

Digital Documents, Work and Technology, three cases of Internet News Publishing

Lars Bo Eriksen
Göteborg University
lars@adb.gu.se

Abstract

The introduction of new information technology can be a rationale for reconsidering what digital products organisations produce and reconsidering how work should be organised around the production of artefacts. In this article a pilot study covering three World Wide Web publishing cases is presented. Through a holistic view on documents, work and technology the three cases are described. The conclusion of the article is that the promises of digital documents are far from realised in the organisations studied and that the products provided are traditionally founded. This can be related to a number of factors but the little involvement of technical skilled people, inappropriate tools and uncertainty on how WWW technology can be integrated with the existing organisation are factors contributing to the weak implementation of digital documents.

1. Introduction

The promise of digital documents is not that of an end to paper, but the promise of new tools that enable humans to collaborate in new ways [1], that enables consumers to get the information they actually need [2], and tools that enables people to change the organisation of the way they produce, store and retrieve information in documents [3]. The Internet is the first infrastructure that enable these possibilities at a large scale. In this article the production of digital documents for the World Wide Web in three Scandinavian newspaper organisations is presented and discussed.

Newspapers are important institutions in the western world of today. As sources of authentication they play a central role in the writing of modern world

history. With a global computer network, the Internet or its successor, some expect this to change within a short period of time (cf. [4, 5]). Newspapers are turning to the WWW for several reasons. One possibility is the use of it to gather information material and do research with the purpose of writing articles, another kind of use is the provision of a WWW service, in other words some sort of digital newspaper. This last approach has gained much attention in the newspaper business and estimates spring 96 were that world wide more than 800 newspapers provide such services [6].

Introducing new technology into organisations can be a rationale for change [7]. This article documents a pilot study of three World Wide Web publishing cases, the aim of the study being to investigate how WWW technology is used to produce digital documents and how work is organised in relation to the documents produced. In this article we focus on the introduction World Wide Web publishing technology. The WWW can also be used for information retrieval but is not the focus of our study. As described by Blomberg, the same technology can have different impact on work in different settings[8], for which reason we have picked three newspapers to study.

2. Research Background

The digital documents produced, stored and retrieved in computer networks are not simply one-to-one analogue to prints on paper. Fundamental differences separate conventional documents from their digital counterpart. With the diffusion of computer networks these differences are becoming apparent to more and more people. Paper documents are typically limited to text and still images, while digital documents can contain sound and live images arranged in hypertext format (i.e. hypermedia). Where paper documents are static entities, digital documents can be "virtual", they can be generated on demand to suit individuals needs, the document content dependent on when and where

they are generated, the document being generated from underlying sources [3,9]. Having escaped the limitations of print on paper authors are no longer in complete control of what the readers view, making it harder to figure out what to write in the first place [1, 4, 10].

With the current use of Internet and World Wide Web, there is a discrepancy between technical potential and actual use. The World Wide Web protocols offer the change to go beyond static documents. The Web language Hypertext Mark-up Language (HTML) was initially designed as a language for communication over the Internet, deliberately designed so simple that computers as well as humans could generate it [11]. What we see today is that not only has HTML been accepted as a language of communication, but it has also become a *de facto* language in which many documents are encoded on the web. As a result, a large proportion of the World Wide Web has become a vast collection of documents in the classic sense: static and even preformatted documents. In this way the potential of Internet protocols as program interfaces is not used, the Internet solely being a distribution channel. Rice, Gruber and others argue that Web browser should not be seen as state of the art software for Internet use, but instead as neat and attractive interfaces for much more powerful applications yet to come [12].

In their article [4] Levy and Marshall discusses digital documents as related to work and technology in libraries. Based on work by Yates [13] they establish and describe dependencies and relations between *documents, work and technology*. In this article we use the framework of Levy and Marshall to discuss how web technology relates to and influences documents and work in the three organisations studied. Any new digital document (library) product will be based on an enabling technology. In the cases of the three newspapers this enabling technology is WWW.

The research approach to collect and analyse data for this article and the motivation for the data presentation is described in section 3. Section 4 presents the three cases. The cross case findings are summarised in section 5. Section 6 contains a discussion of the findings and section 7 concludes the article.

3. Research Method

The research presented in this article was mainly based on qualitative studies carried out in three Scandinavian newspapers. The aim of the study was to contribute to a clearer understanding of what challenges newspapers and other organisations currently face. From this desire to understand the impact of a new

medium in such organisations, an applied research project was established [14].

The organisations chosen for the study were purposefully selected: All are early starters with Internet. The studies were carried out in three different countries: Denmark, Sweden and Norway. The diffusion of Internet has taken different pace in these three countries, our impression (although not clear) being that Denmark lags behind in terms of number of users and in terms of experience in organisations with using the Internet. This factor is expected to contribute to higher variation between the selected organisations. With the terms of Patton this way of selecting cases can be described as intensity sampling [14].

Within each organisation a chain sampling strategy was used to identify people with knowledge and skills of interest for the study. The unit of analysis were individuals. Some of these were interviewed (tape recorded and transcribed) while others were overtly observed in their everyday work. Parts of the interviews could not be taped, since we also talked about the topics when the interviewees showed us around in the newspaper organisations.

The interviews carried out were all based on an interview guide approach [14]. On the one hand we had a desire to get a holistic view of the organisation and the production of the Internet edition, but on the other hand little was known in advance about what was actually going on. The interview guide was used to ensure that we got the information we needed in the interviews, but it did not determine the sequence or structure of the interview. Our aim was to allow new topics of interest to emerge as the interviews went on.

A wide range of questions were discussed covering design aspects, user interaction, current status related to the future expectations, the impacts on everyday work processes, the production of articles and tools used in this process. The results reported, i.e. the case descriptions and the similarities and differences between the papers, have been derived in a bottom up approach.

In two of the three organisations we observed the process that lead to the production of the Internet edition. In both organisations this observation was done within a four hour period. We also studied the Web services of the three papers. As a research method this can most appropriately be described as unobtrusive measures [15], although in principle our access to the Web services could be logged and hence was not purely unobtrusive.

Through an inductive analysis [14] data has been analysed to yield the results presented in this article.

4. Cases

In this section the three cases are presented. For each case a brief history of the Internet publishing service is described to answer question on how and why the services were initiated. For each case we then presents aspects of *products, technology and work* from a perspective based on work by [13] and [4]. The digital product each newspaper provide is described at a general level. The Internet and WWW in itself is considered the enabling technology, but for the organisations to publish on WWW additional technology must be used to the produce the material for the Internet edition. The organisation of work in relation to the product and the technology is also described for each case. Finally each case description presents the rationale of the services in order to give a situational context for the design of the services.

4.1 Jyllandsposten

In number of printed newspapers Jyllandsposten (JP) currently is circulated in 170.000 copies and is thereby the largest newspaper in print in Denmark. Through an aggressive market campaign and an increased focus on journalistic competence (an increase in employees) this status has been achieved over the last two years. In comparison with other Danish newspapers, Jyllandsposten has the largest journalistic staff in Denmark and correspondents around the world. The focus on journalistic competence has meant that less attention has been given to technical aspects of newspaper production. The paper, although its name refers to a specific part of the country, is distributed and read in the whole country. At the time of the interviewing, Jyllandsposten was the only major newspaper in Denmark providing a Internet service.

Jyllandsposten became visible on the Internet in the summer of 1995. Initially the motivation for going on the net, was the establishment of a "computer club", a service that provided Internet access and software to subscribers of the printed newspaper. This evolved through autumn 95 into a service which provided news in brief. In November 1995 it was decided that Jyllandsposten should have a substantial service on the Internet. The 15th of January 1996 "Internetavisen Jyllands-Posten" was launched. The initiative was the work of three individuals involved in the establishment of the first service. During autumn 1995 these three people, one journalist, one with a Master's degree in media science and a student, designed a full blown Internet news service and got upper management approval for their idea. Through the use of a consulting firm specialised in Internet services, the requirements

for the service were balanced with the technical possibilities.

Product

"Internetavisen Jyllandsposten" mainly consisted of replicated articles from the printed newspaper. There was little use of pictures in conjunction with the articles. A search facility enables registered users to search all material published in the Internet edition since it was established. Beyond these facilities, the "JP Computer Club" offered downloadable software. Registered users had full access to all articles, where non registered users had access to a few news articles without any depth. The fee for using the service was 1,800 DKK a year, a price similar to the subscription rate of the printed newspaper. The access control mechanism was in the process of establishment when this study was carried out.

The layout and design of the service reflected the desire to duplicate the printed newspaper. The Internet edition was separated into section and within each section a number of articles was available through an overview bar that was designed with the aim of allowing flexible navigation through the service.

Technology

The technical system used for the production of the printed newspaper was a dedicated text-based interface system ATEX introduced in the early 80'es. At Jyllands-Posten articles for the printed paper were arranged manually. The articles were printed out on transparent film and by hand arranged on pages that eventually become the master from which the newspaper was printed. In comparison with many other newspapers in Denmark the technical level of sophistication was low at Jyllands-Posten.

When an article leaves the hands of the journalist a copy is spooled into a separate directory of the ATEX system. Within this directory a dedicated software program automatically converts the ATEX representation to HTML, the language of the Web. When all articles have been converted and checked for major errors the HTML representations are transferred by file transfer protocol to the Web server, situated outside the organisation at an access provider, for further processing. At the server additional software indexes the articles and arranges them in a file structure that allows consumers to navigate through the material.

Work

The Internet publishing division at Jyllandsposten was formed as a project group with three members. The editor of the printed papers IT section was responsible for the journalistic issues of using of the Internet and in charge of the project. Besides working with the Internet edition he was involved with the production of the IT section for the regular paper. The media expert was responsible for further development of the service and for attracting advertisers. The third participant in the project was responsible for the daily routine of creating HTML versions of the articles and transferring the articles to the Web server. None of these people had any specific technical skills. The initiative for new products was largely the responsibility of these three people. By entering into a dialogue with the consulting firm, the software needed for realising new designs is developed.

The use of the Internet for distributing the contents of the newspaper did not mean any changes in the work of people not active in the Internet publishing division. Being based on the printed paper, the main task was that of transforming the representation of the articles in ATEX to HTML, this process being highly automated through the use of dedicated software. Although the process was supposed to be fully automated, some minor mistakes did occur from time to time, and the staff at Jyllands-Posten had to manually edit the HTML files in order to make the text appear in a proper way. Further the articles needed to be categorised manually for the indexing software running at the server to place the articles correctly into the digital paper.

Rationales

Two distinct rationales existed at Jyllandsposten for having an Internet service. The first and initiating rationale was that of experimenting with the technology. There was a feeling of a need to be up-to-date with the new technology. A combination of future fear and fascination of the technology lead to the establishment of the WWW service, as expressed:

"No doubt what so ever, it is important for a newspaper to be on the Internet. Because it will have an impact, in one way or the other ... without anyone knowing exactly how it will make a difference, therefore you have to be [on the Internet]"

P.N., Jyllands-Posten

Jyllandsposten had a clearly stated short term goal of earning money on the service. Although experience show that this can lead to a dramatical drop in number of consumers (cf. [16]) two arguments are stated in

favour of earning money on the service. Jyllandsposten is specialised with regard of news concerning Denmark, no other Web service in the world¹ offers the same coverage of events and sports in Denmark. Further by being written in Danish it is aimed at people who prefer this language, thereby being specialised with regard of language. The statement being that if not economical sound from the beginning, the service never will be.

Jyllandsposten	Internetavisen Jyllandsposten
URL:	http:www.jp.dk
Age of service when studied	3 month
Estimated accesses a day	4000 people/day
Staff	3 part time employee
Main digital Product	Electronic copies of articles from printed newspaper
Product Technology	External filebased Webserver
Production Technology	Conversion scripts, html-editors.
Work	Development and duplication
Rationale for design of service	distribution to people not having access to printed edition.

Figure 1. Characteristics of Jyllandspostens service.

4.2 Göteborgsposten

Göteborgsposten (GP) is a regional newspaper covering the city of Göteborg and the region of the Sweden in which Göteborg is situated. GP is circulated in 270,000 copies a day. The journalistic strength of GP is the extensive covering of events in the local area and sports. The paper is what journalist defines as a regional market maker, an information source extensively covering the region.

Göteborg city was the host of the world championships in athletics in August 1995. In order to provide information for participants and the public, an intensive use of WWW as a channel of information on the games and the city was established. One of the organisations involved in this was Göteborgsposten. It seemed obvious to the management to use the WWW for distribution of news as well and 14. of august 1995 "GP Direkt" was in existence. In October 1995 the

¹ When the studies were carried out, Jyllandsposten was the only major Danish newspaper publishing on the Internet.

service was heavily redesigned, the amount of content increased and the "fact databases" were established. In November "GP Direkt" started providing updated news on the Internet throughout the day. In November a classified adds service DAGS² was started as the result of a joint venture between SISU³ and Göteborgsposten, Sydsvenska Dagbladet⁴ and Dagens Nyheter. Through December 95 to February 96 additional service on film reviews and more "fact databases" were initiated.

Product

The "GP Direkt" service offered selected articles from the printed newspaper. These were kept on the Web server for a week and then discarded. Besides the mirror of the printed paper a variety of information was available. Lists of film and CD reviews could be accessed, these list were updated regularly and formed an ever growing archive of reviews. As an experiment aimed at schools, a series of "fact databases" were available. These "databases" cover such events as the war in ex-Yugoslavia. The material in the databases was mainly articles from "GP Direkt", but in order to make a consistent source of information, additional documents and text have been added. The "GP Direkt" is free of charge.

Technology

The printed newspaper was produced by the use "Quark Express" a high-tech graphical interface system. Through the use of Macintosh Computers and graphical displays, the newspaper was edited and arranged directly on screen. Images and text was presented in a "WYSIWYG" manner that enables Göteborgsposten to print out the master for the printed newspaper directly from "Quark Express". The process of producing the Internet edition is in itself an add-on to Quark Express.

Two software programs written specifically for Göteborgsposten allows the staff to extract articles from Quark Express. One program is used in conjunction with Quark Express to extract and convert articles. The second program assembles the separation articles into a hyperlinked structure of HTML documents arranged in a file system.

The product provided was based on the use of an external placed Webserver to which files were uploaded on a daily basis.

Work

Five people were, in various degrees, working with the production and maintenance of the Internet service. The manager in charge of the Internet service was responsible for electronic media and electronic publishing. Besides working with the Internet edition, he had several other tasks at Göteborgsposten to attend. Two journalists were the "core" of the Internet edition. They were responsible for maintaining a consistent set of pages and also worked on the establishment and development of "fact databases". And they were responsible for continuously updating the content of the service throughout the day as new stories arrived from their newsbroker, a Swedish equivalent to Reuters. Besides this they updated the different archives on music, film and Internet material. Another journalist is responsible for developing new uses of the Internet, with a special concern to technical issues. The last of the five is responsible for developing the commercial side of the product and selling space for Internet commercials.

The production of the Internet edition contained much trivial work. The aim of the Internet staff was to be "on the net" around 9.30 in the morning. To reach this goal they started at 8 o'clock by deciding what articles to put on the Internet. This was done by going through a printed paper, taking notes of which articles to put on the net. When the selection process was over, one hour was spent on the extraction of the material from Quark Express into preliminary HTML documents. This was done by the use of the add-on for Quark Express, that enabled a journalist to click on text blocks in a particular order, thereby specifying which text was the header, which was the author and which was the main body of text for a particular article. The preliminary HTML files were named in a particular way to indicate their contents. The add-on unit could not fully transform "Quark Express" representation to HTML, therefore the journalists had to check and to some extent rewrite the HTML extracts in order for them to conform to the HTML syntax. When all articles had been extracted, the files were transferred to a PC, where a script processed the files. This script added navigational links to all documents and generates indexes for the Internet edition. This new set of HTML documents was then transferred back to the Macintosh, from where it was transferred to the server situated at an access provider. Thereby the Internet edition became available to the rest of the world. By transferring the articles the process was far from over. Over the next two to three hours the product is polished as bugs are reported (often by users) and further by using HTML editors to insert images and additional headers.

In an attempt to utilise the Web as a resource one journalist each morning appended links to articles when

² <http://dags.gp.se>

³ Swedish Institute for System Development, established 1984, URL: <http://www.sisu.se>

⁴ <http://www.sds.se>

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.