

An Intelligent Information System for Blind People – AI Technology and Philosophical Aspects*

Thomas Kieninger, Norbert Kuhn**,
Kerstin Seidenschwann, Werner Weiss

Deutsches Forschungszentrum für Künstliche Intelligenz
Postfach 20 80
D-67608 Kaiserslautern

Abstract. Communication to exchange information is essential for humans. In many fields of communication this has led to more or less fixed forms that guide or determine the way how people exchange information. This concerns at first a logical layer of the information blocks, i.e., which information has to be transmitted but also a procedural layer describing how the information is send.

A very important domain of communication is the business letter domain. Here, a certain standard has evolved that determines the form of business letters. We distinguish a logical structure and a layout structure that is captured in the international Standard called *Office Document Architecture(ODA, ISO 8613 [1])*.

In this paper we discuss the problem of presenting printed documents to blind people who cannot perceive the information provided by the layout structure of a document. We outline how the ODA Standard can be exploited to allow random access to the logical entities of a printed document. This approach is based on a system that has been developed at the DFKI and which has been described in [5] and which is able to extract the layout information from certain classes of business letters automatically.

1 Introduction

The increasing number of computer applications had strong impacts especially for the office world. Computer based tools are essential in almost any company. They support humans in routine tasks and - what might be even more important - they are a basic means for managing communication among humans, either within the in-house office environment or for communication with other company partners. Computer technology - and in particular computer networks - yield a lot of benefits for business world: information is available on-line, at an

* This work has been supported by the Stiftung Rheinland-Pfalz für Innovation

** phone: (0631) 205-3451, E-mail:kuhn@dfki.uni-kl.de

arbitrary number of different places, and information can be exchanged almost immediately.

For blind people computer technology has substantially improved their communication facilities. Coupling a computer with braille output devices or a speech synthesiser enables blind people to have access to information provided by electronic mail-boxes, electronic bulletin boards or international databases. With an optical scanner and appropriate optical character recognition (OCR-) software even paper bound information can be perceived without the help of seeing people. This technology brought a large amount of autonomy and independence for blind computer users and thus gave reason to assume that it would allow the integration of blind people into an office environment.

Futhermore, there have been efforts to develop integrated computer based products that are particularly tailored for blind people.

A first product of this kind is *ETaB* (Electronic Newspaper for Blind People) which is developed and distributed by the Stiftung Blindenanstalt in Frankfurt/Main (cf. [4]). The subscribers of *ETaB* obtain a daily version of the Frankfurter Rundschau (FR), a major German newspaper. *ETaB* already provides different modes of navigation within a newspaper. E.g., a possibility for keyword search is provided. Because up to now there is no common standard used by the editors of newspapers, the FR is the only german newspaper available in electronic form. However, an attempt for such a common standard is underway. This standard will be based on the Standardized Generalized Markup Language (SGML, ISO 8879, [2]) and should allow the integration of other newspapers in the next version of the *ETaB* system. Also, in the current version of *ETaB* it is not possible to include graphical portions in the presentation of the paper.

This problem, namely the access to graphical information for blind people is tackled in a large project called Graphical User Interface for the Blind (GUIB) which is sponsored by the European Community within the TIDE (Technology Initiative for Disabled and Elderly People) program ([10]). The objective of GUIB is to provide blind people with access to graphical user interfaces like MS-Windows or X-Windows. Within the project, a certain hard- and software solution has been developed to compensate the visual deficiency of the blind. The results of the project have been published in several papers, e.g. [7] or [6].

However, these systems do not apply for office workspaces. In nowadays offices information is to a large extent still paper-bound. Paper documents contain a lot of information which is in part only relevant at different steps of the document processing. Often they provide information that is not even written down explicitly: the layout structure of a letter guides a reader through a document. It helps finding the recipient or the addressee quickly, or it gives information about the contents, e.g. we might infer that a letter from a company that has a table layout is a bill for something they had delivered.

For blind people it is difficult to perceive this kind of information that is not written down explicitly. In principle, they can have access to printed information by the use of scanners and appropriate OCR-software. But, these tools produce an ASCII representation of a document where the layout and thus the logical in-

formation provided by the document is lost. Therefore, within the DFKI project ALV (Automatic Reading and Understanding) a system has been implemented that allows to transform a printed letter into an ODA-based representation. The ODA standard ([1]) is chosen for this purpose because it considers both, the logical and the layout structure of a document.

The project PASCAL 2000³ at the DFKI now has two major goals. Firstly, we want to exploit this kind of technique and integrate it into an office workspace for blind people. Secondly, in a philosophical part of the project we want to analyze the different impacts and the policies of document analysis tools. Within this attempt we want to integrate blind computer users to obtain feedback for ergonomical demands from the beginning of the system development.

In this paper we first describe the document analysis approach and then discuss the basics for the philosophical aspects. We conclude with a summarizy of these points and give an outlook for our nearest research activities.

2 The Document Analysis Procedure

The purpose of the document analysis procedure is to have as input a printed letter which is then transformed into an electronic version where the logical entities are maintained. A common standard which considers these logical entities explicitly is the Office Document Architecture (ODA).

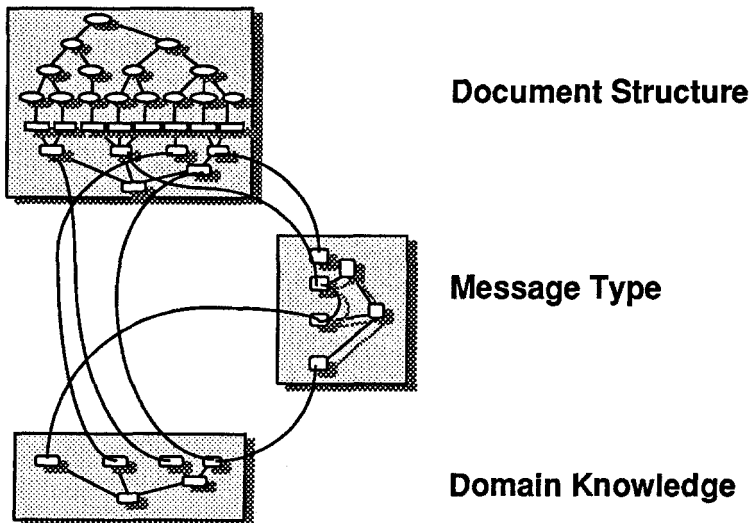


Fig. 1. Information Structure for Business Letters

³ The name is dedicated to the French mathematician and philosopher Blaise Pascal (1623-1662)

Figure 1 shows some different parts of knowledge that are important for *understanding* a single business letter. This means that knowledge of these types has to be modeled to be able to understand business documents automatically, i.e. by a computer program.

The graph in the uppermost box of the figure displays the information by the *document content* as its own, that means the content regardless from 'environmental' information. The square boxes in the middle of the graph correspond to content portions of the document. These are text portions that may be circumscribed by a certain rectangular. Above and below the square boxes we may recognize two disjoint trees that represent two different hierarchies, namely the layout and the logical structure of the particular document. For example, in the business letter the layout objects are pages, blocks, lines, words, etc. while the logical objects comprise elements like sender, recipient, signature, and so on. Via the connection with the content portions we get a mapping of the layout structure onto the logical structure. This process is referred to as *logical labeling*. One central idea in the ALV project at the DFKI has been that the logical labeling is done using only geometrical information, i.e. information of a certain layout object in relation to others. This implies that a hypothesis for a logical label for a particular content portion can be generated before an OCR procedure has started at all (cf. [5]). This approach is very much related to the way that humans perceive a letter. However, as this requires the geometrical information which is only available via visual sensing, this mode of perception is hidden to blind people. Therefore, a first goal in our project is to exploit the existing technique to compensate this deficiency to some degree.

Another type of knowledge that may be useful is the *message type* of a letter. On the one hand this knowledge can be used to understand the logical layout of a letter. For example, in general bills do not comprise a signature, which will make it unconventional to look for it after a letter has been classified as a bill. To classify a letter as a bill we have different approaches. One is to look at the content and the layout of the letter and then come to the result because there are some phrases like 'we have delivered', 'is due till', a.s.o. which make it likely to be a correct guess.

On the other hand, before we receive a bill we usually sent an order to or received an offer from the sender of the bill. Thus, we probably have some background information of the subject of this letter that can be useful in understanding it. This type of information is shown as the *domain knowledge* in figure 1. By this we want to capture all knowledge about processes or about the workflow within an office, a company or in a business correspondence, e.g. the temporal precedence relation between an inquiry, an offer, an order and a bill. If we the dates of a couple of letters are inconsistent according to this relation they are unlikely to fit together into a correspondence. However, if a new document fits in such a chain this gives supporting argument to classify a letter as a bill.

3 An integrated approach: AI and technology-assessment

At the Dartmouth Conference in 1956 John McCarthy presented his vision of an artificial intelligence to the world. Since this birth of a new field in research, a broad knowledge about methods and techniques was composed. The fascination to imitate human cognitive competence, e.g. sensory perception, problem solving or speech and pattern recognition, stimulated constantly the scientific community to increase knowledge and to search for appropriate applications.

The concept of document analysis for paperbound information presented in section 2 is based on AI techniques. Surely, document analysis may have a tremendous impact on the future design of information systems, especially on information systems for blind people. The idea of a supporting communication tool, that enables blind people to work within a modern office environment, is no longer only a benevolent verbal statement; it will become reality in the near future.

Since the original intention of AI is to imitate human cognitive competence on machine processes, this research field is closely connected with anthropological disciplines, e.g. philosophy, linguistics, neurobiology or neurophysiology (cf. [9], p. 1). Therefore, AI projects should be understood as an interdisciplinary approach with the goal to design new communication and information technology as user friendly as possible. The analysis of this interfering framework of different disciplines will help to evaluate the impacts caused by intelligent information systems.

Obviously, the application of AI-techniques bears great potential for the social and functional integration of visually impaired people. Apart from enumerating the hidden chances and the potentials for integration, the enthusiasm should not prevent from also analysing risks for blind users. As shown in figure 3, the multidimensional impact of this technology can be analysed in the light of

- ethical/anthropological aspects as well as
- human factors referring to rehabilitation.

The integration of technology assessment into the system development phase opens the possibility to consciously take into account its amicability towards the human being, society and nature. New technologies' characteristics are hyper-complexity and hyperdynamic which can easily become intransparent for the user. An analysis can show which effects the system has on the self-experience and on the self-esteem of the blind user, if either the computer is a communication partner or if it dominates the user by information abstraction and the information selection process done by the intelligent information system. Part of the human factors that shall be realized is a personality developing work structure as well as a sociotechnical system design. Besides compensating the handicap through knowledge representation the system aims at the integration of blind people into the working world by providing a real workspace.

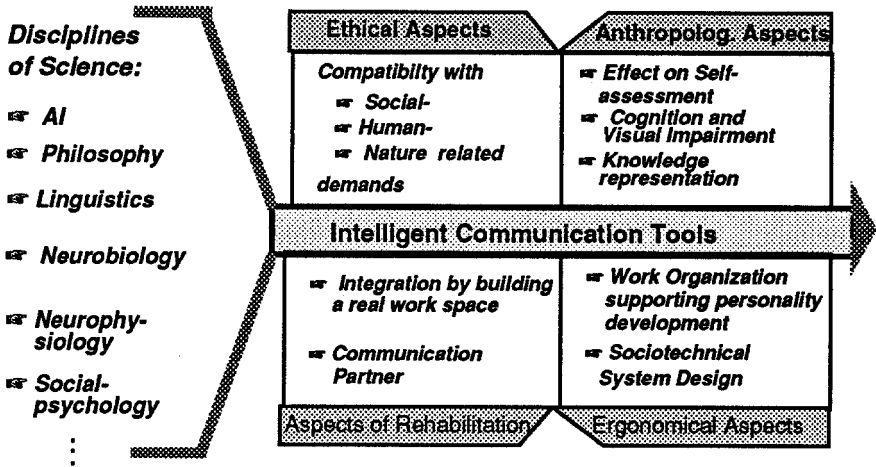


Fig. 2. Chances and risks within aspects of ethics, anthropology, rehabilitation and human factors

Being aware of the ambiguity of intelligent information systems, we have to fulfil an integrated approach of AI-and technology assessment (cf. [8]). Technology assessment can be defined as a systematic analysis, where all impacts and complications, direct or indirect, real or potential, present and future of a technology are defined, evaluated and measured as well as the causeeffect relationships identified. ([3], p. 339) The conceptual decision to assess impacts caused by intelligent information systems must not be misunderstood as a form of "technology arrestment". Rather, this approach will lead to a higher performance of the designed system and greater acceptance by visually impaired users.

4 Conclusion and Outlook

The project PASCAL 2000 is an interdisciplinary project in the fields of Artificial intelligence and philosophy. More specific, it deals with the problem of integrating document analysis tools into an office workspace for blind people.

From the AI point of view its main idea is that paper documents for office communication are arranged in a way that their layout structure corresponds to a commonly agreed logical structure and thereby, supports understanding of the content. This information is destroyed if a document is presented based only on its textual (i.e. ASCII-type) representation. Therefore, we have to use other forms of representation like the one which maintains the logical entities based on the ODA standard for business letters described in section 2. For other kinds of documents (e.g. electronic documents like E-mail letters) other standards can be reasonable assuming that they provide direct access to some kind of logical entities which are directly coupled with content portions of the corresponding document.

The purpose of the project PASCAL 2000 goes beyond dealing with a single document. The overall goal is to exploit such kind of techniques and integrate them into an intelligent office communication system. This system allows random access to single entities of paper bound information and - in that sense - compensates the perceptual deficiency of blind people by using other means of presentation. Then, this person can act as a controller in workflow processes within an office. E.g., the ALV technique can be used to extract an offer from the document data base which corresponds to an order the person is just going to process. We believe that this technology contributes to an enrichment of the field where blind people can work. In general, we will develop plan based, goal oriented query answering mechanisms that enable the system to cope with complex and high level queries for the document data base.

References

1. ISO 8613. Information processing - text and office systems - office document architecture (oda) and interchange format, volume i-iii, parts 1-8, 1988.
2. ISO 8879. Information Processing - Text and Office Systems - Standard Generalized Markup Language (SGML), 1986.
3. B. Bartocha. An instrument for goal formulation and the selection of problems. In M. Centron and B. Bartocha, editors, *Technology Assessment in a Dynamic Environment*, pages 337-356. London/New York/Paris, 1973.
4. Stiftung Blindenanstalt. Elektronische Tageszeitung für Blinde. daily newspaper, Frankfurt a. Main.
5. A. Dengel, R. Bleisinger, R. Hoch, F. Hönes, F. Fein, and M. Malburg. π_{ODA} : The Paper Interface to ODA. Technical Report RR-92-02, Deutsches Forschungszentrum für Künstliche Intelligenz, 1992.
6. K. Fellbaum, K. Crispian, M. Krause, T. Strothotte, M. Kurze, and J. Emhardt. Multimedia interfaces for blind computer users. In E. Ballabio, I. Placencia-Porrero, and R. Puig de la Casa, editors, *Rehabilitation Technology*. IOS Press, Amsterdam, 1993.
7. D. Kochanek. Hypertext system for the blind newspaper reader. In W. Zagler, editor, *Proceedings of the 3rd International Conference on Computers for Handicapped Persons (ICCHP)*, Wien, 1993.
8. W. Rammert. Wie KI und TA einander näherkommen. In *KI*, volume Heft 3, pages 11-16. 1993.
9. VDI Report. Künstliche Intelligenz. Leitvorstellungen und Verantwortbarkeit, 1993.
10. TIDE. Technology initiative for disabled and elderly people. *Amtsblatt der Europäischen Gemeinschaften*, 1993.