

OmniPaper
Smart Access to European Newspapers

User Interface Requirements

Deliverable D5.1
WP5

Version 1.0 (public version)
28/02/2005



1 Content

1	Content	2
2	Document information	4
3	Version history	4
4	Introduction	5
5	Frame of Reference	5
5.1	Characteristics and features of the system	5
5.2	Related systems	5
6	Users	12
6.1	Targeted User-Groups	12
6.2	Grouping and prioritisation of the users	13
6.3	Relevant user characteristics	13
6.4	Description of professional users	13
6.5	Description of private users	14
6.6	Description of system administration users	15
7	Tasks	16
7.1	WRAT	16
7.1.1	WRAT - Web Requirements Analysis Technique	16
7.1.2	WRAT - Results	16
7.2	Scenarios	17
8	Technical Aspects	19
8.1	User Equipment	19
8.1.1	Current Statistics on Software & Configurations	19
8.1.2	Expected equipment	20
8.1.3	Expected Limitations on the user's side	20
9	Metrics	21
10	Requirements	22



10.1	General.....	22
10.2	Internationality	22
10.3	Introduction to system	23
10.4	Personalization	23
10.5	Hierarchical Navigation.....	24
10.6	Relational Navigation.....	24
10.7	Search	24
10.7.1	Simple Search.....	25
10.7.2	Advanced Search.....	25
10.7.3	Search Results.....	26
10.8	Article Overview.....	27
10.9	View Article.....	27
10.10	Printing.....	28
10.11	Special services	28
10.12	Feedback	28
10.13	Help	29
10.14	Technical Development and Implementation	29
10.15	Performance	30
10.16	Accessibility	31
10.17	Failure Safeness	31
11	Conclusions.....	32



2 Document information

Document ID	OmniPaper_WP5_RequirementsDefinition
Title	User Interface Requirements
Short title	
Workpackage	5
Deliverable	D5.1
Subject	requirements definition
Description	This document describes the requirements for the user interface of the OmniPaper prototype.
Security level	public

3 Version history

Version	Date	Author(s)	Description
0.1	2003-05-19	Cure: Johann Schrammel, Peter Wolkerstorfer	First Draft
0.2	2003-07-07	Cure: Johann Schrammel, Peter Wolkerstorfer	revised version
1.0	2005-02-28	KULRD: Bert Paepen	Revised for public version



4 Introduction

This document describes the user interface requirements for the OmniPaper system. First the frame of reference is explicated and user interface solutions of related systems are analysed and the advantages and disadvantages are identified. Then relevant issues for the user interface design such as users, tasks and contexts are taken care of. At last the resulting requirements are specified.

5 Frame of Reference

In this section the frame of reference of the development of the system should be explicated. The information thereby is founded mainly in two sources: First in the project proposal the general outline of the project was determined. Second, at a consortium meeting in Guimares intensive discussions of the OmniPaper system and its functionality took place.

5.1 *Characteristics and features of the system*

OmniPaper will be developed as a web-platform within the scope of the given timeframe. We are aware that service can also be developed to other directions (e.g. tailored to mobile devices). OmniPaper as Internet Portal will be a starting point, with an online/offline part.

The main feature of the OmniPaper Site generally is to search for articles in various local databases of European news-provider, whatever language they have. (The range of languages and their influence is defined further below). OmniPaper is focusing both on experienced and novice users. Also focused search (the user knows exactly what he wants) and unfocused search (the user doesn't know exactly what he wants) are expected. The system has to satisfy both approaches. It has to help the user to find his own focus.

An important point to consider is the presumption, that on the one hand people don't know how to search and specify their searches (user-group specifications will be defined), but on the other hand, people do not have time to learn the proper use of profiling systems. These capabilities might be too complicated. So the system has to have reduced, simplified search-possibilities.

The system should be able to suggest further themes or keywords. Example for the suggestion of further themes are the internet movie database (www.imdb.com) and Amazon (www.amazon.com). The system should also offer contextual links, in order to help the user to focus his search and to define his interests. This recommender functionality should be able to expect what the user wants just by comparing the user-behaviour, without any personalisation-profiles. Again, it should be deemed to be the principle of minimising any user-efforts.

A special characteristic of the system is, that the user should be able to browse connected topics in different layers. This means ways to browse semantically linked topics have to be developed.

5.2 *Related systems*

Basically, all systems using meta-information are related systems, as they have some important comparable aspects which are meeting the OmniPaper objectives as well. In

this section related system solutions are examined and typical user interface elements and solutions are captured to study them with regard to a potential use in the OmniPaper user interface. The most relevant related systems can be clustered into the following groups: digital libraries, online newspapers, search engines, multilingual portals, display solutions for large networked information spaces, recommender systems and automatic translation services.

a) digital libraries and electronic archives

As OmniPaper is a news archive, knowledge and best practice from digital library system should be taken into account. But there are major differences between library systems and OmniPaper: libraries offer more static content, while news are highly dynamic generated. Additionally the categorisation and processing of the content within OmniPaper will be automated, even the categorisation will change dynamically, new categories will be added.

Examples:

Association for Computing Machinery - Digital library (acm.dl),

Austrian National Library (aleph.onb.ac.at)

IEEE Computer Society Digital Library (www.computer.org/publications/dlib/)

Typical user interface solutions and elements:

- simple and advanced search
- login, user identification
- different sorting possibilities
- preview and detailed view of results
- download possibility
- bookshelf
- help
- different viewing options for results
- possibility to discuss/review article
- classification scheme
- make use of relations (e.g. citations)

b) online newspapers

As OmniPaper not only aims at serving as archive but also to provide news “as they happen”, concepts of online newspapers need to be studied.

Examples:

El Pais (www.elpais.es), Spain

Der Standard (derstandard.at), Austria

International Herald Tribune (www.iht.com), International

De Standaard (www.standaard.be), Belgium

The Times (www.timesonline.co.uk), United Kingdom

Typical user interface solutions and elements:

- overview start page with headlines
- different channels
- different possibilities to sort (e.g. by chronology or relevance)
- external links to source
- conceptual separation of actual news and special columns (weather, ...)
- some articles are especially emphasized (teaser)
- difference to print edition
- download whole newspaper as pdf
- news ticker

c) search engines

To provide the users with most proper search interfaces and functionality the offer of professional search engines shall be analysed.

Examples:

Google (www.google.com)

Alta Vista (www.altavista.com)

Teoma (www.teoma.com)

Typical user interface solutions and elements:

- simple and advanced search
- result list
- offering of general search restrictions (e.g. only results in German language...)
- general preferences/options
- refinement suggestions
- display of estimation of quality of search
- spell check feedback

d) Multilingual portals

An important aspect of OmniPaper is that it is an multilingual system. Therefore experiences from other sides shall be used to avoid making errors that aren't necessary.

Examples:

Home page of the European Union (europa.eu.int)

European Patent Office (www.european-patent-office.org)

United Nations Organisation (www.un.org)

Yahoo! (www.yahoo.com)

Typical user interface solutions and elements:

- permanent possibility to select language vs. single decision
- offering documents to download via different links
- parallel/similar display of same content in multiple languages
- mixed display of content (as it is), but the language is explicitly displayed
- automatic translation

e) recommender systems

As artificial intelligence is used to recommend related articles to the user existing solutions and their interfaces are analysed to be able to develop a highly usable solution.

Examples:

Movielens (www.movielens.org)

Amazon (www.amazon.com)

The Internet Movie Database (www.imdb.com)

Typical user interface solutions and elements:

- provide estimation of how much user will “like” recommended thing
- show highest rated recommendation (by categorie, date, ...)
- feedback possibility if recommendation was correct/liked (implicit or explicit)
- explanation on why article was recommended
- possibility to explicitly tell the system preferences

f) automatic translation services

OmniPaper will offer an option to automatically translate the articles.

Examples:

Babelfish (<http://babelfish.altavista.com>)

FreeTranslation (<http://www.freetranslation.com>)

Typical user interface solutions and elements:

- specification of translation direction
- copy/paste to other applications
- possibilities to deal with keyboard restrictions e.g. enter special characters
- translate whole website
- link to more advanced language resources like Thesauri, dictionaries, ...

g) Display solutions for large networked information spaces

OmniPaper will provide the possibility to navigate through the archive in a semantic structure. As the displaying of networked information is highly complex due to restrictions in screen space different approaches that have been undertaken shall be analysed whether they can be applied to the OmniPaper system or not. Due to the special complexity of the area the solutions are described in more detail than the other related systems.

Examples:

Omnigator (www.ontopia.net)

The Omnigator is a technology showcase and teaching aid designed to help understand the power of topic maps. It is a tool designed to display the information that is stored in topic maps for developers. As the designers specify themselves: “The user interface of the Omnigator is emphatically not what we would recommend in an end-user application!”

NicheWorks (www.amstat.org/publications/jcgs/pdf99/wills.pdf)

NicheWorks is a visualization tool for the investigation of very large graphs. NicheWorks allows the user to examine a variety of node and edge attributes in conjunction with their connectivity information. Categorical, textual and continuous attributes can be explored with a variety of one-way, two-way, and multidimensional views.

Matrix Browser (www.swt.iao.fhg.de)

The approach is based on the use of an interactive adjacency matrix for the representation of relations between concepts forming the nodes of the net. These concepts are either displayed as lists or hierarchies on both axes of the matrix. Different interactive markings in the matrix cells represent the relations between the concepts.

Kartoo (www.kartoo.com)

Kartoo is a meta search engine that displays the results in a cartographic representation of information. The particularity of the system is, that in the map relationships between the results are displayed as well.

Vivisimo (<http://vivisimo.com>)

The Vivisimo Clustering Engine automatically organizes search or database query results into meaningful hierarchical folders on-the-fly, out-of-the-box. It interfaces with any search engine or document database, transforming long lists of search results into categorized information without any clumsy pre-processing of the source documents.

Hyperbolic browser (http://www.ulib.org/webRoot/_hTree)

The hyperbolic browser is a focus+context (fisheye) scheme for visualizing and manipulating large hierarchies. The approach is to lay out the hierarchy uniformly on a hyperbolic plane and map this plane onto a circular display region. The projection onto the disk provides a natural mechanism for assigning more space to a portion of the hierarchy while still embedding it in a much larger context. Change of focus is accomplished by translating the structure on the hyperbolic plane, which allows a smooth transition without compromising the presentation of the context.

Typical user interface solutions and elements:

In general the different approaches to deal with big networked information can be clustered into the following two groups:

- a) approaches that use a graphical representation for displaying relationships between concepts and
- b) approaches, that rely on text to communicate the different relationships.

The big disadvantage of the graphical approach is, that only a limited and small set of items can be displayed at one time, so in most cases only a small clipping of the data can be presented. Also typically a lot of screen space is not used. Further disadvantages are difficulties to display and properly mark the different kind of relationships and that re-zooming needs a lot of calculation and therefore can significantly slow down the process of displaying the information.

Disadvantages of the “textual” approach are, that it is difficult to get an overview on the overall structure of the concepts and relationships between them.

6 Users

6.1 Targeted User-Groups

For a meaningful analysis of the user-groups it is necessary to distinct between the prototype and the fully developed system because there are different user groups to expect respectively aimed at. The kind of users depends highly on the exploitation strategy choices that need to be made within the consortium. Therefore this document describes the UI requirements for the prototype only. It could be possible that the prototype will just be a demonstrator for the OmniPaper technology. If the technology would be sold separately, its customers (and users) would be very different.

In the WRAT-discussion regarding potential users the following typical users were mentioned:

- journalists
- researchers
- press-managers
- communication managers
- top-level secretaries
- secretaries
- knowledge managers
- politicians / users with political background
- local administrations

A final arrangement of this users into user groups didn't take place at the consortium meeting, but the following possibly helpful grouping characteristics were discussed:

- final user/ system administrators / news providers (upload information)
- users who don't know what they are looking for / users who know what
- users that want the system to help them find information / users who want to collect news with the system
- b2b, journalists, public companies
- users willing to pay / users not willing to pay
- centralised usage (one person responsible) / distributed usage (everybody uses it)



6.2 Grouping and prioritisation of the users

The targeted users can be grouped into three main clusters:

- professional users (e.g. journalists): this are users that will use the OmniPaper system related to vocational reasons.
- “private” users: the typical user of this group will use the OmniPaper system related to private reasons.
- system administrators: this are the persons, that maintain the OmniPaper system.

The systems most important user group will be the users that are potentially willing to pay for the service (i.e. professional users) as for the final system a distinction of free content and premium content (users need to pay for) is considered.

6.3 Relevant user characteristics

The following user characteristics were identified as being highly relevant for the design of the user interface :

- language knowledge
- computer experience
- experts / non-experts in information retrieval
- professional background
- age
- awareness of news-market / sources behind

This characteristics affect the following points:

- trust
- expecting serious or entertaining approach
- time sensible or not
- willingness to pay

6.4 Description of professional users

The typical user of this group will use the OmniPaper system related to vocational reasons. We assume that the user will have experiences with related software systems and user interfaces. Considering the international background of the projects the user will be from different countries, have different native languages and the cultural background is varying. In the following the most import user characteristics are specified:



Age: The users typically will be grown-ups. Within this age group no special age group is considered to be of special interest.

Sex: Regarding the sex both man and woman are considered to be about equally distributed.

Profession: The professional background of the typical user will deal with information *recherché* and online publishing.

Education: The educational background of the typical user is considered as having at least a very well basic education. Additionally the users are considered to have extensive experiences in dealing with written information.

Computer experience: The users are considered to have good to very good basic computer skills and experiences with typical office programs such as Word, Excel, PowerPoint and are using the internet regularly and therefore are familiar with browser-based forms and applications. Anyhow, the users are not considered to be experts in the field of programming.

Language skills: The majority of users are considered to speak English as a second language. Anyhow, it is not considered that this is true for all the user of this group. Especially when considering users of the eastern European countries it is likely that some of them don't speak English.

Cultural background: The users of the workbench will have different cultural backgrounds. Typically the user will be grown up in a country of the European Union or Eastern Europe. Even though this will be probably the majority of users, it also has to be considered that people from other continents are interested in European affairs and therefore belong to the targeted user group of OmniPaper.

Use frequency and duration:

The professional users are considered to use the system regularly. A typical frequency might be two to four times a month to perform a more extensive *recherché* and visiting the site irregularly in the meantime. The typical duration of a *recherché* session will probably be 10 to 60 minutes.

6.5 Description of private users

The typical user of this group will use the OmniPaper system related to private reasons. We assume that the users experience with related software systems and user interfaces will be strongly varying. As for the professional users a international background is considered. In the following the most import user characteristics are specified:

Age: The users typically will be grown-ups. Within this age group no special age group is considered to be of special interest. Anyhow, children and old persons shouldn't be excluded, because they also might be interested.

Sex: Regarding the sex both man and woman are considered to be about equally distributed.

Profession: The professional background of the typical user will be of a broad range.



Education: The educational background of the typical user is considered to be distributed on all stages. No special focus can be given.

Computer experience: As mentioned above the experience with computers is expected to be varying. In comparison to the professional users, the computer experience probably is less advanced.

Language skills: The users are not considered to speak English as a second language.

Cultural background: The background of the users is considered to be widely spread.

Use frequency and duration:

The “private” users are considered to use the system irregularly.

6.6 Description of system administration users

The typical user of this group will use the OmniPaper system related to vocational reasons. The administrators will have in-detail knowledge about the system.

Age: The users typically will be grown-ups.

Sex: Regarding the sex both man and woman are considered to be part of this group, but men will probably more present due to the distribution of gender in the group of technicians. Anyhow, this should not effect the system design.

Profession: The professional background of this users will be a technical. Profound knowledge on internet technology is to be expected.

Education: Typically this users will have an technical education.

Computer experience: Experts

Language skills: The users considered to speak English as a second language well to very well. Also technical terms are known by them..

Cultural background: European

Use frequency and duration: The administrative users are considered to use the system very regularly.

7 Tasks

Tasks for the user group of administrators are only partially included in this report as there are still a lot of open questions regarding the system architecture and workflow and the required functionalities are not clear defined by now. Anyhow, the main user interface is not affected by this.

7.1 WRAT

7.1.1 WRAT - Web Requirements Analysis Technique

WRAT is developed for usage within a workshop, ideally with most representatives of a developing team (designers, programmers, marketing, business representatives etc). In such a WRAT session the team is guided through discussions on all relevant aspects of a web development. WRAT incorporates a very detailed set of questions and issues (requirements) which have to be considered in any development. These requirements are combined to “views”, (an issue can be found in more than one view) which each represent a specific perspective on the development (e.g. usage aspects view or organization aspects view). While discussing an issue (e.g. screen resolution) it is possible to regard former experiences and results from other projects (anonymous) and include these into ones’ own considerations and decisions.

During the session the statements and decisions are journalised and projected onto a wall in real time. This enables the participants to easily keep track of the discussion process.

In the post-session phase, the administration interface is used to generate a session report with all discussion points, open issues, decisions etc.

7.1.2 WRAT - Results

A first discussion of task relevant issues took place within the WRAT Session. The following related points were discussed:

- Same users can have different tasks in different time
- Users may have different approaches: user that investigate on a special topic vs. users that want to look on the actual news of the day
- a translation button might be offered (similar to e.g. google-translation-function)
An open question occurring here is, if automated translation is developed enough yet to implement.
- articles should be possible to pick directly for presentation
- maps of what is available instead of endless lists may be useful
- the interfaces shall be user configurable
- There will be users that want to get a packet of news every day on their desktop

- Users will want to get management summaries or abstracts only and not the whole text
- People will also want to watch the layout (original source layout?) and not only the content to get context information such as which articles were placed around, how big was the headline, ...
- Users will want to put news together (for example in a “basket”)
- Users should have the possibility to choose the format in which they want to get the content (pdf, rtf, ...)
- A clipping service is considered to provide users with the relevant news on a push basis
- Users may also want to have access to background infos and related news
- Users might just want to read the news versus users might want to use them for further tasks
- dynamic personalisation (learning aspects, behaviour...) shall be considered

7.2 Scenarios

In this section representative scenarios are compiled to identify and demonstrate typical use patterns, tasks and contexts.

1. Browse a topic

Ken, a new system user, is trying to get oriented to the OmniPaper site and to explore a topic of interest. On the first view he wants to get an overview which information is offered by the site. Then he is looking for information who is behind the system, which newspapers are covered and if he can use the whole system for free or if he has to pay for special services. After he has found this information Ken decides to look for information on the political system of Portugal. Therefore he first navigates through the subject classification and looks , which articles are available. He encounters several ones and also follows some of the presented additional links to related articles.

2. Specific item search

Phillipe has been to the OmniPaper site before and wants to find a specific news article that he believes may be available online. He had read the article in an airplane and now only has a partial memory of the newspapers name and remembers, that the article was about bees. He knows that the article was published in a dutch newspaper and from looking up his flight ticket he knows the exact date. Phillipe therefore first directly navigates to the search section of the Omnipaper system. He chooses the advanced search option and specifies all the information he has about this article.

3. Complete news recherche and ongoing updates for specific topic

Drusilla is a seasoned veteran of OmniPaper and as journalist she's working on a major article. It is important that she is kept up to date and not miss anything that is published on this topic. Drusilla therefore “told” the system the topic she's interested in and that she wants to get notified if a new article covering this topic is added to the system, but only if the article is written in German, French or English because that are



the languages she speaks. When Drusilla wants to read the new articles, she navigates to the personalised part of the system and finds all articles in a special folder.

4. News search by author

Mrs. Susanna has special preferences when reading newspapers: there are some authors she really esteems. She wants to find articles from her preferred authors at once no matter in which paper the article was published. She also wants to be able to tell the system which authors she likes and not be forced to always reenter her preferences.

5. What's new

Andrea is working as a secretary. During her work she likes to have a short look on a news portal to keep informed what's going on. She navigates to the system and is taking a quick look at the most important headlines of the day. Andrea discovers the possibility to personalise her startpage. As she is not interested in sports she tells the system to next time not present her the sports headlines on the start page. Also she specifies that she is specially interested in all articles relevant for Austria. So the next time she will enter the system articles about Austria will be shown preferably in the headlines.



8 Technical Aspects

8.1 User Equipment

8.1.1 Current Statistics on Software & Configurations

To provide an empirical basis for the decision on which software and configurations can be expected on the users side the statistics of <http://www.thecounter.com/> in January 2003 shall be used. This statistics combines the traffic reports from more than 1 million customers worldwide.

Browser and Operating System:

Browser		Operating System	
1. MSIE 6.x	52%	1. Win 98	44%
2. MSIE 5.x	39%	2. Win 2000	27%
3. Netscape 4.x	1%	3. Win ME	16%
4. Netscape 5.x	1%	4. Win NT	3%
5. MSIE 4.x	1%	5. Win 95	2%
6. Netscape comp.	1%	6. Mac	2%
7. Unknown	0%	7. Unknown	1%

Screen Resolution and ColorDepth

Resolution		ColorDepth	
800x600	46%	32bit	3%
1024x768	40%	16bit (65K)	43%
1280x1024	4%	24bit (16M)	9%
1152x864	3%	8bit (256 Byte)	3%
640x480	2%	Unknown	0%
Unknown	1%		

JavaScript and Java

Javascript		Java	
Javascript 1.2+:	89%	Java enabled:	88%
Javascript <1.2:	0%	Java disabled:	0%
Javascript false:	10%	Java unknown:	10%

8.1.2 Expected equipment

In general the expected equipment is a standard desktop computer equipped with mouse, keyboard and colour screen.

To ensure most usefulness it is decided that the system should work with even old browsers (version 4). It is expected that the average user will have a colour screen.

8.1.3 Expected Limitations on the user's side

One very serious limitation will be the users different keyboards. This is especially problematic for entering searches.

Further possible problems can be the display of topic maps. As we must take into account that a lot of users will not access the system via a broadband connection it has to be ensured that all functionalities can access fast without difficult graphical displaying. Especially the visualisation of topic maps could be a problem.

Font limitations could raise a problem, but as we are just talking about European languages, it is considered to be a minor problem. The same uniform platform will be developed for each language.

Further limitations in the area of charset have to be considered (ansi/ascii-problems).

Also limitations in speed of internet access have to be considered (2002: 25 million broadband users worldwide = 5% of total Internet user population, Source: Leichtman Research and Parks Associates, Aug. 2002).



9 Metrics

In this section relevant metrics are estimated. This is due to the fact that this numbers influence the best solution for the interface. For example to search within a population of 100 entities requires different mechanisms than a search within a huge population of millions of entities.

<i>Metric</i>	<i>Prototype</i>
<i>Portal Usage</i>	
Average (and peak) number of concurrent users	100 (1000)
Average (and peak) number of searches per hour	100 (1000)
Average (and peak) number of non-search page-hits per hour	100 (1000)
<i>Administration</i>	
How many people will be creating/checking metadata	10-30
How often will articles be updated	at least once to twice/day, there also will be archives that update more frequently (as soon as new articles are available)
How many articles will we end up with after a year	300.000 - 600.000

10 Requirements

Based on the defined user groups, scenarios and the WRAT-session the requirements for the system are analysed in detail in this section.

10.1 General

REQ #1. Easy to remember

Even after some time of not using the system it must be easy to remember how to control the system.

REQ #2. Support experts and novices

The user interface shall support “expert user” as well as novice users - interaction styles.

REQ #3. Continuous working should be supported

E.g. User collects something on Monday. When he/she logs in again on Thursday, the session should be continued there.

REQ #4. Enable learning transfer

The user interface should enable to transfer already learned experiences from other interfaces in the same area.

REQ #5. Easily readable

As the system offers lots of texts it is required, that this text is easily readable.

REQ #6. Concise wording

The User Interface wording should be easy to understand even for non-native speakers.

REQ #7. Graphic support

The wording should be possibly supported by meaningful graphical hints to ease the understanding.

REQ #8. Transparency

The users shall always be informed what happens to the data he/she submits.

10.2 Internationality

REQ #9. Multiple language support

The user shall be able to use the system in their own language. The prototype must support the following languages:



- o catalan
- o spanish
- o german
- o french
- o dutch
- o english
- o portuguese

REQ #10. International way of presenting the information.

As the system aims at an international target group according presentation styles have to be developed.

REQ #11. Cultural-specific designs have to be avoided.

Design solutions that refer to a special cultural background should not be used to ensure comprehensibility for all users.

REQ #12. Local characteristics have to be supported

The system must take specific local characteristics into account. For example character set, date and time formats, ..

10.3 Introduction to system

REQ #13. Comprehensive introduction to OmniPaper

The first time users will expect to get an overview of what is provided by OMNIPAPER on the first page. Newcomers want to be informed what OMNIPAPER is all about. If that information is not satisfactory for the user, he/she will probably leave and never come back again!

REQ #14. Useful start-up page

The problem with most “portal” pages is that they are much too overloaded (“over-newsed but under-informed”). Typical portals are much too difficult to read for typical users. OmniPaper should provide a lean page which offers the relevant categories in sub-pages.

10.4 Personalization

REQ #15. Allow for personalization

Users will feel better when they can adjust some “bells and whistles”, like some colours and icons or a very personal welcome greeting.



10.5 Hierarchical Navigation

REQ #16. Transparent navigation

The user should be aware of the tree structure. The structure should always be present to the user.

REQ #17. Customised information browsing

When browsing, users should see only the information relevant to their needs. E.g. someone searching for news in Greece does not want to see news specific to Germany. Users must however also know that there is some other content which is filtered. We must not hide content from the user!

REQ #18. Suggest content relevant to the user

When the user simply navigates the content, it is also possible to display to him the content most suitable for him.

REQ #19. User controlled display

Users will have the ability to collapse and expand individual nodes. There should also be a 'collapse all'/'expand all' function.

REQ #20. Limit number of elements.

The hierarchical structure shall not have more than 10 entries on a level.

REQ #21. Feedback

The system must provide clear feedback regarding the actual position of the user within the hierarchy.

10.6 Relational Navigation

REQ #22. Communicate semantic relations

The system needs to be able to communicate the different semantic relations between concepts (e.g. broader terms, narrower terms, other related terms) to the user in an easy-to-understand way.

REQ #23. Reduced complexity

The system needs to find ways to reduce the information displayed to the user.

10.7 Search

REQ #24. Proper search mechanisms must be provided.

Search mechanisms could be implemented in several ways. A simple search should be available throughout the site. Advanced search should give the user totally unlimited freedom in combining search criteria.



10.7.1 Simple Search

REQ #25. Simple search

Simple search functionality shall be available on every page.

REQ #26. Boolean operators

The search will process Boolean operators OR, AND, and NOT embedded in a free text statement.

REQ #27. Phrase searching

The search should provide phrase searching, putting quotes around a set of words to only find results that match the words in that exact sequence.

REQ #28. Required and prohibited terms

The search should accept required and prohibited terms, to either require or prohibit words from appearing in the search results process. A generally accepted notation will be used (e.g. + and -).

REQ #29. Stemming

The search will support stemming, where the exact ending of a word (plurals, tenses, etc) does not affect its ability to be found.

REQ #30. Wildcards

The Omnipaper search query language shall support the use of the wildcards '*' ('match all'/'match the rest') and '?' ('match single character') when the according news archive supports this function. Note: The use of wildcards can make searches slow and inefficient. The user should be aware of the implications when using wildcard searches.

REQ #31. Case-Insensitiveness.

The search shall be case-insensitive.

10.7.2 Advanced Search

REQ #32. Advanced Search

Advanced search mechanisms that allow to specify the query in more detail shall be offered. The search engine will accept targeted queries, enabling search terms to target specific metadata elements.

REQ #33. Targeted search criteria

The search engine will accept targeted queries, enabling search terms to target specific metadata elements. The advanced search shall allow to specify search variables in the following elements:



author
timeframe
title
language
publisher, newspaper (source)
abstract
subject classification
keywords

REQ #34. Combination of targeted criteria

The targeted query should enable targeted elements to be used in Boolean combination with each other.

REQ #35. Targeted search functionality

A targeted element should be searchable with the full range of search functionality defined in the simple search section.

10.7.3 Search Results

REQ #36. Display of search results

Search results shall be displayed in the same design as article list in the other views, but additional information and functionalities shall be provided. Search results will be capable of including part of the content of discovered records, to enable informative result displays to be constructed.

REQ #37. Display search query

The entered query must be shown to the user on every results page.

REQ #38. Provide suggestions for refining query

The system shall provide meaningful suggestions for refining the query.

REQ #39. Spell checking

The spelling of the query shall be checked and suggestions for correction if required shall be provided. In the prototype this function only shall be applied, when free language resources can be used.

REQ #40. Number of search results

The search results will indicate the number of records found by the search.

REQ #41. Number of displayed results per page



A mechanism will be provided to enable the user, or portal, to customise the number of search results retrieved for a single page.

REQ #42. Sorting

Possibilities to change the sorting shall be provided; the relevant sorting criteria are relevance, date and language

REQ #43. Search within results

The possibility to search within the results only shall be provided.

10.8 Article Overview

This section deals with the display of an overview article list. Whenever it is required to display such a list the following possibilities shall be provided:

REQ #44. display meta data

Date and time of publishing/update and the source (newspaper, ...) needs to be displayed to provide the context of the article.

REQ #45. short summary/abstract

A short abstract shall be shown to the user to allow for deciding, if the article is of relevance to him.

REQ #46. display links

Links to related articles shall be displayed to

REQ #47. Sorting

Possibilities to sort/filter the articles by relevance, chronology and region are needed.

10.9 View Article

Whenever an article is displayed in detailed view the following requirements shall be taken into account.

REQ #48. mail article

mail article to e-mail address

REQ #49. download

A possibility to download an article in different formats (txt, pdf) is required.

REQ #50. links to related articles

Links to related articles shall be displayed.

**REQ #51.** links to related topics

Links to related topics shall be displayed.

10.10 Printing**REQ #52.** All pages must be easily printable.

A special page version optimised for printing shall be provided for every page.

10.11 Special services**REQ #53.** News Collections

The possibility to collect articles of special interest in a user-specific basket shall be provided.

REQ #54. Clipping/Push services

The users shall be able to order articles of special interest to be delivered to them by mail.

REQ #55. Notification of news update

The user shall have the opportunity to get notified, when a new article matching specified criteria is available in the system.

10.12 Feedback**REQ #56.** The users shall be able to provide feedback on relevant aspects.

To ensure the quality of the system the users shall be provided with the possibility to feedback relevant statements. A general purpose feedback form is required for users to communicate with the Portal Business Group about any issue they choose. This form could include the following fields:

- The user's contact information
- Classification:
 - general feedback;
 - ask questions;
 - report problems.
- Subject
- Required Action:



- none
- answer
- other:

10.13 Help

REQ #57. All user documentation shall be provided online.

The following features are required:

- web-based access in a context-sensitive manner;
- an overview/introduction explaining the system functionalities to novice users;
- every screen, form and function should have a detailed help entry describing how it is used and for what purpose
- a Frequently Asked Questions section (FAQ) shall be provided;
- contact details for obtaining additional information not available online are required.

10.14 Technical Development and Implementation

REQ #58. easy to add new language

In the development of the system it has to be taken care that a new language can be added easily.

REQ #59. Support logging of user behaviour

For the OmniPaper-project, it will be relevant to detect patterns of clicking and learn user requirements in the broadest sense out of that to develop recommendations. Therefore only techniques that allow this logging shall be used.

REQ #60. server-side technologies

Required server-side technologies operation of the system are:

- topic map engine
- html database
- efficient query metadata,
- jsp for creating personalised webpages (or asp),
- webserver,
- scripting engine,



- metadata engine + database,
- firewall,
- redundancy system,

REQ #61. development environment

An open source development is considered to be important. In order to allow further content partners to connect without much efforts, a very simple approach is necessary. Soap is distributable to all platforms.

REQ #62. multi-platform support

The system must support as much platform as possible.

REQ #63. browsers

The system must work without problems on old browser versions (IE v4, Netscape v4).

REQ #64. Screen resolution

The system shall be able to use easily with a 800x600 screen.

REQ #65. Color depth

The system shall be easily usable with 16 bit color depth.

REQ #66. Java and JavaScript

Java and javaScript may be used within the system. Anyhow, all basic functions must be accessible without this features.

10.15 Performance

The OmniPaper system is architecturally designed as a metadata management system with mainly two operation modes: meta mode and direct mode. In the meta mode, the OmniPaper user accesses meta information that is managed and provided by the OmniPaper system and OmniPaper databases, completely detached from the existing systems. In the direct mode, extended queries are performed on the metadata databases at OmniPaper and directly on the existing news archives by utilizing the basic queries on demand.

REQ #67. Response time for meta mode

For the meta mode the a response time shall be below four seconds for 95% of the queries.

REQ #68. Response time for direct mode

For the direct mode we aim to reach a response time below ten seconds for 90% of the queries.



10.16 Accessibility

REQ #69. Location independent

Accessing data must be possible independent from the location of the user. Users should be able to access the system in a unified way no matter where they are. If a user changes the location (office workstation to notebook at a hotel) the user still wants to have the same desktop environment even if the user changes the operating system. Differences may only appear due to device dependencies (e.g screen size).

REQ #70. WAI-Standards

The system shall be designed and programmed according to the "Web Content Accessibility Guidelines 1.0" published by the Web Accessibility Initiative.

10.17 Failure Safeness

Errors in the system may and will occur. It is not desired that such errors can influence the system in a way that it crashes or that it remains in an inconsistent state.

REQ #71. Errors

It should be avoided that errors which occur in parts of the systems influence the whole system in a way that it crashes or is in an inconsistent state.

REQ #72. Failure of the underlying operating system

The underlying operating system may and will crash or break down. It has to be possible to restore a consistent state of the system after a crash of the operating system.



11 Conclusions

This document identifies the requirements for the OmniPaper prototype. The description of the frame of reference showed, that OmniPaper will be a web based system with advanced search possibilities. As main users of the system three user groups were identified: journalists, "privat" users and system administrators. The resulting requirements reflect the main approach: the system shall be easily accessible, standard interface mechanisms as well as advanced and innovate ways of presenting information shall be used and the multicultural target group has to be considered specially.