software/ [accessed 2014-06-08].

[13] von Reibnitz, U. 1991. Szenario-Technik. Instrumente für die unternehmerische und persönliche Erfolgsplanung. ISBN 3-409-13431-X. Wiesbaden: Gabler.

[14] Wilms, F. E. P. 2006. Szenariotechnik. Vom Umgang mit der Zukunft. 1st Edition. ISBN 3-258-06988-3. Bern, Stuttgart, Wien: Haupt Verlag.