Electronic Publishing '01 –
2001 in the Digital Publishing Odyssey
A. Hübler et al. (Eds.)
IOS Press, 2001

Personalized Information Services – an Electronic Information Commodity and its Production

Dipl.-Inform. Thomas Ritz
Fraunhofer IAO
Nobelstr. 12
70 569 Stuttgart, Germany
Thomas.Ritz@iao.fhg.de

Abstract

Personalization has become a buzzword in the last months when talking about the internet and the World Wide Web. Nowadays, a lot of sites promise personalization features, but the services provided differ a lot.

Most of the recent publications utilize personalization as a marketing tool, which helps to address customers in a one-to-one marketing manner. Apart from this, personalization and more personalized information services can be seen as unique information commodities supplied by content providers. In order to maximize benefits, the production tools have to be perfectly integrated into the production process applied for the production of recent mass publications. In this paper, we present a definition and characterization of personalization and resulting information goods and sketch an integrated production framework applicable mainly for publishers to provide personalized information services as a unique product in their portfolio. Finally, the paper discusses usability aspects of personalized information services.

0. Introduction

The broad availability of myriad of information and information sources for consumers and the industry, such as the web, has raised the expectations towards services from information suppliers to new heights. To compete against free online services, value added services need to be developed, for which customers are willing to pay a premium fee. The media industry as the classical content owners have to be considered as key players in this development. This has two main reasons:

1. The already existing customer binding and customer's trust in quality
2. The interest of publisher in raising revenues by reuse of contents

Today users are looking for high quality information which they expect from publishers (NN, 1999). Accordingly, the users trust in the skills of editors, likely known, e.g. from the daily news, to filter information regarding customer's needs (c.f. (Ryan, 1999)). This is fostered by Hess (Hess, 1999), stating that in 1998 10% of all Internet users refer to contents provided by media companies.

Already in 1997, Middelhoff (Middelhoff, 1997) predicted the raising importance of contents as key success factor for the media industry. Due to the fact that the elaboration of contents (editorial work) is very expensive, many others stretch the importance to reuse contents (i.e. (NN, 1999)). Thus, the economies of scale holds true also in the publishing industry (c.f. (Shapiro/Varian, 1998) or (Ryan, 1999)).

But how to create new products out of the existing contents? Berghel (Berghel, 1999) argues that it is a misunderstanding to see quantity and variety as success factors for

electronic publishing. Moreover, Sander (Sander/Kröger, 1998) postulates an enlargement of the magic triangle, known from project management, towards a "magic rectangular" comprising Cost, Time, Quality and, and what's new, Quantity as major success factors. Facing these success factors, Palmer (Palmer/Eriksen, 1999) gives first hints on activities creating real added value and which kind of publisher will benefit:

table 1: value added publishing

Activity	Type of Publisher Benefited	Technical Requirements
Customized Product	All	Agents, "push" technology
Subscription	All	Micro-payment system
Market intermediary	Publishers specialized on topic or region	Customizable use profiles
Enhanced distribution	Regional specialization	Subscriber database, "push" technology
Archive	Depends on content, role in marketplace as authoritative sources	Search capability, micro-payment scheme

Accordingly, Berghel (Berghel, 1999) postulates following activities to come to an "Value-Added-Publishing", using the Internet far beyond a new distribution channel:

- content enhancement
- encouragement of synergy between and among information providers, consumers, and the resources they share
- addition of interactivity and feedback loops to traditional delivery systems
- reorientation of both the information provider and information consumer toward, "process" of publishing rather than a focus on individual products and services
- meta-level analyses and intelligent restructuring of document collections
- Ad hoc document quality ranking and recommending systems

Most of these activities may be addressed by providing personalized information services.

1. Personalised Information services

1.1 Definition

As stated before, lots of services are labelled personalised service. A definition based on common characteristics will result in a generic definition like "Personalization is a service provided based on a user profile". Which individual services generated using these profiles is still uncertain.

Allen as well as Nielsen (sees (Allen, C. 2000) and (Nielsen, J. 2000)) attempt to distinguish personalization and customisation, while looking at the level of control granted to the customer of the service. Nielsen (Nielsen, J. 2000) defines customisation to be "under direct user control" whereas personalization is "driven by the computer" and based "on some form of model of that user's needs". I.e. Bell (c.f. (Bell/Moffat, 1996)) presents such a formal approach, defining vectors that describe the user profile, the document collections and mappings between these measures.

Another definition is given by IBM (IBM Developer Works 2000): *"Personalization* is a *process* of *gathering* and *storing* information about visitors, *analysing* the information, and, based on the analysis, *delivering* the right information to each visitor at the right time". This definition reveals already the process character behind personalization. This makes personalization an information service, matching classical definitions of a service

(c.f. i.e. (Hermes)), what we will call in the following *Personalised Information Service*. But there is still the open question of what might be the right information.

To tackle this problem, Bender ((Bender, 1996)) defines *"Newsworthiness"* as that which is important for the individual and the community (the relation between individual and community interest will be investigated in detail later on). In addition to this, a new dimension appears nowadays. Relying on the paradigm of media independent publishing, it became possible to serve several media from one information source (see i.e. (Matzer, 2000) or (Glotz, P. 2000)). Thus, new information and communication demands are stimulated, making the appropriate media important for profiling users' interests. Sakagami (Sakagami et al., 1998) intends to state more precisely the user interest, focusing on "information freshness". Therefore, he proposes that the "freshness" of information is not only reflecting on the dispatch of information. Furthermore, it has to contain whether the consumer already reads it (for formal models see (Sakagami et al., 1998)).

The definition as follows tries to integrate the dimensions found in definitions cited before:

- process
- user interest (textual)
- demand in schedule
- demand in media

Def. 1: A *Personalised Information Service* is a service towards a customer comprising

1. *filtering* of information out of former *gathered* and *qualified* information regarding users textual *interest*
2. *presentation* of this information using a user defined *time schedule* and *media* appropriate with recent user environment.

As seen before the term interest can be used quite differently. First of all it can focus on the content of the information. But it can also reflect on a time context as shown by the information freshness approach (c.f. (Sakagami et al., 1998)). In real life these preferences are complex structures formed by weighted interest resulting in a profile or preference structure. This shows the similarity to classical economical approaches towards preferences and use (c.f. (Varian, 1991)), leading to an abstract model of interest.

The *media* focuses on the communication channel, which is used to disseminate and present the information for the consumer. Certainly every media creates constraints towards the presentation of the former filtered information (i.e. audio, video capabilities). But also the users' environment can determine the appropriateness. But in the end it is a user preference or a result of user preference and thus part of the user profile.

1.2 Typologies

It was shown that personalized information services are hard to define and a variety of recent existing services match the former given definition. To discuss personalization in more particulars, a categorization of personalization services is needed.

Especially if an IT system learns the user profile from the user behavior, the personalization is not a one step process. In contrary, this process is continued. The gradual process of personalization is sometimes called "creeping personalization" (see (Graham, J. 2000)). Kania (Kania, D. 2000) refers to different models of personalization:

- *Personalized web experience* -- The user's experience itself can be personalized on the fly based on the user profile and their real-time click stream. Basically, these web pages, or parts of pages, are highly dynamic and change during the user's experience.
- *Personalized Information* – Each one of your customers can create their own special and unique version of your web site. Or you can provide personalized recommendations based on a user's unique profile. (This has been the most common form of personalization.)
- *Personalized Service* -- At an additional level of sophistication, you can connect users with personalized customer service for both sales and supporting the customer. These systems can be totally automated or can be a hybrid system that also involves human beings. Dell Computer's order tracking system is an example of an automated personalized service system. An example of a hybrid system is connecting the buyer with an expert customer representative who can make highly personalized recommendations, which are facilitated by the web and/or e-mail.
- *Personalized Community* -- Targeted online communities or discussion forums based on particular interests or needs. These can be built for both business-to-consumer and business-to-business markets. For example, if you are a clothing retailer, you can have several targeted discussion forums within the women's clothing section based on several categories including work formal, work casual, casual, formal/party, accessories, shoes, etc.
- *Personalized Cross-Selling/Up-selling* -- Presenting unique recommendations as users are adding items to their shopping cart, or in an "after-marketing" fashion when a bill is presented or in future e-mail or direct mail communications.
- *Personalized Auctions* -- It is an optimum personalization service where customers get the right product at the right price.

table 2: personalization technologies

Personalization Technologies			
Push by Agent		Pull by User	
Site Control	User Control	Site Control	User Control
Rule Based Matching	Matching Agents	Attribute Search	Menu Select
Feedback & Learning	Community Rating	Full Text Search	Playback
	Collaborative Filtering		

Also Luedi (Luedi, 1997) gives a categorisation of personalization. First he distinguishes push respective pull delivery. Using push *delivery,* the contents are sent actively to the users. In contrast, the users have to take the active part if *pull delivery is* applied. Distinguishing for both methods whether the user defines the filtering methodologies or the site brings Luedi to the following categories of personalization:

- *rule-based-matching:* according predefined rules (i.e. customer clusters) the end user is provided with information
- *matching agents:* an agent searches for information on behalf of the user
- *feedback and learning:* the information system learns from user behaviour and provides information regarding these user profiles
- *attribute search:* the user can search by category (i.e. by industry)

- *full-text-search:* in contrast to search by attributes, the user can search for keywords in the full text (instead of searching for semantic entities like author etc.)
- *community rating:* the quality of information is ranked by the users and others present the information to users regarding the rankings made.
- *collaborative filtering:* as extension to feedback and learning, the learning mechanisms make use of the behaviour of other users (likely with similar profiles), too.
- *menu select:* the users navigate through the information system based on a predefined structure menu.
- *playback:* in order to raise customer binding to a information service, all information regarding the user is used to contact him (i.e. he is welcomed by name when entering a web site).

1.3 Samples

Already in 1996, Bender (Bender et al., 1996) describes sample services performed at the MIT, meeting the former elaborated definition of personalised information services. In the meantime, numerous services emerged. Within the framework of the former mentioned PEACH project a best practice analysis was covering more than 600 web services. All services were evaluated whether the user can tailor one or more of the following items:

- *Topics:* The user can define areas of interest. The information is filtered accordingly to the defined areas. The investigation revealed that there are lots of ways to define the interest profile and to present the information regarding the profile (see results).
- *Layout:* The user selects the media for delivery. This comprises in most cases Web, email, SMS, WAP, PDA or FaxOnDemand services.
- *Sources:* The end user can define the information sources (i.e. newspapers) which should be searched for information. This is important especially for meta search or meta filter engines.
- *Bookmarking:* The end user can use a bookmarking service. This can either mean that the user can bookmark certain areas within one web site or that the end user can manage several web information sources.
- *Quantity:* The end user can define the quantity of information, which should be delivered to him.
- *Times of delivery:* The end user can define a time schedule for the delivery

The study revealed that most services allow the tailoring of the layout and the areas of interest. Surprisingly, the number of services offered for mobile devices as mobile phones or hand held devices was relatively small. The study came up with huge differences regarding the usability of the services provided. For our development we will rely on the best practices found. Concerning the business model applied for the provision of the personalised information service most publishers seemed to be conservative. In most service banner ads are placed and very few services applied business models as pay-per-view.

Another interesting aspect was brought up when analysing methods for the acquisition of user profiles. The study resulted in a categorisation of these methods:

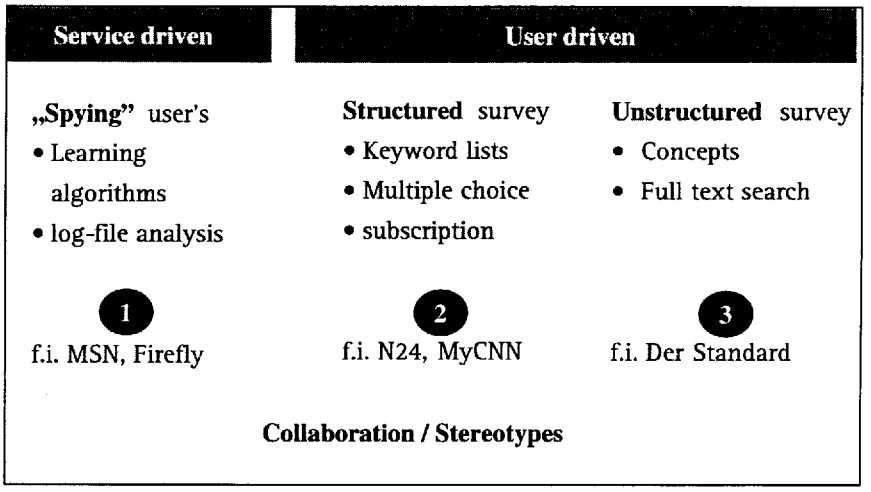

figure 1: three ways to acquire a user profile

2. The production of personalised information services

We assume a generic value chain for the production of personalised information services comprising content acquisition, content processing, content composition, and delivery to

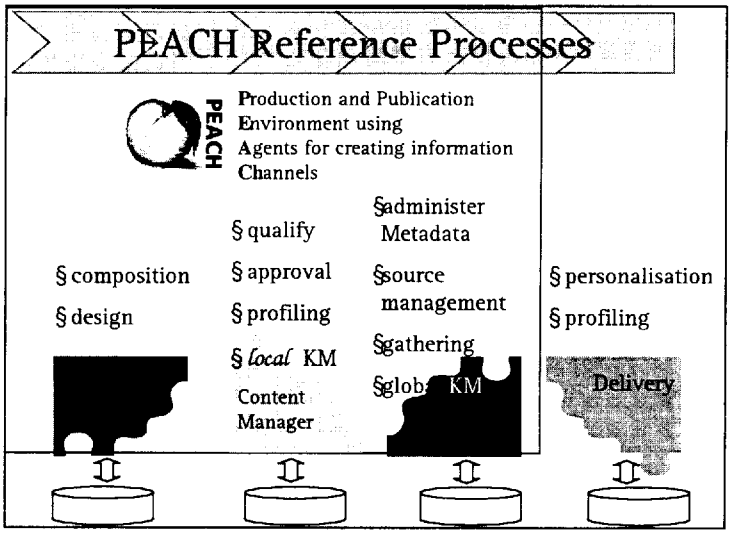

figure 2: Covering the complete value chain for production and provision of personalised information services

the end user. To produce and provide a personalised information service, manifold applications are used today. But from our perspective, there is no tool available covering the complete value chain. Thus, an integration is needed (c.f. (Neuhold, 1998)). As

mentioned before, there are certain key-technologies available today covering parts of the value chain:

To support the content acquisition, we see functionality provided by content management systems. This functionality can be improved by services granted by tools used for information retrieval. The use of (collaborative) agents might help to make the information process more effective. Standards and means used for Digital Libraries and Bibliographic Databases might also deal with fruitful inputs. Editorial tools are well

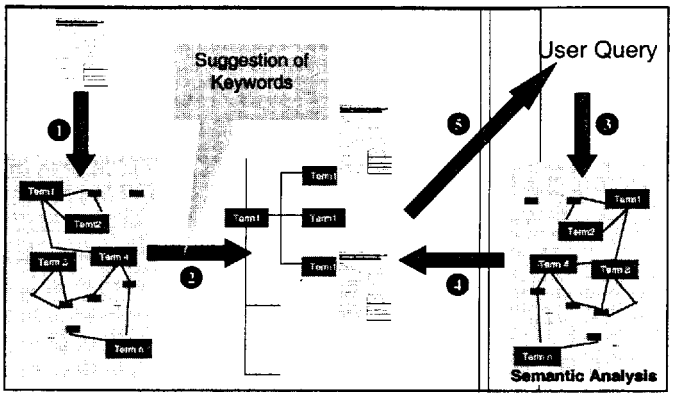

figure 3: Content Qualification and generating user profiles based on semantic nets

developed and perfectly integrated in the production processes of publishers. Therefore, a tool can rely on these services in order to support Content Processing and Composition. Functionality known from Internet portals can be used to transfer the personalised information towards the end user. Therefore, the portal has to be integrated with the content management solution in place.

Further impacts may arise by employing an appropriate business model for placing personalised information services. A survey amongst publishers revealed that this is one of the major concerns of the publishers. Further investigations showed that the publishers would prefer mostly business models close to the business model employed for the recent print products. Thus, ads are mentioned as an important business model and recent researches focus on improving these business models while sticking to the processes used since decades inside the publishing companies.

Within the European funded project PEACH (http://www.vis.iao.fhg.de/peach) we try to attempt this integration and to test the resulting application in a professional test bed. Therefore the PEACH system combines already available systems, defines the integration methods and specifies functionality, which have to be developed from the scratch. We see semantic classification and semantic know how as glue between the different applications involved. In order to deal with personalized information services a user model is crucial for the PEACH delivery system. The same time the semantic know how could be used to define the user profile. The users enter only a vague description of his interests. The semantic know how maps this to the classification scheme (keywords) already applied for classifying the documents. Thus the process could be summarized as shown in *figure 3.*

1. Qualification of contents during editorial process making use of semantically analysis
2. As a result of this qualification, the contents are assigned to a keyword tree

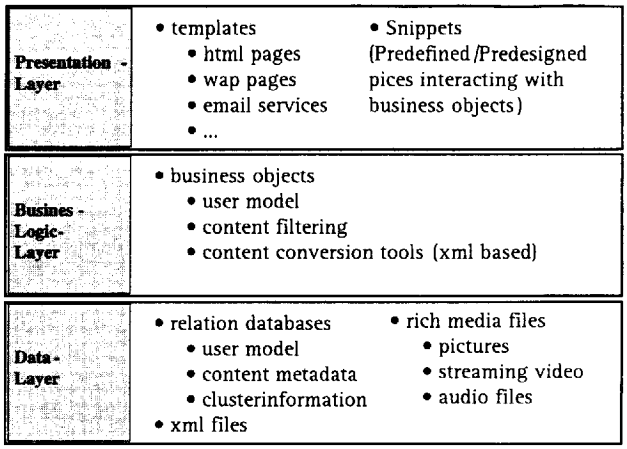

figure 4: The 3-Tier-Architecture

3. User formulates his research interest in his own words. With the help of semantic classification the user query is mapped to the same keyword tree the contents are assigned to these contents are presented to the user, matching his articulated areas of interest (5).

The PEACH delivery system is structured following a 3-Tier approach (c.f. *figure 4*). This approach ensures that the business logic components could be reused for several media, by handling the particular presentation in the presentation layer. The data layer comprises the atomic information. Mostly relational databases will be employed to handle well-structured information as the user model, metadata information etc. Additionally to this, some either unstructured or dynamic changing structured information has to be handled. For text based information (as the text and the structure of articles) we rely on XML structures based on NewsML (c.f. http://www.iptc.org/NewsML). This XML structures could be shown to the

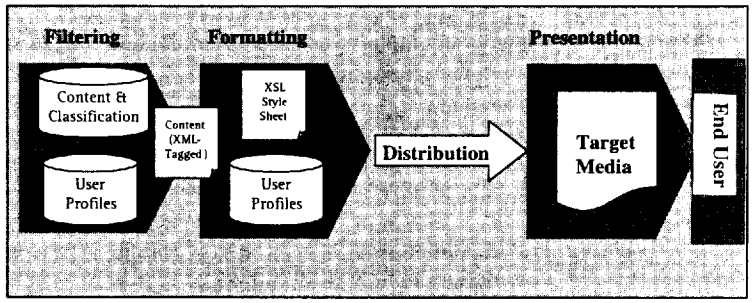

figure 5: The overall process of selecting and formatting contents could be shown as follows

end user by emptying XSL Stylesheets, which filter the information needed and arrange them as defined by the publisher. Furthermore the XML contents may include external entities pointing to flat files as streaming media, pictures and audio files. Again the meta data describing these files are handled employing a relational database system.

The business objects are designed to deal with basic functionality which can be easily embedded in the presentation layer and which encapsulates the data layer from direct access. This is extremely important for critical functions (i.e. dealing with the user profile), cause business rules i.e. for handling privacy issues have to be defined and validated once.

The presentation layer has to deal with interfaces for the different targeted media for information distribution. This layer have to be defined in a way granting a great level of flexibility to the designer of personalized information services. Therefore we came over to design templates for the different distribution media, defining the structure and the look and feel of the service. For designing templates i.e. Macromedias Dreamweaver could be used. To deal with a more modular structure the PEACH system will be shipped with a set of predefined snippets covering the most needed functions and which could be tailored (i.e. designed) regarding the publisher particular needs and requirements.

3. How to Master Interface Design for Personalised Information Services

3.1 Personalization is not the actual goal task users want to perform

The intention of personalised information services is to provide users with individually tailored news and information. It is to optimally support them retrieving information within their scope of interest in the desired coverage, access time, and media. Personalization is also about making online information easier to access, more efficient in use, and providing an individual, personal online experience for the users. Thereby personalization can be considered twofold: as a feature that supports the information supply of a particular information provider, or as an autonomous concept which is to grab and filter information and contents from several information sources and individually present it to the users. In either case the superordinated goal is to help the users getting the information they want in an efficient and pleasing way. Yet the hurdle with achieving this goal is that the information system must first know the users' interests, needs, and preferences. Many of the existing personalization services put the strain of gathering this information on the users. Users are confronted with incomprehensive registration forms, prompting the users to explain their life story in data. Next they have to bother with seemingly endless personalization and customisation features, having to choose from dozens of check boxes or enter cryptic search and keyword strings. The result is that many users either give up before having finished the personalization process or they end up being frustrated because the effective results often stay behind the expectations of the users. Such systems seem to ignore that personalization is not the actual goal task of the users but rather just getting the information they are interested in.

Therefore, personalization services must be easy to set up and use in order to find widespread acceptance with the users. The central unit that is to focus on is the user interface, the visual part of the application with which the users have deal. No matter how complex technical processes are behind an application the user interface is to hide this complexity from the users and to 'translate' pure technical cohesion into a 'language' users can understand. Users don't want to operate an application – they just want to use it. The keyword here is Usability which is determined by the degree of effectiveness, efficiency, and personal satisfaction the users achieve when using any software or interactive application (c.f. (Nielsen, 1999)). Applied to personalised information services that is, effectiveness is ensured by optimal results of the filtering process, efficiency is ensured by requiring the users to perform as less as possible steps to reach their goal, and satisfaction is ensured by making the use of a personalised information service a pleasing experience.

3.2 Characteristics of Functioning Personalization Services

The question is what are the basic characteristics of the user interfaces, which fulfil the demand for effectiveness, efficiency, and satisfaction with personalised information services?

- **They are inviting, convincing novice users from the benefits of the personalization service.** PIS catch the users' attention and immediately reveal their benefits to the users as it is only a few moments that determine whether new users will consider a service valuable to further deal with or rather head on for the next one – which is always just a mouse-click away.
- **They convey credibility and trustworthiness.** Users get more and more concerned about the usage of their private data they are usually asked for when registering with a service. This includes that privacy statements are easy to understand. It gives users the feeling of being taken seriously and treated respectfully.
- **They are easy to set up.** People use personalised services because they want to make getting information easier, avoiding obstacles encounter them when dealing with several news and information resources, and thus save time and intellectual strain. Complex dialog processes for creating personal user profiles and customised interfaces strongly disagree to this claim of the users and make the entire concept appear not authentic.
- **They are accessible via different media.** To optimally fit to the users needs and interests personalised information services are accessible via different media such as cellular phone and personal digital assistant (PDA). At this, operating hand-held devices requires minimum user interactivity as all user preferences are set up by a PC workstation, which provides enough bandwidth as well as sufficient input and display facilities.
- **They are media-adequate.** User-friendly PIS take advantage of the interactive medium hypertext/hypermedia, that is several media such as text, video, and audio support the information where appropriate. Further, they support content-enriching features such as linking to additional sources that provide in-depth information or that correspond to the particular information topic, or support user-feedback, rate information quality and relevance for other users.
- **They are efficient in use.** Contents and information is clearly structured so that user can get their bearings and easily digest the provided information.
- **They are appealing and pleasing.**

4. Conclusions

The presented text showed that personalised information services have to be considered as an information commodity which creates added value for the provider as well as for the information seeker. With reference to the definition presented for personalised information services it became clear that publishers are in an ideal position to provide these kind of services. Nevertheless, it came up that the production of these services follow today an software engineering approach, cause there is no tool existing, which hooks into the well established value chain for the production of information goods. The paper proposed such a tool, developed by integrating several tools, which are already applied in media industry. As glue between these application semantic analysis was proposed. Finally the paper stretched that high level usability of the personalised service is crucial for acceptance by the consumer and is thus an important precondition for all business models applicable for personalised information services.

5. References

Allen, Cliff *Personalization vs. Customization*. 2000 http://clickz.com/cgi-bin/gt/en/pm/pm.html
Bell, Timothy A.H. u. Moffat, Alistair. 1996. *The Design of a High Performance Information Filtering System*. Zürich:
Bender, W.; Chesnais; Elo, S.; Shaw, A. u. Shaw, M. 1996. "Enriching Communities: Harbingers of news in the future." *IBM Systems Journal* 35 (4&5):
Bender, W. 1996. "Community and personalization." *IBM Systems Journal* 35 (34): 367-368.
Berghel, Hal. 1999. "Value-Added Publishing." *Communications of the ACM* 42 (1): 19-23.
Glotz, Peter *Die Zukunft der Zeitung zwischen WAP und UMTS*. 2000 http://www.spiegel.de/netzwelt/medien/nf/0,1518,74766,00.html
Graham, Jeffrey *Creeping Personalization 101*. 2000 http://www.clickz.com/cgi-bin/gt/cz/ems/ems.html?artcile=1087
Hermes, Peter.*Entwicklung eines Customer-Self-Service Systems im Technischen Kundendienst des Maschinenbaus*. Universität Stuttgart
Hess, Thomas. 1999. "Das Internet als Distributionskanal für die Medienindustrie - Entwicklungstendenzen im deutschen Markt." *Wirtschaftsinformatik* 41 (1): 77-82.
IBM Developer Works *Web site personalization*. 2000 http://www-4.ibm.com/software/developer/library/personalization/index.html
Kania, Deborah *Putting Personalization To The Test*. 2000 http://www.clickz.com/cgi-bin/gt/en/pm/pm.html?article=138
Luedi, Ariel F. 1997. "Personalize or Perish." *EM - Electronic Markets* 7 (3): 22-25.
Matzer, Michael. 2000. "Analysten streiten über Essentials bei der Verwaltung von Inhalten." *Computer Zeitung* (14): 17-17.
Middelhoff, Thomas. 1997. "Entwicklung einer Multimediastrategie für Medienunternehmem." *Die Betriebswirtschaft* 57 (3): 411-422.
Neuhold, Erich J. 1998. "Der Informationsarbeiter - heute und morgen." *Der GMD-Spiegel* 3/4 32-35.
Nielsen, Jakob. 1999. *Designing Web Usability: The practice of Simplicity*. Indianapolis: New Ryders Publishing
Nielsen, Jakob *Personalization is Over-Rated*. 2000 http://www.useit.com/alertbox/981004.html
NN. 1999. *Verleger haben nun weniger Angst vor der digitalen Autobahn*.
Palmer, Jonathan W. u. Eriksen, Lars Bo. 1999. "Digital Newspapers Explore Marketing on the Internet." *Communications of the ACM* 42 (9): 33-40.
Ryan, James. 1999. "Stop the press." *Business 2.0* (12): 196-209.
Sakagami, Hidekazu; Kamba, Tomonari; Sugiura, Atsushi u. Koseki, Yoshiyuki. 1998. "Effective personalization of push-type systems --- visualizing information freshness." *Computer Networks and ISDN Systems* 30 (1--7): 53-63.
Sander, Jörg u. Kröger, Carsten. 1998. "Content Management: Intranets richtig aufsetzen." *Diebold Management Report* (12): 14-18.
Shapiro, Carl u. Varian, Hal R. 1998. *Information Rules*. Boston: Harvard Business School Press
Varian, Hal R. 1991. *Grundzüge der Mikroökonomik*. München: R. Oldenbourg Verlag GmbH

6. Biographies

Dipl.-Inform. Thomas Ritz , born 1971 studied computer science and economics at the University of Bonn and obtained his degree in computer science in 1997. Since 1997 he is scientific employee at the University of Stuttgart IAT and the Fraunhofer IAO. He has a long standing experience in publishing industry and focuses his research activities on online communication and information. In several research projects he designed and implemented solutions for electronic information exchange, online communication in virtual communities, branch specific information portals and electronic commerce. Mr. Ritz acts as project manager in the DISMED project (funded by the European Commission in FP 4 Innovation program RSE 067) and in the PEACH project (funded by the European Commission in FP 5 IST programme IST-1999-11345). Besides he joined several public and industry funded research, consultancy and development projects.