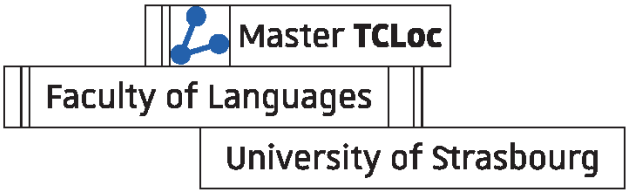


Master in Technical Communication and Localization (TCLoc)

The Relationship Between Content Classification of Technical Documentation and UX Aspects in Documentation Portals

A thesis submitted in partial
fulfillment of the requirements for
the degree of Master of Arts

By

Uta SCHULZ

uta.schulz@etu.unistra.fr

2023

DREW EISENHOWER thesis director

Abstract

Documentation portals that deliver technical documentation of multiple products to multiple user groups carry the company's brand and serve business goals. The users' experience in such portals is therefore an important factor. Findability and context provision typically impact user experience in content-heavy applications. Two content characteristics that can influence these UX-relevant factors are classification and modularization of content. In the past, technical writers handled classification and modularization of technical documentation when preparing technical documentation for delivery. However, in the past decade, processes that typically have been employed to use classification and modularization to ensure findability and context provision in technical documentation were extended by additional technologies or have changed entirely.

Since documentation portals become widespread and integrate more sources and users, this paper investigates the research question: What is the relationship between classification and modularization of content on documentation creation on the one hand and the user experience in a documentation portal on the other hand? To this end, expert interviews with software producers were conducted, and documentary sources reviewed.

The relationship between user experience in documentation portals and content classification and modularization was found to be moderated by use cases. With the relationship clarified, issues that surface in usability tests and user experience research can be addressed in a more targeted way. The relationship can help understand where in a specific process the user experience can be influenced and by whom.

Keywords: technical documentation, user experience, content delivery portals, documentation portals, content classification and modularization

Table of Contents

Introduction

User Experience in Documentation Portals

For companies, it has become convenient to publish technical documentation for several products together with other product related information in one documentation portal to serve several user groups (Antidot, 2020; Ziegler & Beier, 2014). Documentation portals carry the company's brand. The experience users have in such a portal directly connects back to the brand (Lee et al., 2018). Moreover, can user experience issues diminish the return on investment of a documentation portal, for example, by not reducing search time for employees, not reducing support tickets or service time as desired.

It is therefore important for brands to ensure that users from several user groups have a positive experience in the company's documentation portal.

Before the use of documentation portals became popular, creators of documentation, like technical writers, signed for a large part of users' experience with technical documentation, as I will explain in the following paragraphs. I will use script theory to describe in this introduction how this relationship has changed and has made research into the relation between user experience (UX) in documentation portals and content creation necessary.

Further on in this research, I will use expert interviews and findings from literature review to investigate if and how elements that are connected to the creation of technical documentation, such as classification and modularization, are related to aspects of the user experience in documentation portals, such as findability and context provision. I will use cognitive theories and research about the correlation of operational metrics and user experience in such portals to find out how the user experience in documentation portals relates to classification and modularization of technical documentation.

Example situation for a script

When I unpack my new espresso machine, I expect to find a one-pager with instructions on how I can make my first espresso in less than 5 minutes. Alternatively, a QR code to a short video would be OK for me, too.

What is a script?

This kind of situational expectation is called a script. Creators of technical documentation can use it to design documentation that connects the existing mental modal of the user to new information. A script is therefore a means to ensure a positive user experience (St.Amant, 2017), an expectation combined with a learned method, so to say.

Change in script on users' side

When the script changes, the design of the documentation has to be re-evaluated to find out if it has to be adjusted.

Compared to what it was a decade ago, when documentation portals started to become popular (Ziegler & Beier, 2014), the user's script has changed. Currently, users who use digital technical documentation want it searchable the way you search the internet, they want relevant and precise results (Salwitzek & Steuer, 2020). Scrolling through an entire manual is not acceptable anymore in many situations, for example, when end users of products search for how to operate the product, or technicians look for a technical datum.

This change can be explained with a different script that users use, for example, a search engine script in contrast to a print manual script. The details of this script will differ depending on the user group.

When changing from print to web-based documentation, like online help systems, creators of documentation were able to partly address these changed expectations by preserving the modularized structure that previously only served reuse on creation of technical documentation (Soto et al., 2015; Oevermann & Ziegler, 2016). Now, content was also modular when already published. Users, in addition to using the navigational chapter structure, could browse through rather granular full-text search results.

Change in script on creators' side

Creators of technical documentation also have learned expectations. Technical documentation is typically created in a component content management system, in reusable modules of differing sizes (Soto et al., 2015; Oevermann & Ziegler, 2016). To evaluate if a piece of information serves the user's information need, technical writers were able to review the published result of their work. These publications used to have predictable entry points, which were basically created by content creators. Dynamic content delivery as it is supposed to happen in content delivery systems is more complex. It is a challenge for technical writers to predict the situation users encounter before and after searching for information. Therefore, it is more difficult to identify, where and how an adjustment of the content effects the user experience in the documentation portal.

Change in situation

When the situation changes, but the user's script is still the same, this can result in confusion on the users' side (St.Amant, 2021). They will need help to bridge this gap to have a positive user experience. Since it has become convenient for companies to publish technical documentation of several products together with other product related information in one documentation portal to serve several user groups (Antidot, 2020, Ziegler & Beier, 2014), the situation of publishing and accessing technical documentation has changed.

The situation has changed for users:

Landing in or working with such a portal might be part of the script for some user groups, but not for others. Pre-salespeople, for example, might want to aggregate information. They will look for a known script to apply to the situation. That might be the script they have for the Google search engine or the script they have for the Amazon website, for example. End users might look for one

specific piece of information. Ideally, they land on this information just as if they had asked another person. If not, they might type a full-text query into the search field of the application, using a script similar to one of asking another person. For a technician, this situation again might be different.

The situation has changed for documentation creators:

Documentation portals serve several user groups with technical documentation and other materials of several products in a dynamic way, that is, depending on (user-based) variables. This makes the situation in which a user encounters a piece of information hard to predict for documentation creators. When the situation is hard to predict, it is hard to address a gap that might exist between a user's expectation and the encountered situation. Consequently, it makes it hard to design content for a positive user experience.

This raises the questions: Who owns the user experience in documentation portals, and how? What are the variables that influence findability or context provision and thus the user experience. To what extent can creators of technical documentation influence the user experience in documentation portals?

Hypothesis and Research Question

I will therefore hypothesize that aspects of the user experience (such as findability and context provision) on information seeking in documentation portals depend on classification and modularization on content creation and ask the following research question:

- What is the relationship between classification and modularization of content on documentation creation on the one hand and the user experience in a documentation portal on the other hand?

Objectives and Purpose

To get an idea of the aspects of this presumably multifaceted relationship I will take up several viewpoints on classification:

- the viewpoint of documentation portal software producers
- the viewpoint when creating technical documentation and delivery processes

- the viewpoint of users and cognitive aspects
- the viewpoint when measuring user experience aspects in a documentation portal

These viewpoints reveal touchpoints between classification and users in documentation portals that can characterize the relationship.

With the relationship clarified, issues that surface in usability tests and user experience research can be addressed in a more targeted way.

The results of this thesis can furthermore inform the creation and selection of items for testing and researching UX aspects such as findability and context provision in documentation portals.

Scope and Limitations

There are many different kinds of technical documentation portals. I will describe the limits of the concept *documentation portal* for this thesis in the chapter *Concepts*.

In this thesis, I will investigate where user experience in documentation portals has touchpoints with content creation, specifically with classification and modularization thereof. However, it is out of scope of this thesis to evaluate the user experience in specific documentation portals.

Moreover, only the part of classification and modularization of content that is directly related to serving content to different user groups will be investigated, not the topic in its entirety. Technical writers as content creators are assumed to have an influence on the user experience of technical documentation. This by the quality of their writing for the user and the correct use of classification and module size. However, it is out of scope to evaluate the role of technical writers in terms of user experience in depth.

There are many ways to create and deliver content in documentation portals. In this research, I will review content classification on three levels of content intelligence, as will be explained in the chapter *Literature Review*. However, it is out of scope of this research to review methods of content delivery in documentation portals comprehensively.

Also, the review of literature regarding operational use metrics in documentation portals as predictor for user experience aspects is limited to research that appeared to have relevant findings about some navigational elements and possible relationships to classification and modularization and does not include the typical range of web analytics and all possible navigation elements.

The sample size for the expert interviews is 14 people and 7 software products. This purposive sample is limited to mainly German and European experts. Out of 7 software companies, 6 were German and 1 international.

In this paper, I review literature that connect UX to information retrieval (IR) to analyze a possible connection to classification and modularization of content, specifically technical documentation. However, as a result of user experience research in technical documentation portals being sparse, I also reviewed information retrieval systems that only share some of the characteristics of documentation portals, such as query search and navigation, but no filter facets, or delivery of information from one source, but not product related information.

Concepts

Following, I will explain the main concepts and acronyms used in my research question and in the course of the present thesis.

Documentation Portals

Documentation portal is a rather broad concept, even if I limit it to the use in technical documentation.

Other Words for Documentation Portal

In technical communication, documentation portals are also known as content delivery portals, content delivery systems, CDP, information portals, documentation platforms, documentation hubs, self-help portals and more. They are software products for mobile and desktop devices used by various professional and non-professional user groups.

Characteristics of Documentation Portals

In this research, I will consider documentation portals that serve documentation for several physical products in one application for one or more user groups. The purpose of these documentation portals is to dynamically deliver technical documentation to various ends on request. Content modules, their size, relationships and capabilities characterize these documentation portals and are related to my research. Content comes from one publisher, that is, from one company, as opposed to a search engine.

Documentation Portals Not Considered in this Research

I will not consider software applications that serve static documents only, where content modularization by definition does not play a role. Furthermore, I will not consider product centered software applications, like digital twin applications since they do not put the technical documentation in the center of the application and come with different requirements. I will not consider documentation applications that serve the technical documentation of one specific instance of a product, for example, an application serving all the user documentation for a person's car.

Content Creation or Documentation Creation

In this thesis, I will define content creation and the creation of technical documentation very broadly. It includes researching and understanding the user and forming content according to user needs, beyond writing by structuring and classifying it in preparation to content delivery. However, a delivery process that employs, for example, an ontology based on use cases that will determine the documentation portal, independently of the content structure, is not considered a part of the content creation. In contrast, a taxonomy that is leveraged to pre-annotate content to train a model for automatic content tagging is linked to the search function of a documentation portal on content delivery and is therefore considered to be part of the content creation process.

Classification and Modularization of Content

Classification and modularization in this thesis refer on the one hand to the activity of tagging content with content classes and determine the size of the information units to be tagged, and on the other hand to the state of the content being tagged with classes and divided into information units of a certain size. Modularization is an inherent part of conventional content classification since a decision must be made about the size of the information unit to be tagged.

Content Classes, Modules, Module Size

In this thesis, a content class is a structural element of a text unit or information unit that is machine-readable, and which makes the text unit identifiable and selectable, such as metadata. A module in this thesis refers to a text unit or other information unit, while not necessarily the smallest one. In conventional systems, it would refer to a so-called self-contained topic. Its size

can for example be expressed as word count. In this thesis, a module is not limited to an aggregation of several information units into one module but can also be one granular base unit.

User Experience

When users interact with an application, such as a documentation portal, they experience various pragmatic and hedonic aspects of this activity. These aspects can be seen as qualities of the experience users have from the interaction, wherein the experience is not limited to the time of the interaction or the presentational layer of the application. Aspects of content in documentation portals such as findability or context provision are expected to influence user experience aspects such as perceived ease of use, value or credibility. I derived this conceptualization of user experience for the present thesis from sources, such as Schrepp (2019) and Norman (2013), as the part of the multifaceted concept of user experience that is most relevant in the context of this thesis.

Technical documentation

Information products in technical communication, such as product information, also known as technical documentation, can consist of information for use of a product range for specified users and collateral information (administrative, design, compliance or market research information, metadata, data, for example) (Union, 2016). The user-dependency of technical documentation content is multidimensional. A horizontal dimension includes a user range that relates to a product's life cycle (installation, operation, administration, maintenance). A rather vertical dimension includes a user range that relates to skills (rookie, expert) and yet another dimension to the user role (end-user, retailer, purchaser), for example.

Acronyms

IR = information retrieval
CDP = content delivery portal
UX = user experience
CCMS = component content management system
UI = user interface
AI = artificial intelligence

Literature Review

UX and Technical Writing

Technical writing and user experience design overlap. Both are user-centered processes in how they aim to help people to perform tasks. Technical writers and UX designers are users' advocates in product lifecycle processes.

Ziegler (2022) lists new competencies of technical communicators that emerge due to new technologies. With content delivery portals (which I call documentation portals in this thesis) and dynamic content provisioning gaining ground, the approach to content access has changed and become more situational and user centered. As resulting competencies for technical communicators, Ziegler names among others: defining the sequence of access and retrieval processes, navigation, pre-knowledge, context detection and development of rules to map metadata from component content management systems (CCMS) to search facets in the content delivery portal (CDP).

St.Amant (2022) describes a context-based approach that takes situational variables into account by employing script and prototype theories and shows how this approach can enhance the design thinking approach for technical communicators. St.Amant (2017) also details how prototypes are relevant for user experience design and how these prototypes can vary according to the users' context. In the present thesis, the relationship of content classes and user experience in documentation portals will be investigated. In such portals, users make sense of navigational elements by reading the labels and expecting a certain concept. This expectation relates to the said prototypes and depends, according to St.Amant, on the user context.

The Viewpoint When Creating Technical Documentation and Delivery Processes

In this section, I will review the current state of technologies used in delivering content to portal applications. I will focus on the point in the respective processes where classification and modularization play a role on content creation or for content delivery to documentation portals.

Documentation portals can provide content dynamically based on target group and context by adapting information selectively to users' situations and making it accessible (Salwitzek &

Steuer, 2020). Typically, content for technical documentation is created in a CCMS and published dynamically to a documentation portal. In the portal users have dynamic access to the content, that is, depending on the context or permission of the user, different content is displayed and can be retrieved. Further elements typically are facetted search, navigation and full-text search (Ziegler, 2020). Additionally, user-based recommendations like related topics can be displayed (Salwitzek & Steuer, 2020).

The above elements for content retrieval can be found in many documentation portals. However, how the functionality is enabled and implemented can vary significantly (Ziegler, 2020).

Content is delivered with the objective to meet a user's information need in a certain situation. To this end, information must be identifiable and selectable by computer programs. Therefore, information needs to be tagged with classes. The user's information need can be expressed in user stories, use cases or other process definitions, which translates to a specific delivery case (Ziegler, 2019; Ziegler, 2021; Wilhelm & Ziegler, 2021).

In technical documentation creation there are currently two paradigms around (among others): on the one hand content creation in XML-based CCMS structuring content by assigning metadata, on the other hand graph-based CCMS, where content modules are objects in a knowledge graph. The latter provides many degrees of freedom for content delivery and further application of content as data. However, ontology modelling is often perceived to be costly and complex, and not yet as widely spread as CCMS.

CCMS enable modularized structured content creation to maximize reuse (Oberle & Ziegler, 2012). This is because content follows product complexity, standardization and component-based "mass-customization" of products (Ziegler, 2016). Content in technical documentation is therefore structured in topics, typically in a semantic XML structure that can be standardized or a custom model. Tagging and semantic structuring of the (reusable) topics is applied by technical writers on content creation (Ziegler, 2020; Oeverman & Ziegler, 2016).

The delivery of technical documentation to content delivery portals, on the other hand, can be divided into paradigms based on so-called information or content intelligence levels that lend themself to structure the methods used to leverage content classification in documentation portals.

Ziegler (2019b) describes three levels of content intelligence: native intelligence (content metadata are directly processed on delivery), augmented intelligence (additional semantic layer) and artificial intelligence (automated extraction and processing of metadata using machine learning). Following, I will review relevant research for each level that connects classification of content and user experience in documentation portals.

On the level of native intelligence of information Salwitzek and Steuer (2020) investigate the transformation of CCMS content for delivery in a documentation portal. They use metadata with XSL transformation to, on the one hand, create HTML documents to publish in the portal and to, on the other hand, create a facet file to enable structured search. The XML facet file contains all the metadata used for search (such as filter facets) under the premises that the content has been classified with matching metadata. In other words, using an XSL transformation, one facet can match the exact same metadata in the CCMS or various metadata in the CCMS or match CCMS metadata, but have a different name. With this method, content classes and modules can be leveraged unchanged for search and navigation in the delivery portal, or they can be modified by name change or aggregation according to user needs, for example.

In dynamic content delivery, information can be provided on a granular level. The required context can be delivered in several ways. On the level of augmented information intelligence, one method is the use of microDocs and semantic correlation rules (Ziegler, 2021), where a primary object correlates to secondary objects by selection rules. These selection rules can be seen as a use case.

MicroDocs can bridge the context gap between granular topics and whole document. They can also serve as intermediate maturity level between structured content and knowledge graph representation, although microDocs are not limited to having a graph structure. They can be aggregations of topics from the CCMS delivered as package into the documentation portal. Another option is to assign semantic metadata to content for enhanced retrieval in the portal, leveraging the relations of the semantic layer (Salwitzek & Steuer, 2020).

Burkhardt and Clesle (2020) develop an ontology and a documentation portal without using the content classification at all. They create use cases to model domain knowledge in an ontology on the one hand and to design front end metadata and user interface (UI) on the other hand. The ontology is connected to the front end of the documentation portal using queries. Furthermore,

they leverage the ontology to enhance the search experience with synonyms, use case based questions and quick tiles. In the search results, topics are displayed with metadata from the content, such as title, chapter or version. They also provide the possibility to browse the ontology itself to explore the information space.

In technical documentation, retrieval of information that is relevant for specific user needs relies foremost on semantic metadata (Wilhelm & Ziegler, 2021). On the level of artificial intelligence, Wilhelm and Ziegler test IBM Watson for the use of artificial intelligence (AI) in technical communication, comparing a customized model with the generic model for automated content classification, extraction of metadata and knowledge, among others. As one of the steps in the workflow, they describe the creation of a custom metadata taxonomy, a so-called type system, that defines the domain's entity and relation types for the classification of entities. The hierarchy levels of this taxonomy depend on the required granularity of the information to be tagged for later retrieval.

A typical use case for such automatically tagged content is the search function and the facets in documentation portals, where the indexed entity types and the entities themselves are made searchable for users (Reußner, 2018). In a case when creating a classification semi-automatically, Reußner identifies entities of interest and their aspects based on user group characteristics. He reports to additionally take design aspects of faceted search into account at an early stage of the classification process since faceted search is also a goal of his classification. To enhance the search experience, Reußner furthermore suggests the use of personalized tagging by users. By providing a range of tags to limit the variety, in addition to the possibility to create new tags, such a folksonomy may identify classes that enhance the user experience.

The Viewpoint of Users and Cognitive Aspects

Cognitive mechanisms and theories can relate information seeking to user experience and content properties, such as classification and modularization. Some of the studies I reviewed refer to such cognitive mechanisms. Therefore, I will review research on information foraging theory, relevance theory, relevance judgment and script and prototype theory to analyze possible links between user experience and content classification or modularization in the reviewed studies.

Information Foraging

Documentation portals are software applications that provide information. They typically integrate a search interface. Users come here to seek information. How users seek information can be explained by employing information foraging and sensemaking theory (Pirolli & Card, 2005). Moreover, this theory can help in designing search interface applications. Information foraging can answer the question “When do people give up the pursued search trail?” Sensemaking can answer the question “What do people need in order to find meaning in a collection of search results?” Information foraging theory draws on optimal foraging theory or how animals decide what to eat (Pirolli & Card, 1999). Information foraging helps users find relevant information. They follow an information scent and decide how long they will follow this information scent on a certain patch by performing a cost/benefit analysis (Russell-Rose & Tate, 2013).

An information scent is the perceived value that information has for satisfying a user’s information needs (Budiu, 2020). The information scent is the stronger, the more of relevant trigger words search results contain (Spool et al., 2004). In a search user interface, information scent can be derived from descriptive titles relevant to the user, highlighted trigger words in the search results and clear labeling of result with relevant categories (Russell-Rose & Tate, 2013). Users perform a cost/benefit analysis on how long to pursue a certain search path. According to the marginal value theorem (Charnov, 1976), there is an expected gain to a search activity in a certain information patch, that is, a search location. After spending time searching and getting returns, the expected within-patch future gain diminishes to a point below the expected gain in another patch (Pirolli & Card, 1999), that is, an easier catch. Users thus make a tradeoff between comprehensiveness and timeliness to save energy on attention (Russell-Rose & Tate, 2013).

Horiguchi et al. (2012) state that a hierarchical menu structure should be designed to be consistent with user expectations and that hierarchical menus should be designed considering syntactic dependency structures. For an instruction manual of a device, they compute information scent of each menu item and of terms representing user tasks. They did this for an organized menu and for a menu with randomized structure, respectively. User tests prove that information scent distribution can predict which item users select as long as they are unfamiliar with the menu design, and that such unlearned users have to base their navigation decision on the lexical

similarity between menu labels and their own search intent. Horiguchi et al. (2017) develop a method using a software agent that simulates unlearned users' navigation decisions following an information scent to predict problematic menu hierarchy.

User-based Approaches to Classification of Information

Hjørland (2013a) critically reviews literature on how effective user-based or cognitive views have been for classification and indexing of information. He concludes that user-based and cognitive approaches cannot contribute to core questions of knowledge organization, such as sorting information units into classes or deciding on synonymy. He acknowledges that user studies were able to provide one learning: A common finding with the introduction of online systems was that searchers preferred textual search. Classes, in contrast, were not considered user-friendly. Hjørland demonstrates how studies of users are not helpful for creating user-friendly knowledge organization, pointing out four issues.

1. User studies show what is not useful for users, but not how to improve it.
2. User studies reveal typical information behavior, e.g., avoidance behavior. But certain information or quality of information cannot follow common behavioral trends.
3. Generalizing user studies that involved information is problematic.
4. Users' knowledge is closely connected to the information they are looking for. They cannot be seen as an external factor.

He continues that semantic relations are context-dependent and cannot be deduced by universal cognitive principles. The debate in the reviewed literature was basically discussing if concepts should be studies based on psychological studies or based on cultural and domain-specific studies (Hjørland, 2013a).

Byström (1999) researches how perceived task complexity (work task), information types and source examination relate, and finds that neither task complexity nor the need for multiple information types is related to the increase in external information source use. On the other hand, growing task complexity increases the use of people inside the organization and decreases the use of internal documentary sources. Her first one of the resulting 11 statements is: "As soon as information acquisition requires an effort people as sources are more popular than documentary sources."

Prototypes in Concept Theory

In essence, the above is indicating that finding the right word for the right person is not an easy task. Gabora, Rosch and Aerts (2008) try a context-sensitive approach to concept theory while summarizing the history and current state of concept theory. Concepts help us sort instances into classes or categories, and thereby compare and understand situations. People's categories form around prototypes that people have. In contrast to concepts, however, prototypes are less demanding in that they need not have defining attributes in common with other category members and boundaries defined. They can be based on experience, culture, goals or any other aspect. Prototypes are very sensitive to context (Gabora et al., 2008).

Script and Prototype Approach to Context-Sensitive UX

A practical approach to this context sensitivity of understanding offers St.Amant (2018) by combining script theory with prototype theory in an adapted form. A script is a type of schema that people's brains create over time to save processing capacity by summarizing a complex activity into one cognitive unit. The script is tied to a particular setting which needs to be recognized to trigger the script. This is where prototypes of place come into play. People try to match characteristics of the setting to a prototype they have (St.Amant, 2018). This matching is commonly known as expectation. A mismatch is a confusing situation. Based on the script and prototype theory, questions are identified that can guide context-sensitive UX research. These questions relate to items, their characteristics, location and use, presence of other people and their roles, and entry and exit points and conditions of the setting (St.Amant, 2017; St.Amant, 2018), that is, the user context. That the user experience in a certain application such as a documentation portal depends on the user context is relevant for the present research since documentation portals are meant to provide content for several user groups, as the expert interviews will reveal.

Relevance, Relevance Theory and Relevance Judgment

One of the most important functions of documentation portals is information retrieval (IR) for users. Saracevic (2016) describes relevance as the basic notion underlying all IR systems. Providing relevant information to users according to their queries, profiles and information needs is the objective of any IR system. The principle underlying any classification and categorization and other systems to control information is aboutness, not relevance. It can be organized to fulfil

specific needs, but search, in contrast to IR systems, is not an integral part. Relevance, on the other hand, as the underlying notion of search, is a built-in human mechanism that comes with cognition, widely understood by all people. Relevance is created by systems out of queries, algorithms and information. Users derive relevance based on the relation of context and information.

In contrast to the aboutness of classifications, relevance is always a relation. It explains “what makes an input worth picking up from the mass of competing stimuli” (Wilson & Sperber, 2004.)

Relevance theory (Wilson & Sperber, 2004) centers two principles: The Cognitive Principle of Relevance and the Communication Principle of Relevance. In other words, the human mind prioritizes relevance, and assumes the intent of relevance in communication.

Saracevic (2016) analyzes 21 studies that evaluate relevance judgment criteria and finds seven groups with criteria such as topic, quality, depth, type, authority, value in use, confidence (see *Appendix B* for the list). Another finding was, that people apply multiple criteria interactively and do not infer relevance on topicality alone.

Chu (2011), in a study of relevance judgement factors as part of the 2007 Legal TREC track interactive task of 80 factors, finds specificity, ease of use and subject matter on the first three ranks.

Xu and Chen (2006) in a study measuring five relevance judgement factors find topicality, reliability and understandability on the first three ranks.

The Viewpoint When Measuring User Experience Aspects in a Documentation Portal

Following, I will review relevant literature about how operational measurements of user behavior in IR applications involving search and navigational elements correlate with UX aspects. Facets and navigational elements can mirror content classes and modules and the measurements might characterize a relationship between content classes and UX aspects of IR application, which I will detail in the chapter *Discussion*.

Gwizdka and Spence (2007) ask what searching behavior can tell us about the difficulty of information tasks. They use one single large website for the tasks studied. They correlate

operational measures of information search with objective task complexity and subjective post-task difficulty. They find objective path complexity to be determined by factors such as length of navigation path to target information, complexity of navigation choices on each page, information assessment by relevance judgement (target) or information scent (non-target). This to reflect the cognitive effort to seek information. Subjective post-task difficulty is found to be correlated to, for example, time on task and number of pages visited, which is related to number of judgements and navigation decisions to be made and thus to the physical and cognitive effort. A high similarity of users' navigational path to the optimal path indicates a high efficiency. Users perceive tasks the easier, the higher the search efficiency and the lower the search effort. The perceived task difficulty is found to be correlated to objective task difficulty, but the data suggests that other user-related variables factor in as well, such as cognitive ability, experience, domain knowledge.

Findability is one of the most important qualitative factors in help applications since the volume of technical content is increasing fast. Gao et al. (2020) conduct a findability study in a help center application (Alibaba Cloud). They argue that the unofficial three-click-rule for web design mirrors the impatience of users. Users come with their own queries in mind, so that measuring objective findability, based on conventional precision and recall, is impractical. The objective findability, they describe here, is the system relevance as opposed to the user relevance in Saracevic (2016). The study measures post-task perceived findability and task complexity on a 7-point Likert scale and with user search behavior measures.

Gao et al. (2020) find that since user behavioral data correlates to a large extent with perceived findability, user behavioral data can be used to locate hard-to-find documents in the help center and examine user paths for improvement. Such data are search type and search complexity, number of clicks made to find a target, navigation time, time spent after search, number of search results browsed. The average time spent on each result page correlates only with subjective task complexity.

Bodrunova and Yakunin (2018) assess how user satisfaction is connected to task complexity and perceived menu complexity. They find that the menu complexity has no impact on users' navigation behavior with tasks of low complexity. In contrast, for high complexity tasks, the perceived menu complexity increases and leads to non-productive search behavior. With high menu complexity users turn to the older menu types not using tag and contextual menu and users'

search behavior lost consistency. The authors find more than two navigational instruments on a page to be counterproductive.

Dorfhuber and Ziegler (2017) investigate how content relevance analysis can provide insights about users of documentation portals and what content they use to find indicators that are relevant to content delivery improvement. One of the indicators was path analysis based on prior content classification as well as the order of navigational elements. They state that content classification can provide for a more precise segmentation for the analysis.

UX of Facets and Facets of Content

A known obstacle to using technical documentation is the vocabulary gap between how a certain user group is able to describe what they search for, and the vocabulary used in the documentation portal. Providing users with pre-defined classes and metadata can mitigate this problem (Reußner, 2018). In search interfaces of applications for dynamic content retrieval, such as in documentation portals, facets typically reflect metadata.

Facet analysis provides a method for creating faceted classes based on users' interests and the aspects of those interests (Hjørland, 2013b). It seems self-evident that studying and involving users is at the base of user-friendly design. Yet, successful systems like the Google search are not build on user studies. The latter was inspired by bibliometric links between papers (Hjørland, 2013a).

Hearst (2006) finds hierarchical faceted metadata a highly understandable data model for search interfaces. It bridges the complexity between hierarchy and full knowledge representation. He assumes that the user interface for search has as one of its design goals to at all times retain a feeling of control and understanding on the user side, which are typical UX factors. As an example, he mentions the eBay Express interface, where designers determined in advance which facet sets are of most interest for the user. He proves, that the details of the design of search interfaces have a high impact on its success.

Gollub et al. (2019) presents a facet filter method wherein facets broaden the search after a query, in contrast to widely used facets that narrow the search results. The facets are generated from a taxonomy that originates in the metadata of the document set. They test a prototype app versus a baseline system using the UEQ questionnaire. For simple search tasks, the system assessment

result was similar. For complex search tasks, however, the new system with the so-called scoped facets scored significantly higher in the user experience questionnaire than the baseline system.

Methods

Primary Data: Expert Interviews

I conducted interviews with a homogeneous purposive sample of experts of content delivery systems used in technical communication. For the sample, I chose content delivery software vendors and consultants on the Tekom fair for technical communication in Stuttgart, Germany, from 8th to 10th of November 2022. Selection criteria for the experts were knowledge of the system's architecture and the required content structure, as well as knowledge of the types of end users of the portal.

In expert interviews with software vendors and consultants, I attempted to find patterns that indicate a relationship between classification and modularization of the content and the user experience in documentation portals, dependent or not on the respective architecture. I attempted to collect themes that might help to clarify to what extent content structure, content creation and content delivery each play a role in the user experience in documentation portals. To this end, I obtained the experts' opinion about if and how elements of classification and modularization in documentation portals depend on user groups, and if and how user needs ought to be considered on content creation to be reflected in the documentation portal. I asked 7 open and 10 closed questions.

After interviewing 14 experts (12 datasets (2 interviews were with 2 interviewees each)), representing 7 software producers and 3 consultancies, data saturation was reached. Of the companies, 8 were German and 2 international. The interviews were conducted in German and English and translated into English if necessary.

Analysis of Documentary Sources

The research that I screened in the chapter *Literature Review* reveals conclusions and results that interact with each other. Cognitive theories such as relevance theory or information foraging theory can put the results of other research in a perspective that can help understand the relationship between classification and modularization of content and user experience in

applications such as documentation portals. I will therefore analyze this research about content delivery, metrics in portal-like applications and cognitive theories to inform my conclusion on a relationship between classification and modularization of content and user experience in documentation portals.

I used the library of the Strasbourg University, ResearchGate, Google Scholar, Journal of Usability Studies and the publications of the institute for information and content management (i4icm.de) to find relevant research in English and German. I used free publications preferably published after 2015. I used my list of viewpoints from the *Introduction* to sort candidate publications.

Discussion

I will start with the discussion of the results of my expert interviews. For detailed results see *Appendix A*.

The Viewpoint of Documentation Portal Software Producers

The interviewees had a heterogeneous view on a possible relationship between content classification and modularization and user experience in documentation portals, as well as on the character of the relationship. Those interviewees, who saw a direct relationship, that is, who linked facets directly to product and content classes, who stated that the module size depends on topics and that technical writers are responsible for context, represented software products that leveraged content and product classes directly into the documentation portal using the respective taxonomies.

When content was delivered using an added layer or semantic technologies, interviewees tended to be less clear about the relationship between content classification and modularization and users and do not connect the user experience directly with content creation.

Some interviewees indicated that a holistic view of user behavior has an impact on designing user experience. They questioned the benefit of a pull approach with granular information for the user and promoted moderating the content for the user.

Interviewees all seemed to agree that usability testing is possible only with the content implemented in the documentation portal, which is a strong indicator that content and usability are connected but does not allow conclusions about the role of content classification.

From the expert interviews, I can conclude, that the technology used in content delivery is decisive for content classification and modularization having a direct relationship with UX aspects in documentation portals. In systems that directly leverage content classes and do not aggregate nor split modules, such a relationship exists and makes classification on content creation responsible for some UX aspects. It is currently common to use such systems, however, the majority of the experts represented systems, where such a relationship is indirect or can be bypassed completely, such as ontology-based systems with use case based delivery scenarios.

According to the interviewees, there are documentation portals where the navigational elements directly reflect the classes and modules in the component content management system (CCMS). However, to link the user experience of a documentation portal to the classification and modularization of content, these elements have to be used by users. If they avoid those content-retrieved interaction elements (such as facets) and fall back on full-text search, users might have no direct touchpoints with classification and modularization of content. It seems evident that the mere existence of such navigational aids does not imply a relation between UX aspects and classification or modularization. Hearst (2006) states that the success of a search interface is highly sensitive to the design details and backs this up by examples where design was improved by usability testing, not changing any of the underlying technologies.

Several of the interviewees stated that usability tests of documentation portals are typically not done. However, usability test would be a measure to become aware of users avoiding navigational elements and lead to improve more popular search types, such as full-text query search.

I will continue with a discussion of the reviewed literature about how the conclusions and results interact with each other and how they support, or not, the hypothesis of a relation between UX in documentation portals and classification and modularization of content.

Relevance Judgement to Identify a Relation Between UX and Content Classification and Modularization

In the chapter *Literature Review*, I retrieved a number of high-ranking relevance judgment criteria (Saracevic, 2016; Chu, 2011; Xu & Chen, 2006). The reason was that user-perceived relevance appears to be an important factor for successful information retrieval. Various of the factors that influence relevance judgment imply a relationship with classification and modularization of content in documentation portals. Topicality of an information object in a search result list as well as in a document may be represented by tags, for example, of product lifecycle categories such as installation or maintenance, or of a product name. Tags of authors or last-update dates may relate to the factor of reliability affecting relevance judgment. Tags in search result lists and in the content itself oftentimes are generated from content classes. Hence, is a relationship between classification of content and relevance judgment in documentation portals present.

Factors like specificity/amount of information, appropriateness to situation and depth or scope may be related to modularization of content. The smaller the document to be assessed, the less processing effort will be required to judge relevance. The more specific, the easier to identify a good match. An information module with several topics will make it harder to judge if the relevant information is present than one with only one specific topic. Hence, is a relation between modularization and relevance judgement in documentation portals present. Relevance judgment is one of three factors in objective task complexity (Gwizdka & Spence, 2007). Objective task complexity correlates with subjective task complexity, the latter being relevant to the user experience in information retrieval systems.

The difficulty with assessing relevance judgment to measure an UX relevant aspect like subjective task complexity in documentation portals is that relevance judgment by users is influenced by many interacting criteria that depend on the user context. Relevance judgment needs to be assessed by people for specific tasks and is hard to measure by analytics software. The sample size would be rather small and, in many cases, not representative for the user base. This makes user relevance judgment as a measurement inefficient. In contrast, system relevance for measuring IR system performance is compared to a previously judged relevance value for a standardized query (Saracevic, 2016).

Dorfhuber and Ziegler (2017) attempt to find out how relevant content is to its users. They suggest path analysis based on prior content classification, and precise content use segmentation for analytics leveraging content classes. This establishes a relationship between classification of content, and the possibility to measure and improve relevant aspects of user experience.

Users find a label the more relevant to their goal, the higher the semantic similarity between menu label and search intent. For users unfamiliar with a specific hierarchical menu, information scent of menu labels can be seen as an UX-relevant parameter in a documentation portal. Information scent can be computed and hierarchical menu labels problematic to usability can be predicted by simulation (Horiguchi et al., 2017). Menu labels may be directly derived from content classes. In such a case, content classes are linked to the information scent of menu labels and difficulties in navigation can be predicted by simulation.

Measurements as Characteristic of a Relationship Between UX and Content Classification and Modularization

The work of Gwizdka and Spence (2007) suggests that the optimal path length and complexity of navigational choices (link labels, visual design) since they correlate with objective task difficulty, affect the user experience. For IR applications where navigational elements are retrieved from content classes, a relation between content classification and subjective task difficulty can be assumed. However, though subjective task difficulty correlates with objective task difficulty, there seem to be other user-dependent factors influencing subjective task difficulty and thus the user experience.

One finding of the findability study in a help application (Gao et al., 2020) is the correlation between the use of several types of search functions (so-called search complexity) and perceived findability. In the Alibaba Cloud Help application those types of search functions (tags, side navigation, content hub) share the same classification and modularization of content. Hence, the UX factor perceived findability in the documentation portal (help center) can be influenced by classification as well as modularization of the content. A prediction model using search complexity and other behavioral measurements may identify hard-to-find documents. However, it does not characterize classification and modularization specifically.

Facets Make Them Run?

The question if facets and navigational aids are actually useful to users in search comes back in Hjørland (2013a) where user studies reveal that searchers prefer textual search. Classes, in contrast, are not considered user-friendly. Also, Byström (1999) states that people are more popular as sources than documentary sources, and this correlates positively with the amount of information types needed. Bodrunova and Yakunin (2018) explicitly mention a "script" of habitual behavior that searchers seem to fall back on when menu complexity is increased.

In contrast, Hearst (2006), Gollub et al. (2019) and Reußner (2018) and nearly all the interviewed experts find facets in documentation portals to be helpful.

What seems to be a contradiction at first can be explained with relevance theory and script and prototype theory. Users bring the expectation to an application that it provides relevant information for them. Facets that match users' conceptualizations, mental models and situational needs can easily be judged as relevant or not and are perceived as helpful. On the contrary, when there is a vocabulary gap between users and facets or a gap in categorization or other expectations, users perceive facets as confusing.

If facets are leveraged directly from content classification without targeting a use case, then facets are not user-centered, but content-centered, not targeted and only randomly helpful for users.

Script and prototype theory provides a basis to develop methods for a user-centered approach to delivering content, such as the method St.Amant (2022) suggests. Such a method, to be effective for content delivery in documentation portals, should be applied to enhance delivery use cases, which can be different from content creation, depending on the technology used.

Technologies on Content Delivery and Use Cases

Use cases appear to be the link between content classification and user experience in documentation portals. In the chapter *Literature Review*, technologies on different levels of content intelligence were discussed. On the native level of content intelligence, it is common to leverage the content structure, that is, the classification and modularization of the content directly as search and navigational elements into the documentation portal. However, all three levels provide the possibility to leverage use cases as a base for search and navigational elements.

On the native level, use cases can be implemented in the XSL transformation. On the augmented level, selection rules in semantic correlation layers, or ontologies can be used to reflect use cases. This can be achieved by microDocs, where content is transformed and packaged, or it can be achieved independently of the content by using the content as data layer, to name two examples of the many available possibilities. On the artificial intelligence level, when creating the entity types and their level hierarchy, and when pre-annotating the training material with a certain vocabulary, the decisions made should be based on use cases as they have an impact on users' search experience.

Summary and Conclusion

Research about the relationship between UX in documentation portals and content creation is necessary since documentation portals have become mediators between content creators and users. Content classification and its use on delivery in applications is a common topic in information retrieval literature. However, there is a gap in the literature on research about UX and search in technical documentation portals and the connection to content creation, classification and modularization. The present thesis contributed by collecting experts' viewpoints on the topic, and by investigating possible characteristics and underlying principles of this relationship such as analytical measurements, relevance judgement or information scent. Also, how the relationship depends on different technologies was investigated.

Touchpoints

1. Touchpoints from experts' point of view

classes/modules	element in documentation portal
role	login
various classes	preassembled search queries
chapters, topics	tree-like side navigation
metadata such as product model	facets/menus

2. Measurable UX aspects of elements in IR applications that may be retrieved from classes and modules (selection)

UX aspect	measurements	elements in UI	source
findability/subjective task difficulty	relevance judgment such as by topicality or reliability	tags in search results	Gwizdka & Spence (2007), Saracevic (2016)
context provision/ subjective task difficulty	relevance judgment such as by specificity, amount, scope, depth of information	information module size	Gwizdka & Spence (2007)
content relevance/accessibility	path analysis, order of navigational elements	navigation, menu	Dorfhuber & Ziegler (2017)
findability	information scent	hierarchical menus	Horiguchi et al. (2017)
findability/subjective task difficulty	optimal path length, complexity of navigational choices	navigational elements	Gwizdka & Spence (2007)
post-task perceived findability	search complexity, find time, number of pages visited	search, menu, side navigation, content hub	Gao et al. (2020)
satisfaction	task complexity, perceived menu complexity	top menu, side navigation, popular links, content hub	Bodrunova & Yakunin (2018)
helpfulness/findability	usability tests, UEQ	facets	Hearst (2006), Gollub et al. (2019)

Hypothesis and Findings

The present thesis could not prove the hypothesis to be true in general, however, the hypothesis is true under the condition that classification is directly leveraged for use in navigational and search elements in documentation portals.

1. The relationship between classification and modularization of content and the user experience in documentation portals is typically mediated by use cases. In content delivery, use cases are technically expressed by semantic data structures and require classification. This classification can in some cases coincide with classification and modularization of content on creation.

2. Use cases can be implemented for content delivery with all technologies reviewed in this research, such as XSL transformations, semantic correlation layers or ontologies. In other words, a user-centered approach to content delivery can be realized independently of the technology used.

3. Classes and modules can be leveraged into the documentation portal directly without name change, aggregation or splitting, or indirectly via a use case based classification middle layer. Use case classification and content classification need to cooperate in that use case relevant information must be made explicit by tagging the content with respective classes. This can happen on content creation or in other processes.

Hence, a relationship between classification on content creation and UX aspects in documentation portals exists, but only when content classes are directly leveraged into the portal without being moderated by a use case. This conventional content-centered method of delivery is one of the methods currently in use.

4. Facets typically originate directly from metadata of a document set. However, the way facets are implemented can have a strong impact on the user experience in documentation portals. This is backed by research evaluating two different facet implementations using the UEQ survey (Gollub et al., 2019) and by usability tests of design variations (Hearst, 2006). Hence, the classes that facets render share their

impact on UX aspects in documentation portals with the design of the facet implementation.

5. Experts found facets in technical documentation portals to be important for users. Some experts promoted a holistic view of users with their habits and context. This is in line with a context-sensitive approach based on script and prototype theory. Script and prototype theory can be used to inform use case creation methods that take the user context into account, not only on a situational and behavioral level, but also on a lexical level to inform facets and navigational elements.

6. Information scent can predict and simulate unlearned users' search behavior in hierarchical menus and be used to find pain points. Information scent can be seen as a characteristic of the underlying classification of a hierarchical menu. Such a menu may have its origin in a use case based structure or come directly from content classes. On the contrary, analytical measurements such as optimal path length, find time and others could not be clarified as characteristic for a relationship between classification and modularization of content and UX aspects in documentation portals. They can help to find pain points in UI elements, but do not reveal how they are related to content classification.

7. When use cases moderate the user experience, there is no direct relationship between classification and modularization of content on creation and UX aspects in documentation portals. This is in line with the notion of aboutness underlying any classification on the one hand, and the notion of relevance underlying information retrieval on the other hand.

Further research

A new edition of the standard IEC/IEEE 82079-1 was published in 2019 (Tillmann, 2020). Usability requirements for dynamic information delivery are specifically outlined in chapter 8.4 and are generally in line with the notions discussed in this thesis. However, recommendations for content structuring do not refer to the notion of different classifications for content organization and for use cases. The standard rather tends to recommend structuring the content according to target audiences among other criteria. Further research into the applicability and possible

shortcomings of the standard's content structuring recommendations with regard to different paradigms of use case based delivery of technical documentation might be valuable.

There are cases when classification and modularization of content on creation impacts the user experience. This depends on the technologies deployed for content delivery. More research into the roles that take part in creating the user experience with regard to classification for content delivery is therefore necessary.

Classification that is leveraged as navigational elements and in search in content-heavy applications such as documentation portals impacts UX aspects. However, other design aspects also weigh in. Research is necessary about the extent to which labels and search classes in content-heavy applications influence UX in comparison to other UI aspects.

Research is scarce about how users would like to use or could successfully use navigational elements and search functions in content-heavy applications, specifically of technical documentation in portal applications. More research might help inform user-centered creation of classification for content delivery, its relevance and its points of usage in the UI of the documentation portal.

Bibliography

Antidot. (2020). *Dynamic delivery. What is it and why does it matter.* https://www.fluidtopics.com/wp-content/uploads/sites/3/FluidTopics-WP-DynamicDelivery-202011en.pdf. Whitepaper. Accessed December 22.

Bodrunova, S., & Yakunin, A. (2018). *Impact of menu complexity upon user behavior and satisfaction in information search* (pp. 55–66). https://doi.org/10.1007/978-3-319-92046-7_5

Budiu, R. (2020). *Information scent: How users decide where to go next*. https://www.nngroup.com/articles/information-scent/ Accessed January 2023.

Burkhardt, D., & Clesle, S. (2020). Ontologies and use case based planning of content delivery. *SHS Web of Conferences*, *77*, 03001. https://doi.org/10.1051/shsconf/20207703001

Byström, K. (1999). *Task complexity, information types and information sources: Examination of relationships*. Tampere University Press.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology*, *9*(2), 129–136. https://doi.org/10.1016/0040-5809(76)90040-X

Chu, H. (2011). Factors affecting relevance judgment: A report from TREC legal track. *Journal of Documentation*, *67*, 264–278. https://doi.org/10.1108/00220411111109467

Dorfhuber, J., & Ziegler, W. (2017). *IN30: Content relevance analytics: Was lehren Delivery Portale über unseren Content und die Nutzer?* Accessed January 23. https://www.i4icm.de/fileadmin/content/HSKA/03_Vortraege/IN30_Content_Relevance_Analytics_Was_lehren_Delivery_Portale_ueber_unseren_Content_und_die_Nutzer_Dorfhuber_Ziegler_.pptx

Gabora, L., Rosch, E., & Aerts, D. (2008). Toward an ecological theory of concepts. *Ecological Psychology*, *20*(1), 84–116. https://doi.org/10.1080/10407410701766676

Gao, Z., Gao, Y., & Yu, J. (2020). What makes it findable? An exploration on user search behavior and the findability of technical documentation. *2020 IEEE International Professional Communication Conference (ProComm)*, 154–160. https://doi.org/10.1109/ProComm48883.2020.00031

Gollub, T., Hutans, L., Al Jami, T., & Stein, B. (2019). Exploratory search pipes with scoped facets. *Proceedings of the 2019 ACM SIGIR International Conference on Theory of Information Retrieval*, 245–248. https://doi.org/10.1145/3341981.3344247

Gwizdka, J., & Spence, I. (2007). What can searching behavior tell us about the difficulty of information tasks? A study of web navigation. *Proceedings of The Asist Annual Meeting - P ASIST ANNU MEET*, *43*. https://doi.org/10.1002/meet.14504301167

Hearst, M. (2006). *Design recommendations for hierarchical faceted search interfaces*. ACM SIGIR Workshop on Faceted Search. Accessed December 22. https://flamenco.berkeley.edu/papers/faceted-workshop06.pdf

Hjørland, B. (2013a). User-based and cognitive approaches to knowledge organization: A theoretical analysis of the research literature. *Knowledge Organization*, *40*, 11–27. https://doi.org/10.5771/0943-7444-2013-1-11

Hjørland, B. (2013b). Facet analysis: The logical approach to knowledge organization. *Information Processing & Management*, *49*, 545–557. https://doi.org/10.1016/j.ipm.2012.10.001

Horiguchi, Y., An, S., Sawaragi, T., & Nakanishi, H. (2012). *Analysis of menu selection behavior using information scent model*. 1514–1519. https://doi.org/10.1109/SCIS-ISIS.2012.6505334

Horiguchi, Y., Kojima, N., Sawaragi, T., & Nakanishi, H. (2017, June). *User simulation to inspect menu hierarchy design using information scent model*.

Lee, H., Lee, K. K., & Choi, J. (2018). A structural model for unity of experience: Connecting user experience, customer experience, and brand experience. *Journal of Usability Studies*, *14*, 8–13. https://uxpajournal.org/wp-content/uploads/sites/7/pdf/JUS_Lee_Nov2018.pdf

Norman, D. A. (2013). *The design of everyday things*. Basic Books.

Oberle, C., & Ziegler, W. (2012). *Content intelligence für Redaktionssysteme*. technische kommunikation, H. 6, S. 48-54. Accessed December 22. https://www.i4icm.de/fileadmin/content/HSKA/02_Publikationen/ICM/CMS-Kennzahlen_mit_der_REx-Methode_Ziegler.pdf

Oevermann, J., & Ziegler, W. (2016). Automated intrinsic text classification for component content management applications in technical communication. *Proceedings of the 2016 ACM Symposium on Document Engineering*, 95–98. https://doi.org/10.1145/2960811.2967153

Pirolli, P., & Card, S. (2005, May). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis* (Vol. 5, pp. 2-4).

Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*, 643–675. https://doi.org/10.1037/0033-295X.106.4.643

Reußner, L. (2018). *Classification of technical documentation*. https://doi.org/10.13140/RG.2.2.19251.53280/2

Russell-Rose, T., & Tate, T. (2013). *Designing the search experience : The information architecture of discovery* (pp. 23–45).

Salwitzek, C., & Steuer, C. (2020, May). *Data transformations from CMS to CDP enriched by semantics*. https://doi.org/10.1051/shsconf/20207703006

Saracevic, T. (2016). The Notion of Relevance in Information Science: Everybody knows what relevance is. But, what is it really? *Synthesis lectures on information concepts, retrieval, and services*, *8*(3), i-109. https://doi.org/10.2200/S00723ED1V01Y201607ICR050

Schrepp, M. (2019). User experience mit Fragebögen messen. In S. Fischer and Hess (Ed.), *Mensch und Computer 2019 - Usability Professionals*. Gesellschaft für Informatik e.V. and German UPA e.V. https://doi.org/ 10.18420/muc2019-up-0101

Soto, A. J., Mohammad, A., Albert, A., Islam, A., Milios, E., Doyle, M., Minghim, R., & Ferreira de Oliveira, M. C. (2015). Similarity-based support for text reuse in technical writing. *Proceedings of the 2015 ACM Symposium on Document Engineering*, 97–106. https://doi.org/10.1145/2682571.2797068

Spool, J. M., Perfetti, C., & Brittan, D. (2004). *Designing for the scent of information: The essentials every designer needs to know about how users navigate through large web sites*. User Interface Engineering.

St.Amant, K. (2022). Context, cognition, and the dynamics of design thinking: Cognitive methods for understanding the situational variables affecting usable design. *Technical Communication*, *69*, 27–39. https://doi.org/10.55177/tc796562

St.Amant, K. (2018). Reflexes, reactions, and usability: Examining how prototypes of place can enhance UXD practices. *Communication Design Quarterly*, *6*, 45–53. https://doi.org/10.1145/3230970.3230976

St.Amant, K. (2017). Of scripts and prototypes: A two-part approach to user experience design for international contexts. *Technical Communication*, *64*(2), 113–125.

St.Amant, K. (2021). Context, cognition, and communication: Understanding how the psychology of location affects health and medical communication. *European Scientific Journal, ESJ*, *17*(30), 8. https://doi.org/10.19044/esj.2021.v17n30p8

Tillmann, M. (2020). *Umsetzung der IEC/IEEE 82079-1 ed. 2 : Unter berücksichtigung anderer branchenspezifischer normen* (R. S. S. Fritz Michael ; Claudia Klumpp ; Alexander Kurz ; Martin Rieder ; Schmeling, Ed.; pp. 1 Online–Ressource (144 Seiten)). tcworld. http://dx.doi.org/10.25716/thm-74

European Union (2016). *The "blue guide" on the implementation of EU products rules 2016*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.C_.2016.272.01.0001.01.ENG. Accessed January 23.

Wilhelm, A., & Ziegler, W. (2021). Extending semantic context analysis using machine learning services to process unstructured data. In *SHS Web of Conferences* (Vol. 102, p. 02001). EDP Sciences. https://doi.org/10.1051/shsconf/202110202001

Wilson, D., & Sperber, D. (2004). Relevance theory. In *UCL Working Papers in Linguistics* (Vol. 14, pp. 606–632). https://doi.org/10.1002/9780470756959.ch27

Xu, Y., & Chen, Z. (2006). Relevance judgment: What do information users consider beyond topicality? *Journal of the American Society for Information Science and Technology*, *57*, 961–973. https://doi.org/10.1002/asi.20361

Ziegler, W. (2020). Extending intelligent content delivery in technical communication by semantics: microdocuments and content services. *SHS Web of Conferences* (Vol. 77, p. 03009). EDP Sciences. https://doi.org/10.1051/shsconf/20207703009

Ziegler, W. (2016). *European perspectives on technical conununication: Drivers and concepts ofContent management systems in the age of globalization and mass customization*. Official Journal of Japan Technical Communicators Assiciation JTCA Accessed January 2023. https://www.i4icm.de/fileadmin/content/HSKA/02_Publikationen/ICM/TCSymposium2016English.pdf

Ziegler, W. (2019a). Delivery zwischen kontext und content. *Technische Kommunication*, *H. 6*, 58–61. Accessed January 2023. https://www.i4icm.de/fileadmin/user_upload/tk0619-Ziegler.pdf

Ziegler, W. (2019b). Drivers of digital information services: Intelligent information architectures in technical communication. *ACM Chapter Proceedings on Educational Technology, Language and Technical Communication Theme: Information Design and Management*, 47. Accessed January 2023. https://www.aijcr.org/_files/ugd/c51627_5dd7fad7e8b44967a45f9d3f6a959af5.pdf#page=54

Ziegler, W. (2021). Semantic correlation rules as a logic layer between content management and content delivery. *SHS Web of Conferences* (Vol. 102, p. 02007). EDP Sciences. https://doi.org/10.1051/shsconf/202110202007

Ziegler, W. (2022). New Roles and Competencies in Technical Communication Induced by Semantics and Analytics. In *SHS Web of Conferences* (Vol. 139, p. 02004). EDP Sciences. https://doi.org/10.1051/shsconf/202213902004

Ziegler, W., & Beier, H. (2014). *Alles muss raus*. technische kommunikation, H. 6, S. 50-55. Accessed December 22. https://www.i4icm.de/fileadmin/content/HSKA/02_Publikationen/ICM/tk614_ZieglerBeier_AllesMussRaus.pdf

Appendix A Interviews

Items and Answers

Apart from contact questions, I asked 7 open and 10 closed questions:

1. **Q4** What is your job in your company?

The interviewees had the following functions in their respective companies: Solutions Architect, Business Development, CEO (2), Sales (2), Customer Success Manager, Consultant (4), Account Manager (2), Product Manager

2. **Q5** There is a ready-made version of the software that clients can use out-of-the-box with some configurations. Yes/No

This question targets the assumption that one type of portal can serve a multitude of users, data sources and use cases, while customization does not go further than configuration. User experience can be different in different user groups. If there is a relationship between the user experience and the content creation this had to be addressed for the different user groups in one out-of-the-box portal.

All 7 software producers have an out-of-the-box solution, or are developing one.

3. **Q6** What are the main reasons of your clients to start a documentation portal project? (multiple selection)*
   A = self-help portal for product end-users
   B = documentation portal for in-company use
   C = documentation portal for internal and external use

This question targets the number and variety of use cases of documentation portals.

4 => self-help portal for product end-users
2 => documentation portal for in-company use
10 => documentation portal for internal and external use

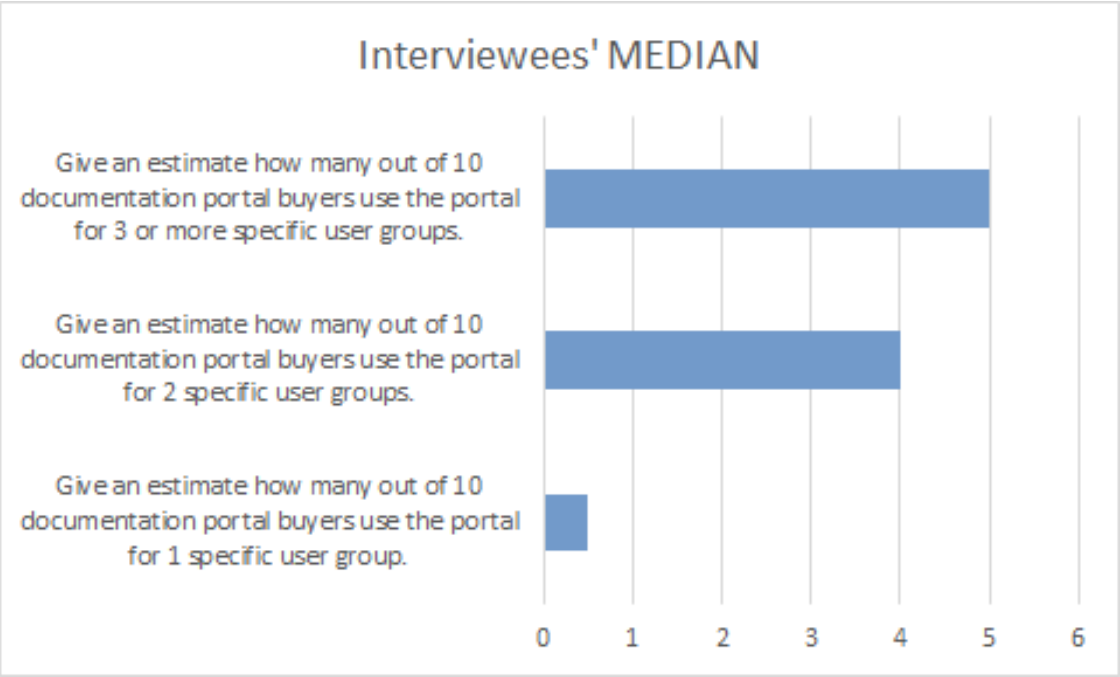
4. **Q7** What other main reasons do companies have to for starting a documentation portal project? (open question)

This question targets the goals of documentation portals in terms of content and users.

3 => pull together information from several user groups
2 => digital transformation
2 => topic oriented information provision
2 => single point of truth
1 => feed information back into the system for quality improvement
2 => integration of other systems
1 => automated content delivery
1 => commercialize information
1 => trainings
1 => data service

5. **Q8-Q10** 3 Questions: Give an estimate how many out of 10 documentation portal buyers use the portal for 1/2/3 specific user groups.

3 questions to inquire how many different user groups these portals might serve:



Qu8-Q10.png

6. **Q11** More ideas about user groups for documentation portal projects?

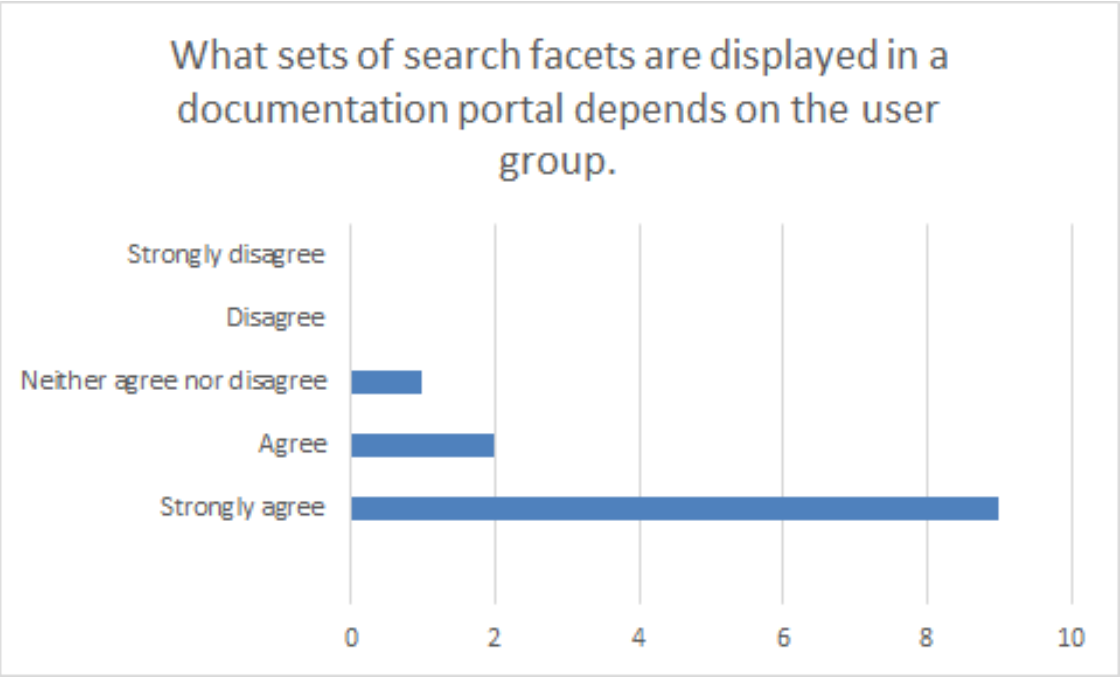
The questions about user groups might reveal relations between users and content or system design.

5 => many user groups is the purpose of the system
1 => people use portals differently
1 => output in the portal differs according to logged-in user
2 => clients don't know their user groups exactly from start

7. **Q12** What sets of search facets are displayed in a documentation portal depends on the user group.

Strongly agree - Agree - Neither agree nor disagree - Disagree - Strongly disagree

This question targets the relation between facets in the portal UI and users.

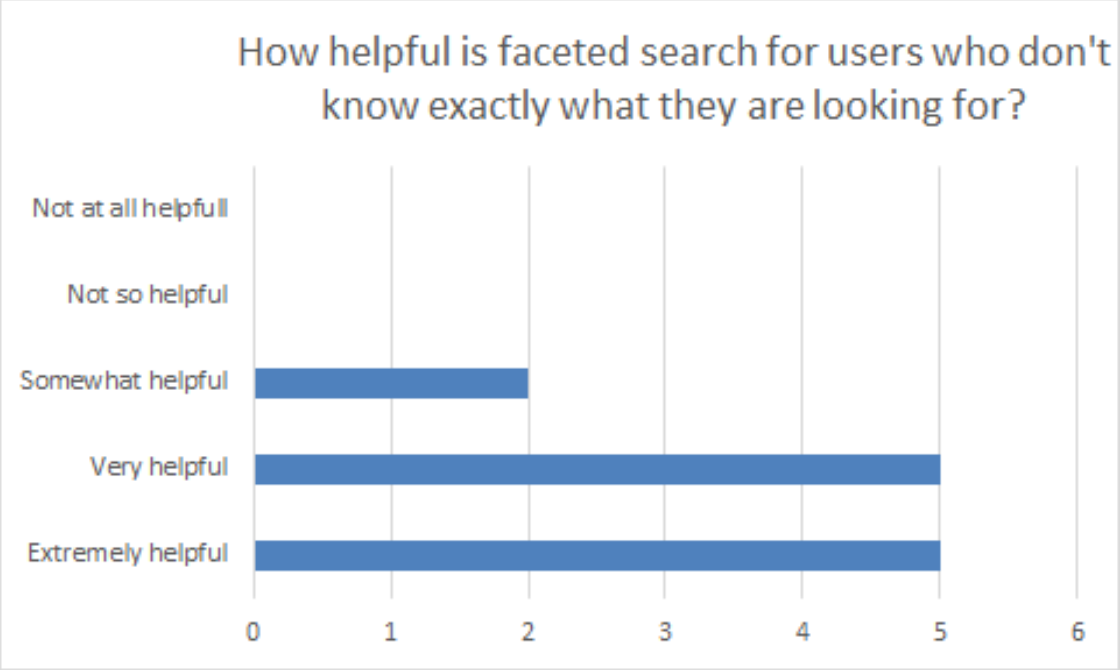


Q12.png

8. **Q13** How helpful is faceted search for users who don't know exactly what they are looking for?

Extremely helpful - Very helpful - Somewhat helpful - Not so helpful - Not at all helpful

This question targets the relation between facets and UX in portals.



Q13.png

9. **Q14** Give an estimate how many out of 10 documentation portal buyers could do with query search only, without faceted search?

This question targets the usability of facets as a specific means of user interaction in search.
Median = 0, 7 said 0, 1 said 1, 1 said 2, 2 said 3, 1 said 7

10. **Q15** What else can you tell me about how important facets are for certain user groups and how they are interdependent?

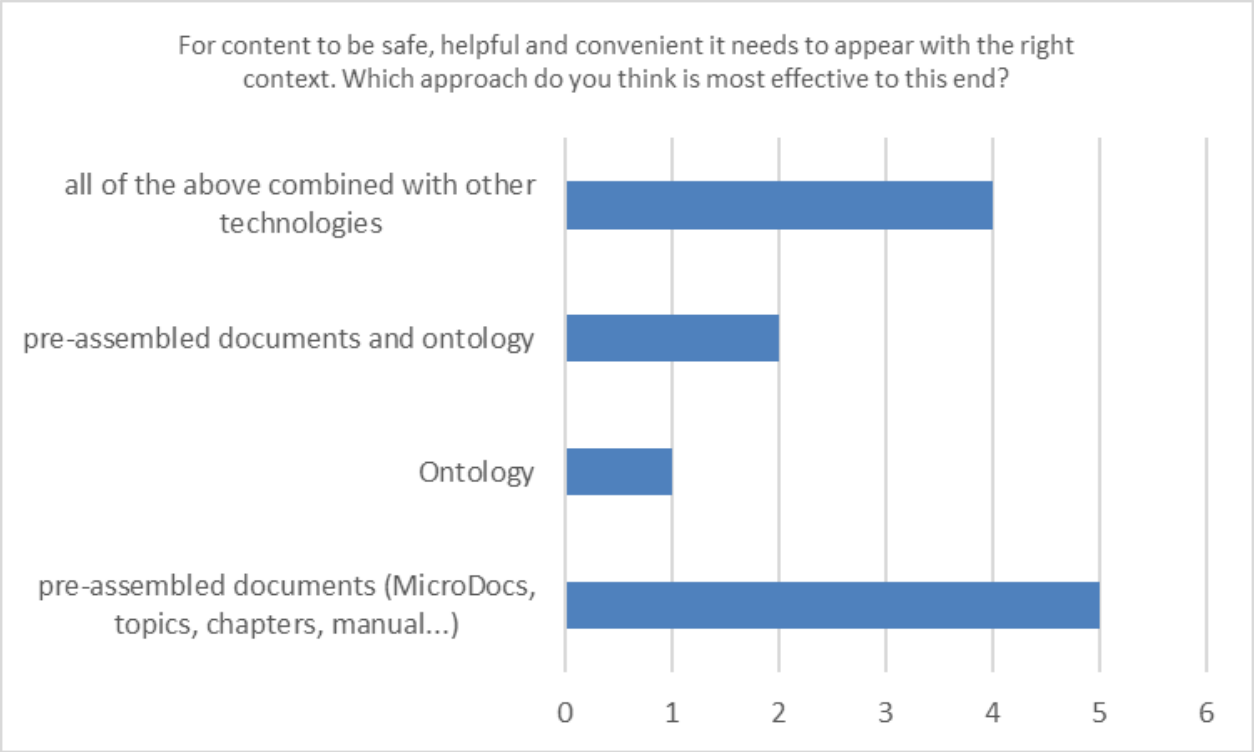
This question targets the existence of a relation between facets and users.

4 => facets come from product
4 => facets (should) come from users/use cases
5 => facets come from CMS
2 => facets are installed in an extra layer

11. **Q16** For content to be safe, helpful and convenient it needs to appear with the right context. Which approach do you think is most effective to this end?

A = pre-assembled documents (MicroDocs, topics, chapters, manual…)
B = ontology
C = pre-assembled documents and ontology
D = both of the above combined with other technologies

This question targets the delivery of information units in their context, assuming that modularization and context are related. The question insinuates that context is related to UX.



Q16.png

12. **Q17** Is there a knowledge graph involved in your system?

This question might serve for correlation.

Yes => 6
No => 1

13. **Q18** More ideas about context approaches?

This question was supposed to reveal approaches and assumptions or viewpoints on delivering contextual information since the lack or abundance of information context can affect the user experience.

If not otherwise indicated by "(#)" it is one person per statement:

- People have learned to use documents that someone else has moderated for them
- When users collect their own books what happens with the safety information?
- One has to imagine the emotional state of such an end user. You want to fall back on well-known concepts.
- There has to be a way for users to make sure they are in the right places.
- Neighboring topics are also relevant.
- Topics are shown in the context of a manual
- Context is the responsibility of the content creator
- Ideally, topics describe all cases with sufficient context. BUT…
- For full-text search and metadata [from CCMS], first step: "What questions do the users have?"
- Set up your own graph-based navigation structure based on task/use case/product/product family
- individual searches represented by different navigation structures
- grouped search queries according to use case
- additionally separate layer via metadata
- There has to be a way for users to make sure they are in the right places. (repeated in this group)
- MicroDocs come from the info model of the ontology.
- Building ontologies requires a high effort

- Missing context was the trigger for developing the knowledge graph based CCMS in the beginning

14. **Q19** How does the module size/granularity of content relate to user needs?

Targets the viewpoint of interviewees on the relation between user and modularity of content.

If not otherwise indicated by "(#)" it is one person per statement:

- User needs connect back to module size/granularity of content.
- The accepted text unit is … well as far as possible graphical. Users prefer videos because they are shorter.
- There are different modules depending on the target group. This is independent of the CDP. So, regardless of whether this target group gets the content from the CDP or from somewhere else.
- Modularization depends on the target group.
- To begin with, one question should be answered with an answer from the CMS [means one module]. This also for the sake of usability.
- Content does not necessarily need to be modular according to the user. This can also be solved by recommendation in search.
- Using the modularization of the content in the documentation portal is not a priority.
- When modularizing, consider content creation vs. content production, e.g., mark content at sentence level, but then "cut" it according to user group during production. -> modularization on creation might serve different purpose than on production
- If the CDP requires smaller modules, this can be done using metadata or sematic technologies
- One question is answered with one module. (#2)
- The technical writer defines the user groups for the text and creates content for different groups.
- To begin with, one question should be answered with an answer from the CMS [means one module.

15. **Q20** In how far are usability tests of the portal done?

This question targets the relation between UX and content in documentation portals.

If not otherwise indicated by "(#)" it is one person per statement:

- I've never seen any user testing done
- not yet
- Usability falls by the wayside with these projects. Also, because there are no experts.

- some clients do
- Proof of Concept
- Proof of Value
- We defined use cases by brainstorming and analyzing existing CMS content from clients.
- No, but there was this master student testing the system: 40% use faceted search to filter, 50% browse in the results.
- Usability tests are recommended to clients during the design phase
- we recommend clients to start with the user
- Clients can use Matomo
- The customer gets the system for 3 months to test. Then he can carry out usability tests himself and have the system adjusted.
- The clients do that, we are out of this
- Usability is the client's business.

Interview Summary

Out-of-the-box solutions

All interviewed software producers had an out-of-the-box solution (or were building one). For different clients and content, the portal would have a similar architecture of delivering content. This suggests that the user needs will be met by content characteristics. Out-of-the-box portal solutions typically provide findability supporting features like facets, navigation and query search. Facets as well as query search can retrieve information from content classes. Navigation is often based on content modules, see for example Bodrunova and Yakunin (2018). The use of out-of-the-box solutions for portals is therefore a strong indicator that in such portals the user experience factors like findability or ease of use are related to classification and modularization of content.

Many user groups

5 of the interviewees said that having many user groups is the purpose of the documentation portal. All of them confirmed that portals actually have several user groups. Having content for several user groups delivered to one technical documentation portal implies that a solution for dealing with user-dependent relevance have to be implemented. In one interviewee's system this was solved by login. There are many solutions beyond login to deliver relevant content for a specific user group dynamically. User group information may be mapped to content classes in a certain layer or user groups may be implemented as metadata in the content. One interviewee stated that with only one user group the portal does not have any added value.

Usability tests

4 interviewees said that they never saw usability tests done and 4 said they do proof of value. 6 stated that usability tests are the responsibility of the buyer. This is an indication that documentation portals, for dedicated usability test, need real content. In other words, the user experience depends on the content, without being limited to this. This allows for the conclusion, that the content, its properties or the delivery process or combinations thereof are responsible for a part of the user experience in the documentation portal.

Facets

11 of 12 interviewees found facets to be important. 4 connected facets to products. Product classes are typically part of the content. 5 connected facets to users or use cases. 5 said facets come from the CCMS and 2 said that facets are installed in an extra layer. Typically, role facets ("admin") and use cases in the form of life cycle information ("maintenance") are part of the content in the form of classes, attributes or metadata. However, user and use case may also refer to a login function or to aggregated content on delivery side ("onboarding") and would in this case be indirectly connected to content classes, for example, by use of an algorithm or mapping ("extra layer"). Hence the interviewees see a relation between content and facets in the documentation portal and consider facets to be important for the user.

Information Context

The interviewees related context to the user experience (findability, control, cognitive/emotional component), to modularization, to classification of content and to various possible technologies. I would like to cite 2 responses that show how user experience in content delivery requires a holistic view of the situation:

> "People have learned to use documents that someone else has moderated for them."
>
> "One has to imagine the emotional state of such an end user. You want to fall back on well-known concepts."

3 interviewees associated context with people's behavior, questioning the benefit of a pull approach with granular information for the user. 1 person, when ask about context, said that there has to be a way for users to make sure they are in the right places. Another associated context with MicroDocs coming from the information model of the ontology, which places use case dependent classification in a middle layer.

5 interviewees related information context to individualized search and had various specific solutions, such as pre-build queries represented by a facet, a QR code or any element. 1 interviewee considered navigation based on tasks, use cases, product family or product to represent context. Another representation was an additional user or use case dependent metadata layer. This relates metadata to user-based findability.

4 interviewees related context to neighboring topics and considered technical writers to be responsible for context. This indicates a connection between content modularization on creation and the UX relevant aspect of context provision.

Module Size

Content is produced in information modules for the sake of reusability. I asked the interviewees how they relate module size to user needs. 5 interviewees related the module size or the modularization to user needs or said it to be user dependent. 2 responded that one topic should answer one question (of the user). 2 interviewees said that module size in the CCMS is not necessarily related to user needs. These user needs in terms of module size can be addressed by recommendation in search, semantic technologies, or are not a priority at all, because modularization serves reuse on content production.

Appendix B Relevance Judgement Criteria

Saracevic (2016) analyzed 21 studies that evaluated relevance judgment criteria and grouped them as follows:

- Content: topic, quality, depth, scope, currency, treatment, clarity.
- Object: characteristics of information objects, e.g., type, organization, representation, format, availability, accessibility, costs.
- Validity: accuracy of information provided, authority, trustworthiness of sources, verifiability, reliability.
- Usefulness or situational match: appropriateness to situation, or tasks, usability, urgency; value in use.
- Cognitive match: understanding, novelty, mental effort. Link to previous knowledge.
- Affective match: emotional responses to information, fun, frustration, uncertainty.
- Belief match: personal credence given to information, confidence.