



COVER SHEET

Brelage, Christian and Recker, Jan and Mueller-Wienbergen, Felix (2006)
Navigational Design of Web Information Systems – Framework Development and
Case Study. In Ljungberg, Jan and Andersson, Magnus, Eds. *Proceedings 14th
European Conference on Information Systems*, Goeteborg, Sweden

Accessed from: https://eprints.qut.edu.au/secure/00004634/01/59_Paper.pdf

NAVIGATIONAL DESIGN OF WEB INFORMATION SYSTEMS – FRAMEWORK DEVELOPMENT AND CASE STUDY[†]

Brelage, Christian, European Research Center for Information Systems, Leonardo-Campus 3,
48149 Muenster, Germany, christian.brelage@ercis.de

Recker, Jan, Queensland University of Technology, 126 Margaret Street, Brisbane 4000,
QLD, Australia, j.recker@qut.edu.au

Mueller-Wienbergen, Felix, European Research Center for Information Systems, Leonardo-
Campus 3, 48149 Muenster, Germany, felix.mueller-wienbergen@ercis.de

Abstract

As prior empirical and conceptual work indicates, success and usability of web information systems is subject to navigational design. Although web information systems are not a new phenomenon and were examined intensively in the past, holistic in-depth investigations of navigational issues seem to be arbitrary rather than theoretically and conceptually founded. In particular, we argue that we are lacking an appropriate description language serving as a shared conceptualization of our subject of research. We present a conceptual framework for describing and assessing web information systems and their navigational capabilities. Moreover, we provide some empirical evidence of its applicability by reporting on a case study that pays special regard to navigational design and usability improvement.

Keywords: Web information systems, navigation, conceptual framework

[†] This research has partly been funded by the German Federal Ministry of Education and Research (BMBF) and the German Aerospace Center (DLR) with the research project “Internet Economy and Hybridity - Modelling of Hybrid Information Systems“ as part of the EU research cluster “Internet Economy” (grant number 01 AK 704).

1 INTRODUCTION

Due to a number of advantages (e.g., ease of use, usage by simple browsers, easy information linking), information systems using common WWW-techniques (e.g., HTML, XML, CSS) have become an essential part of information system architectures by now. Many organizations operate intranets and websites to achieve an easy distribution of information to their employees or customers. Flexibility and, above all, comprehensive information navigation structures are the greatest benefits of web information systems. In general, web information systems are socio-technical systems used to support the goals of an organization by facilitating information access through web technology. They do so by providing a distinct concept of information access enabling individuals to satisfy their information needs efficiently and effectively. Web information systems include those systems, whose software systems comprise web applications rather than traditional applications.

Despite several technical enhancements, mainly hypertext has been used in web technology so far. The origins of hypertext can be traced back to Bush (1945) and Nelson (1965). In general, a hypertext is a graph that contains nodes (contents) and links between them. The act of switching from one document to another by following a link is called navigation. Navigation mainly determines the users' ability to retrieve desired information. If the navigational design does not reflect the users' context perception, usability and effectiveness of information retrieval will be reduced significantly. In this regard, as early as 1990 was user disorientation identified as a major problems of using hypertext (Nielsen 1990), reflected by the well known phrase "lost in hyperspace". In order to use a system efficiently, users always have to know where they are, where they have been and what they can do (Nielsen 1999). However, not all links are equal. While some are used to structure the whole information supply (categorizing), some imply a relationship between contents, and some others are merely created to improve usability (e. g., shortcuts).

Web information systems are not a new phenomenon; however, the description of navigation and its usage for information retrieval remains incomplete. We argue that we know very little about navigational design of information spaces and its utilization for satisfying information demands. We further argue that assessing navigation structures requires a more detailed investigation into the conceptual design of such systems, in particular a theoretically founded conceptualization for describing and assessing navigational design. Otherwise, evaluation and interpretation of empirical results remain arbitrary.

Accordingly, we are addressing the following research questions in this paper:

- Are the methods that are currently applied suitable to assess the navigational design of present-day web information systems? As our literature review in section 2.1 and practical experience shows, this is not entirely the case.
- What kind of conceptualization, capable of reflecting the complex nature of web information systems, is needed by researchers and practitioners in order to achieve an agreement on their subject of research? The answer is given via the systematic construction of a conceptual framework in section 3. We provide means for describing and analyzing navigational capabilities of web information systems that can be used by researchers and practitioners alike to improve and assess navigation in information spaces.
- Does the conceptual framework provide sufficient means for describing and analyzing navigational functionality of complex web information systems holistically? We provide some empirical evidence of the applicability of our framework in form of a case study in section 4.

The remainder of our paper is as follows. Section 2 contains our literature review on empirical and conceptual work on navigational design as well as a brief description of our research approach. Next, we firstly construct our conceptual framework (section 3) and then apply it to the web information system ASInfo (<http://www.asinfo.de>) (section 4). Section 5 concludes the paper with a brief outlook including directions for further research.

2 RELATED WORK & RESEARCH APPROACH

2.1 Literature Review

Considering the relevance of navigational design quality for web information systems, there are surprisingly few scholarly papers concentrating on aspects of navigation and navigational design. Papers focusing on these aspects of web information systems can be roughly distinguished by their research approach: (1) *Empirical papers* present studies on content and design of web information systems as well as on their influences on user satisfaction and their ability to retrieve desired information. (2) *Conceptual papers* mostly deal with modelling aspects of web information systems.

Regarding empirical work on web information systems, Huizingh (2000) presents an empirical study on the composition of web sites. In order to describe web information systems, he distinguishes content from design. The navigation structure is one aspect amongst others in the design dimension of his framework and can be divided into four categories: Tree, tree with back-to-home page button, tree with horizontal links, and extensive network. The majority of web-sites analyzed in this study have a tree-structure or tree with back-to-home page button (Huizingh 2000, p. 131). The complexity of web-sites was found to be correlated to the industry-type (Information industry most complex, services and products the least). Critical factors for web site success are presented by Liu and Arnett (2000), who identify system design quality as one important factor of web site success but do not describe navigational design quality on a detailed level. Bucy et al. (1999) conduct an extensive study on web page complexity and its correlation with site traffic. Rather than navigational design, graphical design issues (usage of colours, frames etc.) and its effect on page complexity are investigated in detail. Khan and Locatis (1998) perform a study on link display and density, and their effects on information retrieval effectiveness. Since their study is based on a quite simple information system, they point out that additional research is needed to investigate effects of information structuring in large, deeply structured information spaces. Using the two factor model developed by Herzberg (1987), Zhang and von Dran (2000) identify navigation as a hygiene factor. Thus, the absence of a clearly defined and appropriate navigational structure is a “dissatisfier” for users surfing the Web.

Regarding conceptual work on web information systems, Farkas and Farkas (2000) present guidelines for web navigation. They specifically address the problem of designing hierarchies and optimal ratios of breadth and depth and distinguish different types of linking, which are adapted in our conceptual research framework (see section 3.2). Larson and Czerwinski (1998) survey several studies about optimal ratios between depth and breadth of hierarchies. Rivlin et al. (1994) present algorithms for (semi-) automated identification and construction of hierarchies in hypertext environments using textual content analysis. Since hierarchies have several advantages (e.g., the ability to be laid out in plan and the inherent concept of aggregation and abstraction), they are intensively investigated with concern to navigational purposes in information spaces (Furnas & Zacks 1994). Furnas and Zacks’ (1994) concept of multitrees was adapted and integrated in our conceptual framework. Van Dyke Parunak (1989) discusses the problem of outlying complex graphs in a flat structure and correlates graphs with navigation. Quantitative analysis and description of such graphs in hypertext are surveyed by Dhyani et al. (2002). Nielsen (1990, 1999) wrote articles in which navigational issues with regard to user interface design are discussed, culminating in his pessimistic statement that the web will suffer a severe usability melt down in 2000, unless some skilful manoeuvring is applied. He sees reasons for inefficient information retrieval on the web in poorly designed user interfaces, partly caused by inappropriate navigation structures. Nielsen’s vision did not entirely come true; however, some of the issues he describes remain unsolved.

Contemplating these related research approaches, we conclude that the empirical work clearly indicates that an adequate navigational design is crucial to a successful implementation of web information systems. However, since the focus of the discussed papers has usually been broad (e.g., web site success, user satisfaction), navigational design issues have only been considered one variable

amongst others and have hence not been investigated in appropriate detail. We argue that this is due to problems related to the detailed description and appreciation of the navigation structure of information spaces and its characteristics. In particular, the navigational structure of complex systems that contain several thousands instances of contents, which is quite common in the Web and intranets today, seems to be very hard to describe. In this context we argue that researchers are lacking a well defined conceptualization of web information systems that pays special regard to navigational issues. The discussed previous conceptual efforts provide useful insights on usage and formalization of navigation structures, some of which we will extend and integrate in our framework. Yet, we argue that we need more complex structures and differentiations of linking types in order to reflect the complex nature of web information systems and their usage in contemporary organizations.

2.2 Research Approach & Background

In order to provide a sound and well defined conceptualization for researchers and practitioners alike, we apply the language critic approach developed by Kamlah and Lorenzen (1984). The core of their approach is the notion that people have to agree on a common language in order to exchange thoughts or beliefs about concepts of a real or imaginary world. In doing so, they form a language community. Without a shared understanding of things, objects and their relations, it is impossible to disclose the world in a scientific and reasonable manner. Kamlah and Lorenzen's work on a systematic, step-by-step construction of common languages serves in this research as a theoretical foundation for the construction of our conceptual framework.

This approach is closely related to the philosophical discipline of ontology referring to the representation of things and their relations in the world, and to means for achieving truth and knowledge about them (Bunge 1977). However, the construction of a common language in Kamlah and Lorenzen's sense is a process that is recurring permanently among people, whereas ontology construction (philosophically) is a singular, unique act by definition. Nevertheless, the conceptualization respectively the common language serves the same purpose as an ontology which is often referred to as a "shared conceptualization" (Gruber 1993). In the information systems discipline, Kamlah and Lorenzen's work was firstly applied by Holten (2003) who constructed a modelling method for information warehouse projects using the language critic approach.

We are applying Kamlah and Lorenzen's approach as follows. We are constructing a detailed conceptualization of web information systems so to achieve a common understanding of our universe of discourse. When doing so we pay special attention to navigational design aspects. The conceptualization serves as our common language and allows us and fellow researchers to derive reproducible results whenever web information systems are designed or analyzed. We will prove the feasibility of our conceptual framework in a selected case. In addition, we derive some exemplary metrics which may be used in further empirical work to assess the navigational design of existing web information systems.

3 DEVELOPMENT OF A FRAMEWORK FOR ANALYZING WEB INFORMATION SYSTEMS

3.1 Information Space Conceptualization

The following conceptualization is partly based on the work by Becker et al. (2003) who present a modelling approach for web information systems. Each web information system contains *Content*. Content represents information objects that are stored in the system, for instance documents, videos, or scripts that compile information from databases dynamically. Contents may be structured (*Contentstructure*), reflecting their complex nature, for instance, different sections of an article or of

related articles. Overall, the conceptualization is corresponding to hypertext theory (Content as nodes, Contentstructure as links between them).

Navigationnodes and *Navigationnodestructure* comprise an arbitrary web structure that can be used for navigation as well. In contrast to the concept “Contentstructure” which implies a semantic relation between two contents, links between navigation nodes do not imply any semantic dependency or relationship between the contents which can be accessed at each navigation node. Usually, web information systems compile sets of contents based on their metadata annotations, e.g., the latest 10 entries in the systems (“newest entries”), or the 10 contents that have been accessed most frequently (“user favourites”). These compilations do not necessarily comprise content that is in some way semantically related.

Hierarchies have traditionally been used in order to structure and classify phenomena from a real or imaginary world. Due to their (technical) advantages and their inherent concepts of aggregation and abstraction, they were intensively investigated in research on hypertext (see section 2.1). Hierarchies reflect an appropriate ordering of the things in the world and are, if properly engineered, intuitively understandable for users. In order to adapt and extend this concept of hierarchies, the concept of *Referenceobject* is introduced. A reference object represents an object or concept from a real or imaginary world that is relevant in the given context of the information system. A subset of reference objects are *Dimensionreferenceobjects*, which are ordered hierarchically and characterized by an explicit assignment to exactly one *Dimension*. Dimension reference objects can further be divided into leafs and nodes. Each of the contents is assigned to at least one reference object, implying that this content is relevant whenever an individual desires to retrieve information about the particular reference object. Thus, reference objects classify the contents.

Dimensions can be understood as navigation trees (hierarchies) in web information systems. Since the Web grows constantly at high rates and the amount of data increases perpetually, it is very reasonable to utilize more than one hierarchy in order to structure the information in the system using different classification schemes (e.g., “Products” and “Time”). Thus, several dimensions may be used in a system, allowing for alternative searching paths for users. Moreover, information can be characterized and classified, respectively, much more appropriately using more than one hierarchy structure. If alternative hierarchies are provided, users, especially non-experts that are not familiar with the expert terminology, may locate information more efficiently. Taken together, the dimensions form a multidimensional space which we refer to as the *information space*. The concept of information space is derived from the work of Holten (2003) who introduced this conceptualization for the specification of management views in information warehouse projects. The concept *Combinedreferenceobject* represents a reference object that is composed of at least two reference objects from at least two different dimensions. As we will see later, this vector uniquely identifies each coordinate in the information space.

In order to reduce the amount of information (e.g., for personalization purposes), it may be necessary to partition the information space into smaller subspaces. The concept *Dimensionsscope* is used to tailor a hierarchy horizontally (selecting aggregation levels) or vertically (selecting “branches” of the hierarchy tree). Dimension scopes can be combined to *Dimensionsscopecombinations* which represent multidimensional subspaces of the information space.

Figure 1 clarifies the concept of information space. Note that the process of identifying and constructing reference objects and their relations is a critical task for information system engineering and necessarily requires human intuition and creativity. In the end, an unfavourable definition of dimensions and reference objects induces a wrong conceptualization of the world and the things that are in it. Consequently, users may not be able to use the system efficiently. However, this process is not the main focus of our paper.

In addition to the introduced constituting concepts of an information space, we deem personalized information delivery and access as an additional requirement in order to reduce information overload

and to improve information retrieval efficiency. The concept *Personalizationobject* is used to represent personalized information access. A personalization object may be defined by *Roles*, *Tasks*, *Users* or *Organizational Entities*. By assigning a personalization object to a dimension scope combination, only a subspace of the information space is deemed relevant for the specific personalization object. Thus, the amount of data users are confronted with can be reduced significantly. The same approach is used to depict access restriction (concept *Accessobject*). These concepts allow for a comprehensive description of advanced capabilities of web information systems. However, as these concepts are not implemented in the investigated case (see section 4), we do not discuss them further.

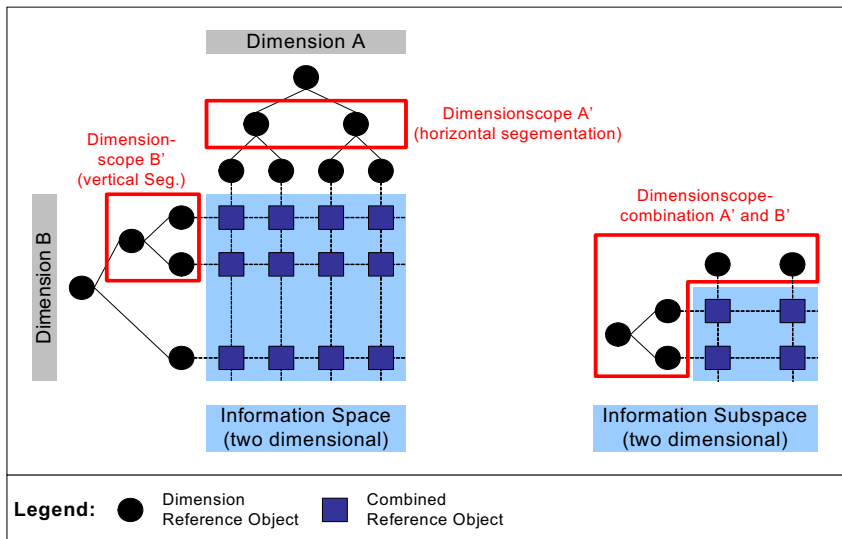


Figure 1: Information Space Conceptualization

3.2 Types of Navigation

Contrasting and extending prior work on navigational issues, we differentiate five different types of navigation in our framework. Firstly, content structures can be used to link information that is semantically related to each other (*content level linking*). Secondly, navigation nodes are used to compile containers of contents or shortcuts within the information space (*navigation node linking*).

Hierarchies are used to structure reference objects as representations of concepts from a real or imaginary world. By descending the hierarchy, users are constantly refining the amount of information that is displayed to them. Thus, hierarchies provide easy and intuitive means for finding information in large information spaces. This third type of navigation is different from content level linking or navigation node linking as users are interacting with objects from the classification scheme and not with content or navigation nodes. We refer to this type of navigation as *dimension hierarchy linking*.

The information space is spanned by dimensions, with each node in the space being uniquely identified by a vector of reference objects (combined reference objects). These combined reference objects can be used for navigational purposes as well: Firstly, links between inter-related combined reference objects cross the complete information space. Secondly, combined reference objects enable users to navigate across more than one dimension simultaneously. They can build complex vectors of reference objects. Only content that is assigned to all elements of that vector is displayed to the user, thereby reducing the amount of information significantly. Yet, this type of navigation can rarely be found in the Web today. As an example, ThyssenKrupp, one of the largest companies for industrial goods and services in the world, bases its ThyssenKruppBase (<http://base.thyssenkrupp.com/>) on this

navigational approach. We refer to these linking types (the forth and fifth in our approach) as *combined reference object linking* and *browsing*.

The different navigation types and the usage of the information space are illustrated in Figure 2. The dimensions are taken from ASInfo (and translated to English), which is investigated in detail in the section 4. Note that we included combined reference object linking and browsing for illustration purposes; the ASInfo case does not implement navigation for combined reference objects.

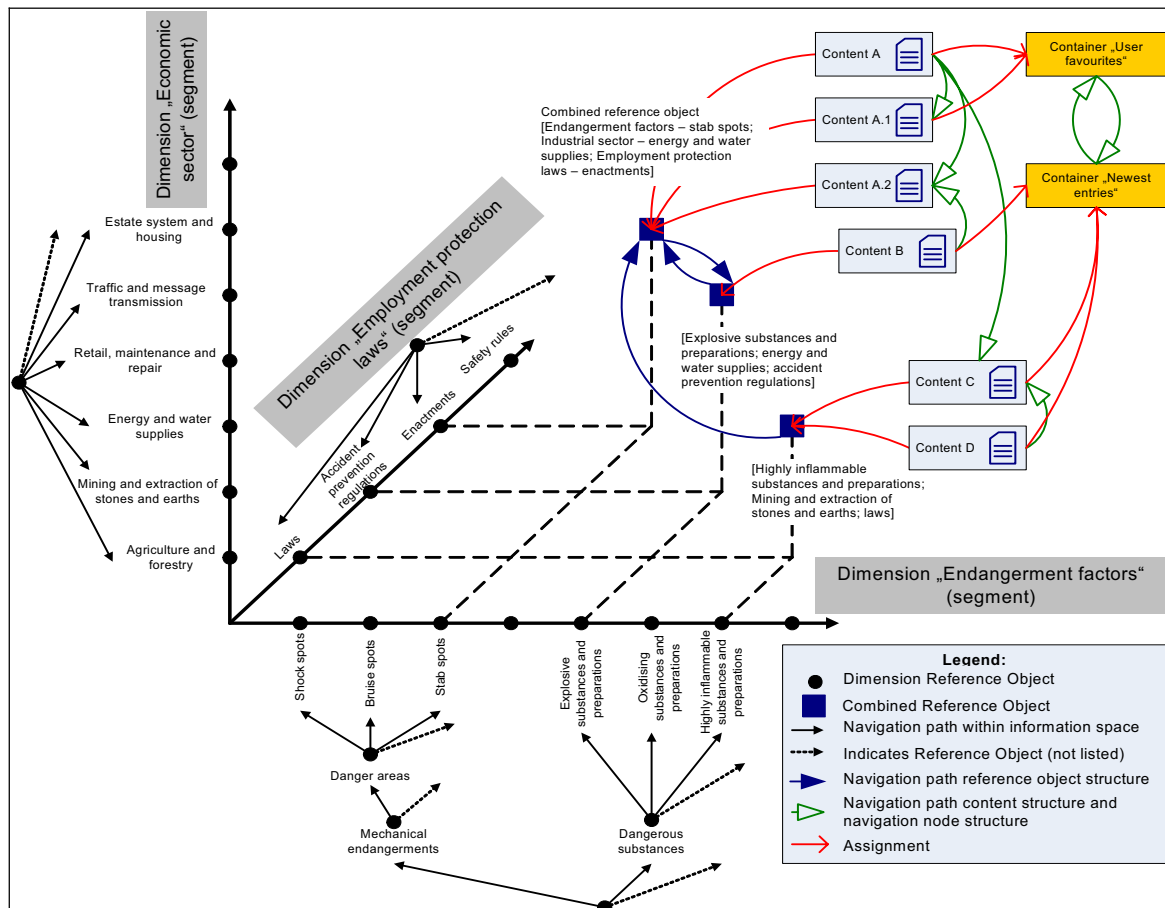


Figure 2: Navigation Types

We argue that a dedicated differentiation of navigation types is important to web information systems design as current systems have reached a level of complexity that requires a more detailed description and investigation reflecting their complexity. Prior research implicitly assumed that links are all equal. In contrast, we argue that, depending on type and structure of content, some navigation types are better suited than others. Some information systems may use content level linking intensively while others (such as ASInfo) do not. More importantly, different types of linking imply different statements about the contents or objects being linked. For instance, switching reference objects changes the context substantially, thereby altering the selection of content, whereas content level linking does not necessarily change the entire context.

4 CASE STUDY

4.1 Case Description

The web information system ASInfo (<http://www.asinfo.de>) provides a comprehensive overview on documents concerning maintenance of industrial health, safety standards, and provisions on labour. The documents provided include, for instance, regulations on wearing helmets on construction sites and handling information on explosives or chemical substances. The system integrates information from more than 130 web-sites of external information providers. Their documents are scanned on a regular basis by the system, indexed and integrated via linking in the system. The indexing algorithm, similar to common web search engines, is based on assessing document similarities and key word occurrences. ASInfo does not provide any information by itself. The system can be accessed without registration or payment by the public and is the most comprehensive library of its kind in Germany.

Users of the system are mainly security appointees of companies who may neither have an appropriate training (e.g., an university degree in chemistry or medicine) nor are experts (especially in small and medium sized organizations). The information providers are specialized institutions with numerous experts in chemistry, industrial engineering, medicine etc. This expertise gap results in users being not always familiar with the terminology being used by the information providers. Consequently, matching information demand and information supply is rather difficult. Moreover, different information providers are using different classification schemes that may be redundant or even inconsistent. Integrating all information resources under one common classification scheme is the biggest benefit of ASInfo. Our analysis of ASInfo focused the evaluation of navigational capabilities, the identification of unsolved issues, and the provision of recommendations for improvements.

4.2 Application of the Framework & Analysis

We applied the framework described in section 3 to ASInfo as follows. Firstly, we identified the counterparts of our conceptual framework concepts in ASInfo and collected data on them from the system. Thereby, we were able to uniquely identify and analyze the navigational capabilities of ASInfo. Since we were not able to access the database of ASInfo directly, we collected the data as displayed in ASInfo itself (e.g., the amount of contents for each reference object). Moreover, a randomized sample of contents (n=240) was selected in order to ascertain information about content length and type. Quantitative measures (see Appendix 1) were derived and computed in order to describe dimensions and the information space on a detailed level.

In our analysis of ASInfo, six main dimensions were identified, namely Endangerment Factors (German: Gefährungsfaktoren), Single Topics (Einzelthemen), Employment Protection Laws (Arbeitsschutzrecht), Research (Forschung), Economic Sector (Branche) and Processes for maintenance of industrial health and safety standards (Arbeitsschutzprozesse). Based on research on earlier work on web metrics, see for instance (Botafogo & Rivlin & Shneiderman 1992, Dhyani & Ng & Bhowmick 2002), we constructed and computed measures in order to describe the dimensions. The results differed enormously. We elaborate on one example in order to illustrate the approach: The dimension “Endangerment Factors” contains 185 reference objects (42 nodes, 143 leafs), whereas the dimension “Research” contains only the root node. Since each of the contents is assigned to at least one reference object, it is reasonable to investigate the *colonization* of reference objects with contents (see Appendix 1 for the definition of colonization). Based thereon, the relative share of the total amount of reference objects for each dimension can be contrasted with the relative share of contents assigned to reference objects in a given dimension. The result gives us information about the distribution of contents and reference objects in the information space. Moreover, the variation coefficient of contents for each dimension can be used to analyze the colonization of nodes (amount of documents that are assigned to each node) in each dimension. Figure 3 gives some results.

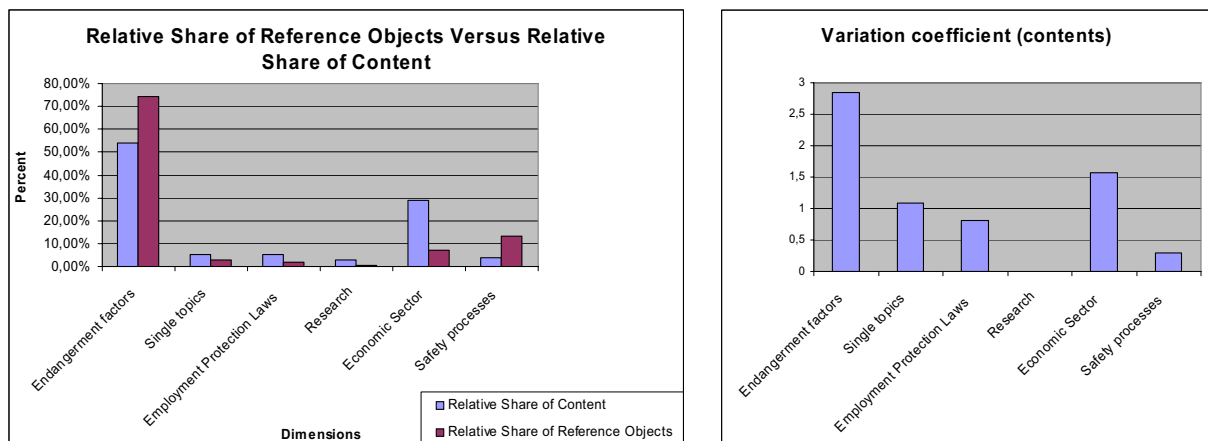


Figure 3: Selected Key Figures

As shown in Figure 3, the dimension “Endangerment factors” comprises nearly 75 percent of all reference objects that are used in the system but contains only 54 percent of all content assignments. The dimension “Economic sector” on the other hand contains 29 percent of all content assignments while comprising 7 percent of all reference objects. On average each node in the dimension “Economic sector” contains over 3000 contents, whereas the average of contents in “Endangerment factors” is 554. Thus, information retrieval effectiveness for users may be lower in the dimension “Endangerment factors”. Also, the variation coefficient (contents) is relative high in both dimensions, indicating that the colonization of nodes varies to a large extent in comparison to other dimensions. Note here that dimension scopes or dimension scope combinations are not used in ASInfo.

ASInfo currently only provides navigation via dimension hierarchy linking. Other types of navigation are not implemented, which, however, may improve the usability of ASInfo (see section 4.3). Thus, users can merely use singular reference objects for navigation purposes. In total, 249 reference objects are used in six dimensions. Two reference objects do not contain any information. If combined reference objects (selection of more than one reference object) were used for navigation, there would be far more possibilities of producing different classification vectors. We refer to this theoretical navigation potential as the *power* of the information space (see Appendix 1). Potentially, ASInfo can provide more than 11 million distinct vectors of reference objects. Each vector can be seen as a predefined searching query that generates the same result as if the elements of the vector have been used singularly in a searching dialog. The main advantage of browsing versus searching is the dedicated explication of the structure of the information space. Thus, users who are not familiar with the structuring of information, or users who do not know what information they really need may still be able to find the desired information by constantly refining their search query.

4.3 Discussion and Limitations

Since dimensions mainly determine the ability to retrieve the desired information efficiently, some suggestions for further improvement can be derived from their analysis. The dimension “Research” organizes thousands of contents but it is not structured at all. Thus, a decomposition of this dimension may improve user information retrieval efficiency substantially. On the other hand, the dimension “Endangerment factors” is deeply and broadly based structured. This may result in the problem of “getting lost” in the dimensions. It may be reasonable to subdivide this dimension into smaller ones. However, the reorganization of dimensions requires the inclusion of a domain expert in order to ensure semantic consistency and adequacy of the resulting information structures.

Another suggestion for improvement is the construction of subspaces and personalized information delivery, which have not yet been considered in ASInfo. Usually, not all documents are relevant for

users (e.g., all content concerning “universities” and “chemistry”). The construction of subspaces for personalized information delivery can improve information retrieval efficiency and effectiveness while the overall structure of the information space remains unchanged.

Furthermore, combined reference objects can be used for navigation by allowing users to build complex search queries via the selection of reference objects from more than one dimension (refer, for instance, to the ThyssenKruppBase). While the structuring of the information space remains unchanged, the amount of information that is relevant to a user query decreases significantly. Thus, information retrieval efficiency may further be improved.

In summation, our framework was very useful in order to analyze ASInfo comprehensively with special regard to navigational issues. The analysis provided evidence for problems that may arise for users. Thus, the analysis can be used for further improvement of the system. However, a more detailed investigation of user behaviour and searching strategies is still outstanding.

Regarding limitations of our approach, the framework allows for the construction of personalized views on information spaces. However, the structure (ordering and linking of objects) is still statically determined and can neither be individually (re-) constructed nor mapped to individual ontologies or other models of knowledge representation. Knowledge management applications, however, require the capturing and mapping of different ontologies in order to match information demand and supply more efficiently. These issues have fervently been debated in literature since the emergence of the semantic web, see for instance (Chaffee & Gauch 2000, Doan *et al.* 2002).

The meaning and purpose of content mainly impacts the structure of an information space. While being optimal from a technical point of view, balanced navigation trees are not necessarily reflecting a natural or intuitive ordering and construction of things and objects in a real or imaginary world. Thus, enforcing balanced navigation trees or other arbitrary information structures that are not ontologically founded or justified may result in severe user confusion and decreased system usability. From a knowledge management point of view this phenomenon can be described as ontological adequacy. The question, whether an information structure is ontological adequate, denotes an important and interesting issue for further research.

At current, the framework implicitly assumes that the information space and its structures are fixed and do not change substantially over time. Given the growth and dynamics of web environments in general, this assumption may be unrealistic. Adaptable or evolvable navigation structures are needed in order to reflect the dynamics of the world. Thus, the framework has to be extended in order to provide means for describing and formalizing navigation structures that evolve over time while still maintaining consistency and ontological adequacy. However, studies in cognitive psychology suggest that user perception of the world is relatively stable and does not change as quickly as the environment (Edelman 1998). Thus, our framework can still serve as a useful assessment approach.

5 CONTRIBUTIONS & CONCLUSIONS

We described the construction of a conceptual framework for web information systems and applied it to a web information system *ex post* in order to describe and analyze the navigational means provided to its users. Our framework was found to be capable of describing current web information systems holistically on a detailed level and provided a starting point for the identification of further improvements. In general, a conceptualization of web information systems using the framework can serve as a standard for further empirical or conceptual work on web information systems (e.g., comparison of web-sites within a specific industry sector or a deeper investigation of navigation within large information spaces). The application of the framework in our case produced accurate as well as interesting results.

However, the broad variety of web information systems and their diverse application areas (including, for instance, knowledge management applications, simple web-sites, intranet applications) complicates

the construction of a holistic framework. Different system types differ to a large extent in structure and user behaviour. For example, news portals or entertainment sites differ substantially from systems like ASInfo and, consequently, require different types of navigation and linking. These differences have not yet fully been reflected in our approach, which, however, we deem applicable to different types of web information systems.

Additionally, the act of navigation is not comprehensively described through “clicking links”. User searching and information retrieval strategies have to be analyzed in more detail in order to fully understand navigational problems of web information systems. Since the structure of information reflects a perception of the world and the things in it, they have to be ontologically adequate in order to provide efficient and effective means for accessing information. Moreover, different users may have different perceptions of the world (for instance, experts versus non-experts), which have to be matched to the information structure of the system. The concept of ontological adequacy is a very interesting and challenging research question, on which we will work in the future.

References

- Becker, J., C. Brelage, M. Thygs and M. Ribbert (2003). Conceptual Design of WWW-Based Information Systems. In Proceedings of the 11th European Conference on Information Systems (Ciborra, C., R. Mercurio, M. De Marco, M. Martinez and A. Carignani, Eds.), Naples, Italy.
- Botafogo, R.A., E. Rivlin and B. Shneiderman (1992). Structural Analysis of Hypertexts: Identifying Hierarchies and Useful Metrics. *ACM Transactions on Information Systems*, 10 (2), 142-180.
- Bucy, E.P., A. Lang, R.F. Potter and M.E. Grabe (1999). Formal Features of Cyberspace: Relationships between Web Page Complexity and Site Traffic. *Journal of the American Society for Information Science*, 50 (13), 1246-1256.
- Bunge, M.A. (1977). *Treatise on Basic Philosophy Volume 3: Ontology I - The Furniture of the World*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Bush, V. (1945). As We May Think. *The Atlantic Monthly*, 176 (1), 101-108.
- Chaffee, J. and S. Gauch (2000). Personal Ontologies for Web Navigation. In Proceedings of the Ninth International Conference on Information and Knowledge Management (Agah, A., J. Callan and E. Rundensteiner, Eds.), pp. 227-234, ACM Press, McLean, Virginia.
- Dhyani, D., W.K. Ng and S.S. Bhowmick (2002). A Survey of Web Metrics. *ACM Computing Surveys*, 34 (4), 469-503.
- Doan, A., J. Madhavan, P. Domingos and A. Halevy (2002). Learning to Map between Ontologies on the Semantic Web. In Proceedings of the 11th International World Wide Web Conference (Lassner, D., D. De Roure and A. Iyengar, Eds.), pp. 662-673, ACM Press, Honolulu, Hawaii.
- Edelman, S. (1998). Representation is Representation of Similarities. *Behavioral and Brain Sciences*, 21 (4), 449-498.
- Farkas, D.K. and J.B. Farkas (2000). Guidelines for Designing Web Navigation. *Technical Communication*, 47 (3), 341-358.
- Furnas, G.W. and J. Zacks (1994). Multitrees: Enriching and Reusing Hierarchical Structure. In Proceedings of the SIGCHI Conference on Human factors in Computing Systems (Adelson, B., S. Dumais and J. Olson, Eds.), pp. 330-336, ACM Press, Boston, Massachusetts.
- Gruber, T.R. (1993). A Translation Approach to Portable Ontology Specifications. *Knowledge Acquisition*, 5 (2), 199-220.
- Herzberg, F. (1987). One More Time - How do you Motivate Employees? *Harvard Business Review*, 65 (5), 109-120.
- Holten, R. (2003). Specification of Management Views in Information Warehouse Projects. *Information Systems*, 28 (7), 709-751.
- Huizingh, E.K.R.E. (2000). The Content and Design of Web sites: an Empirical Study. *Information & Management*, 37 (3), 123-134.
- Kamlah, W. and P. Lorenzen (1984). *Logical Propaedeutic. Pre-School of Reasonable Discourse*. University Press of America, Lanham, Maryland.

- Khan, K. and C. Locatis (1998). Searching through Cyberspace: The Effects of Link Display and Link Density on Information Retrieval from Hypertext on the World Wide Web. *Journal of the American Society for Information Science*, 49 (2), 176-182.
- Larson, K. and M. Czerwinski (1998). Web Page Design: Implications of Memory Structure and Scent for Information Retrieval. In *Proceedings of the SIGCHI Conference on Human factors in Computing Systems* (Karat, C.-M., A. Lund, J. Coutaz and J. Karat, Eds.), pp. 25-32, ACM Press, Los Angeles, California.
- Liu, C. and K.P. Arnett (2000). Exploring the Factors Associated with Web Site Success in the Context of Electronic Commerce. *Information & Management*, 38 (1), 23-33.
- Nelson, T.H. (1965). A File Structure for the Complex, the Changing and Indeterminate. In *Proceedings of the ACM 20th National Conference* (Winner, L., Ed.), pp. 84-100, ACM Press, Cleveland, Ohio.
- Nielsen, J. (1990). The Art of Navigating Through Hypertext. *Communications of the ACM*, 33 (3), 296-310.
- Nielsen, J. (1999). User Interface Directions for the Web. *Communications of the ACM*, 42 (1), 65-72.
- Rivlin, E., R.A. Botafogo and B. Shneiderman (1994). Navigating in Hyperspace. Designing a Structure-Based Toolbox. *Communications of the ACM*, 37 (2), 87-96.
- van Dyke Parunak, H. (1989). Hypermedia Topologies and User Navigation. In *Proceedings of the 2nd Annual Conference on Hypertext and Hypermedia* (Akscyn, R., Ed.), pp. 43-50, ACM Press, Pittsburgh, Pennsylvania.
- Zhang, P. and G.M. von Dran (2000). Satisfiers and Dissatisfiers: A Two-Factor Model for Website Design and Evaluation. *Journal of the American Society for Information Science*, 51 (14), 1253-1268.

Appendix

Name	Symbol	Calculation	Description
Colonization	Col _{abs}	$\sum_{j=1}^{\#T(is)} \#CE(t_j)$	Returns the sum of occurrences of colonized tree elements across all hierarchy trees within an information space.
Variance coefficient (contents)	VC _c	$\frac{\frac{1}{\#T(is)} \sum_{j=1}^{\#T(is)} \left(\sum_{i=1}^n \#C(e_{ij})(t_j) - \frac{1}{\#T(is)} \sum_{j=1}^{\#T(is)} \sum_{i=1}^n \#C(e_{ij})(t_j) \right)^2}{\frac{1}{\#T(is)} \sum_{j=1}^{\#T(is)} \sum_{i=1}^n \#C(e_{ij})(t_j)}$	Returns the relative standard variation of the number of contents related to the mean of number of contents of an information space.
Power	P _{abs}	$\left\{ \prod_{j=1}^{\#T(is)} \left[\sum_{i=1}^{\text{Max}_i \{ \#HL(t_{ij}) \}} \#N(hl_{ij}) + 1 \right] \right\} - 1$	Returns the potential number of materialised search vectors.
<i>Legend</i>			
e _{ij}	Element: Denotes tree element i of hierarchy tree j.		
hl _{ij}	Hierarchy level: Denotes hierarchy level i of hierarchy tree j.		
l _{ij}	Leaf: Denotes leaf element i of hierarchy tree j.		
t _j	Tree: Denotes hierarchy tree j.		
#C(x)(y)	Content-Function: Returns the absolute occurrence of contents assigned to a tree element x of hierarchy tree y.		
#CE(x)	Colonized element-Function: Return the absolute occurrence of colonized tree elements of a hierarchy tree x.		
#HL(x)	Hierarchylevel-Function: Returns the hierarchy level of a tree element x.		
#N(x)	Node-Function: Returns the absolute occurrence of nodes on a hierarchy level x.		
#T(x)	Tree-Function: Returns the absolute occurrence of hierarchy trees in an information space x.		

Appendix 1: Selected measures