

REPORT



Digitalisation and Information Technology

ToKeN

Access To Knowledge and its enhancement Netherlands
Symposium, March 18, 2005 Eindhoven

Netherlands Organisation for Scientific Research

PROGRAMME OFFICE TOKEN

Netherlands Organisation for Scientific Research (NWO)

Physical Sciences

Dr.ir. Christiane Klöditz

P.O. Box 93460

2509 AL The Hague

The Netherlands

Phone: +31 (0) 70 3440 561

Fax: +31 (0) 70 3440 861

E-mail: token@nwo.nl

Website: www.nwo.nl/token

Held every year since 2002 this year's symposium took place in the Auditorium of the Technische Universiteit Eindhoven. The day was organised jointly by J.F.Schouten School for User-System Interaction Research Eindhoven and NWO Physical Sciences. About 75 people attended the meeting.



RECEPTION DESK AND PARTICIPANTS

TOKEN¹ – ACCESS TO KNOWLEDGE AND ITS ENHANCEMENT NETHERLANDS

ToKeN is a multidisciplinary NWO research programme involving specialists in both cognitive and computer science. It focuses on the ability of individuals to retrieve relevant knowledge and information from computer systems and to derive implicit knowledge from raw data. The overall aim is to develop methods and techniques to optimise the interaction between human users and advanced multimedia information systems. At present twenty ToKeN-projects are ongoing in the three application domains ToKeN aims to cover, namely Education & Culture (six projects), Healthcare (five projects) and Law enforcement & the Judicial system (nine projects).

In his opening speech the chair of the Programme Committee, **Prof. Jaap van den Herik** (Universiteit Maastricht) outlined the background, the context of and progress within ToKeN. He emphasized that making knowledge accessible in the three application domains is one of the main challenges of the programme.



PROF. JAAP VAN DEN HERIK, CHAIR PROGRAMME COMMITTEE

¹ Previously known as ToKeN2000 the programme has been renamed to ToKeN in 2004.

The **keynote speaker** was Univ. Prof. Dr. Manfred Tscheligi, director and co-founder of the Center for Usability Research & Engineering (CURE) in Vienna. In his presentation "User Experience Management - Interdisciplinary Strategies for Next Generation Services" he accentuated that design knowledge which is available has to be matched with experience knowledge. According to Tscheligi the concept of user experience is more than a new word. However, since there is not much research carried out on this he called on researchers to work on this concept.



UNIV.PROF. MANFRED TSHELIGI

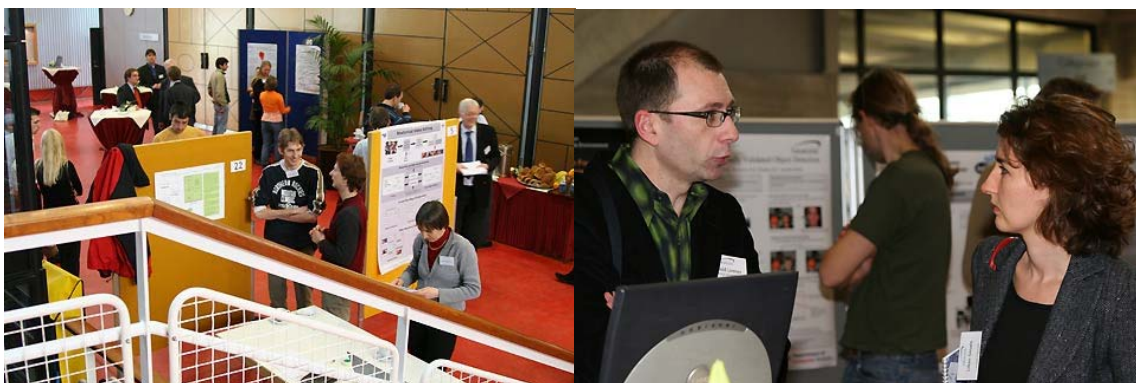
After a short break the symposium continued with **flash presentations** of the eight new projects on Law enforcement & the Judicial system. Also six of the ongoing projects were illustrated in **detailed lectures** by directly involved PhD students and postdoctorals. The programme of the day as well as abstracts of all presentations are included in Appendix 1 and 2 of this report.

According to the tradition on the ToKeN-events, lunch was served accompanied by a **lunch concert**.



MUSICAL INTERMEZZO DURING THE LUNCH BREAK

After the lunch more than twenty posters and demonstrations were presented at an extensive **poster session** (see Appendix 3 for a list of the posters).

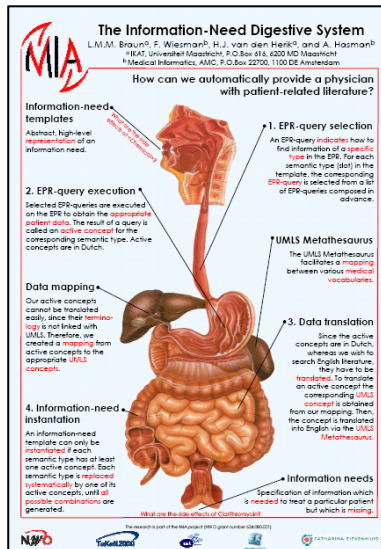


IMPRESSIONS OF THE POSTER SESSION

In his concluding lecture on thoughts of potential users **Dr. Hans van Eekelen**, chair of the users committee ToKeN, drew the attention on the challenge of keeping the balance between fundamental and application oriented research. He also stressed on the importance of coherence of the larger ToKeN-projects. Both are no trivial tasks also seeing the fact that most of the projects involve various research fields and are spread over several universities in the country. Finally, he explicitly called on the ToKeN-researchers to publish extensively and at the same time make the publications better accessible for potential users by using a minimum of jargon.



DR. HANS VAN EEKELLEN, CHAIR USERS COMMITTEE



THE WINNING POSTER

Also a tradition of the ToKeN-days is a poster contest where the best poster is nominated by the participants. This years winner was the poster entitled *The Information-Need Digestive System* by L.M.M. Braun, F. Wiesman, H.J. van den Herik, A. Hasman.



PROF. DON BOUWHUIS

The chair of the day, **Prof. Don Bouwhuis** concluded the day with words of appreciation to the speakers of the day, to the participants and to those involved in the organisation of yet another interesting and inspiring ToKeN-day.

Appendix 1: Programme March 18, 2005

- Chair: Prof. dr. D.G. Bouwhuis, J.F. Schouten School for User System Interaction Research,
Technische Universiteit Eindhoven (Member Programme Committee ToKeN)
- 9:45- 10:15 *Registration and welcome with coffee/tea*
- 10:15 *Opening*
- 10:15 - 10:30 *Flashback ToKeN March 2004 – March 2005*
Prof.dr. H.J. van den Herik, Universiteit Maastricht and Chair Programme Committee
- 10:30 - 11:15 *Keynote presentation: User Experience Management - Interdisciplinary Strategies for Next Generation Services*
Univ. Prof. Dr. Manfred Tscheligi, Center for Usability Research & Engineering Vienna and ICT&S Center, University of Salzburg
- 11:15 - 11:35 *Coffee/Tea Break*
- 11:35 - 12:20 *ToKeN/Law enforcement & the Judicial system: Presentation of eight new ToKeN-projects*
- 12:20 - 12:40 *ANITA: Law, technology and practice: exchanging crime-related information within a new legal regime*
Mr. H.H. Kielman and Mr. W. Koelewijn, Leiden University, Faculty of Law
- 12:40 - 14:40 *Lunch Break and Poster Session*
- 14:40 - 15:00 *I-CATCHER: Intensive-Care Access to Terminology & Course of Health Exploration and Retrieval*
Dr. A. Abu-Hanna, Academic Medical Center (AMC) at the University of Amsterdam, Department of Medical Informatics; Dr. M.C.A. Klein, Vrije Universiteit Amsterdam, Artificial Intelligence Section
- 15:00 - 15:20 *AUTHENTIC: Bridging two worlds*
Drs. I.E. Berezhnoy, Universiteit Maastricht, Institute for Knowledge and Agent Technology
- 15:20 - 15:40 *NARRATOR: Nominalizations; an example of detailed semantic analysis for information retrieval in Narrator*
Drs. H.G.B. Reckman, Leiden University, The Leiden Centre for Linguistics (ULCL)
- 15:40 - 16:00 *Coffee/Tea Break*
- 16:10 - 16:30 *EIDETIC: Context-based object detection for CIBR*
Ir. N.H. Bergboer, Universiteit Maastricht, Institute for Knowledge and Agent Technology
- 16:30 - 16:50 *MIA: Towards Better Health Care with Agents*
Dr. F.J. Wiesman, Academic Medical Center (AMC) at the University of Amsterdam, Department of Medical Informatics
- 16:50 - 17:15 *Thoughts of potential users*
Dr. H.A.M. van Eekelen, Chair Users Committee ToKeN
- 17:15 - 17:25 *Closing remarks*
Chair of the symposium
- 17:30 *Reception*

Appendix 2: Abstracts

Abstracts flash presentations of eight new ToKeN projects in the application domain of Law enforcement & the Judicial system

- **EVIDENCE: Making sense of evidence: software support for crime investigations**

PI: Prof. dr. mr. H. Prakken, Universiteit Utrecht (Email: henry@cs.uu.nl)

Crime investigation is a difficult and laborious process that is prone to error and the costs of mistakes can be high. This project will develop software support to tackle two problems that often occur in complex criminal cases: lack of overview of a mass of evidence by initial investigators and lack of transparency of case files for subsequent investigators, prosecutors and fact finders.

A demonstrator prototype will be developed of software with which crime investigators can visualize and analyse their reasoning about a case. Such software will support investigators in seeing patterns, discovering new relationships or inconsistencies, and identifying missing evidence. It will also enable subsequent investigators, prosecutors and fact finders to gain a better understanding of a case. The software will be built as an extension to evidence data modelling software currently used in six Dutch police regions. The effects of the software on an investigator's understanding of a case and the quality of file transfer will be measured in detailed user experiments.

The software will be based on recent insights that the only viable manner in which police investigators can structure the gathered information is through stories about what happened, linked to the available evidence and hypotheses with evidential arguments. These insights will be further developed and made precise by clarifying how stories can be anchored in the available evidence by building arguments with general knowledge. Thus the necessary theoretical foundations will be provided for software that is theoretically sound and respects the practical constraints of crime investigation.

- **DALE: Data Assistance for Law Enforcement**

PI: Dr. W.A. Kusters, LIACS, Leiden University (Email: kusters@liacs.nl)

The rapid growth of available data in all regions of society requires new computational methods. Besides traditional statistical techniques and standard database approaches, current research known as Data Mining (DM) uses modern methods that originate from research in Algorithms and Artificial Intelligence. The main goal is the quest for interesting and understandable patterns. This search has always been and will always be a critical task in law enforcement, especially for criminal investigation, and more specific for the fight against terrorism.

Data Mining, sometimes also referred to as Knowledge Discovery in Databases (KDD), can be defined as "the non-trivial extraction of implicit, previously unknown and potentially useful and understandable knowledge from data". Databases from law enforcement applications are usually large, and contain data with varying types, including free formats. The data is temporal, dynamic and noisy; often information is explicitly hidden.

Research therefore focuses on semi-structured data and pattern bases. In particular association rules will be used. The goal is to compose a framework for data mining in law enforcement, fed by problems that arise from this area and in close cooperation with domain experts. The systems developed must be such that future developments can be embodied. Emphasis is - for the moment - on text oriented databases. The approaches are however such that, e.g., multi media databases can be incorporated at a later stage. Research will both start from known theories and from questions arising from daily practice.

The research is carried out at LIACS, the computer science department of Universiteit Leiden, in cooperation with KLPD, Dutch national police at Driebergen. The project starts on 1 March 2005. The main research effort will be provided by two PhD researchers.

For more information, please contact the project leader dr. W.A. Kusters, kusters@liacs.nl.

The full project proposal can be found at <http://www.liacs.nl/home/kusters/dale/>.

- **ACCESS: Agent-based Criminal Court Electronic Support Systems**

PI: Prof. dr. F.M.T. Brazier, Vrije Universiteit Amsterdam (Email: frances@cs.vu.nl)

Software agents are autonomous entities that can pro-actively process information in distributed digital environments. The unique pilot project of the courts in Amsterdam and Rotterdam in which digital dossiers are used by judges and public prosecutors is currently based on static information.

This project proposes to extend the digital dossier with additional information from different sources, based on needs identified in interaction with the Rechtbank Amsterdam. Agents may have partial control over the information access and processing process, currently in hands of human legal experts. Integrity of data and processes is essential. Integrity management is one of the main elements in system design. This project proposes to define the requirements for integrity management in (semi-)closed environments.

Close collaboration with the Rechtbank Amsterdam is essential. A prototype agent system will be designed for the purpose of exploration of options: both technically and legally. The results will also include open legal and technological issues with respect to the use of agents in more open environments.

- **CASSANDRA: Context-Aware SenSing for Aggression Detection and Risk Assessment**

PI: Prof. dr. D.M. Gavrilă, Universiteit van Amsterdam (Email: gavrilă@science.uva.nl)

Increasingly, surveillance technology is fielded to help safeguard public spaces such as train stations, shopping malls, street corners, in view of mounting concerns about public safety. Traditional surveillance systems have a human operator in a control room monitor a wall of CCTV screens for specific events that occur rarely. This is not only cost-inefficient but also error-prone, as the number of cameras increase.

Advanced surveillance systems enrich the collection process by automatically filtering-out spurious information and presenting the operator only those parts of data which are likely to be security-relevant. Existing systems have limited capabilities; they typically perform video-based intrusion detection, and possibly object tracking, in fairly static environments.

The CASSANDRA project involves the more ambitious goal of human activity recognition in dynamic environments, in the context of automatic aggression detection. Because events associated with the build-up or enactment of aggression are difficult to detect by a single sensor modality (e.g. shouting versus hitting-someone), we propose to combine audio- and video-sensing. CASSANDRA will first address the individual sensor technologies, and subsequently, the combination of audio and visual primitives to form high-level, cognitive descriptions of scene activity, incorporating prior-knowledge about application context. The final system will be validated in realistic field tests.

- **LAYLA, the Laymans Legal Assistant: Assisting consumers in business-consumer conflicts by diagnosing, advising and mediating**

PI: Dr. R.E. Leenes, Tilburg University (Email: r.e.leenes@uvt.nl)

Alternative dispute resolution, especially supported by IT, is gaining political and judicial attention as can be witnessed by recent policy whitepapers and statements from the Minister of Justice. In practice, there are numerous barriers that hinder the use of online and IT supported alternative dispute resolution (ODR and ADR). The Layla project aims to address some of these barriers and stimulate the actual use of ODR.

The Layla project will develop online tools that help consumers in business-consumer conflicts to clarify their position, assess their rights and obligations, determine ways to resolve their conflicts and provide different online (mediated) conflict resolution methods.

Layla consists of two modules: a consumer advice module and an ODR module. The consumer advice module diagnoses the preferences and capabilities of the user and provides legal advice, both material and procedural, tailored to the user's preferences and capabilities. The module allows the user to understand their legal position in the dispute and to make informed choices on the way to resolve her conflict: resign, negotiate, mediate, arbitrate or start litigation. If the user feels her conflict can be resolved by negotiation or mediation, she, and her opponent enter the ODR advice module which tries to determine which type of ODR/ADR is best suitable for resolving the conflict.

This advice depends on an analysis of the possible implicit barriers the parties have with respect to alternative dispute resolution. If the barriers can be overcome, the parties can enter one of the ODR components that assists them in resolving their conflict.

Layla aims to improve the information dissemination to consumers, increase the number of satisfactory resolved consumer conflicts, and reduce the number of cases brought to litigation.

- **TRIGRAPH: tri-modal writer identification**

PI: Prof. dr. L.R.B. Schomaker, University of Groningen (Email: l.schomaker@ai.rug.nl)

Automatic identification of the writer of a questioned sample of handwriting enjoys a renewed interest. Although the achievable performance levels will be less impressive than is the case in DNA or iris-based person identification, the criminal-investigation process is often helped tremendously by forensic handwriting analysis. There exist three groups of basic methods concerning the features which are derived from scanned handwritten samples in forensic procedures:

- 1) Interactively measured features by human experts using a dedicated graphical user-interface (GUI) tool;
- 2) Manual or semi-automatically derived character-based features which are related to the 'allograph' subset which is being generated by each writer;
- 3) Fully automatic methods using features which are computed from a region of interest (ROI) in the image.

Of these methods, the automatic approach (especially the third group), has been treated with scepticism by practitioners within the application domain. This is due to the complexity of scanned image samples of handwriting which are found in forensic practice. However, recent advances in image processing, pattern classification and the increased availability of computing power allow for a substantial improvement of current procedures in forensic practice.

In the Trigraph project, we will contrast and combine these three approaches to writer identification in order to understand the differences between human and machine performance in finding a list of plausible writer identities for a given sample. It is to be expected that the complementary nature of the information which is derived from the three methods will lead to a considerable improvement in the performance of the writer-identification process as a whole if a suitable information-integration technique is used. The results of this project can be used in writer verification (knowing whether two pieces of handwriting are produced by the same writer) and may also have applications in historical document analysis. An example in this latter domain concerns the estimation of the writer's age for an undated sample of a well-known author for which a sufficient number of dated samples exist.

The research will be performed in a close cooperation between the AI institute of Rijksuniversiteit Groningen, the Nijmegen Institute for Cognition and Information and the Netherlands Forensic Institute in Rijswijk.

- **IPOL: Intelligence led policing. An effective law enforcement strategy or a threat to civil liberties and the principles of criminal law?**

PI: Prof. mr. Th.A. de Roos, Tilburg University (Email: T.A.deRoos@uvt.nl)

The introduction of intelligence led policing, focussing on crime risk reduction, using probabilistic profiles to stop criminals (especially terrorists) and other actors from reaching their targets, in the area of criminal law and other areas of law enforcement has been so far a mainly silent operation. This silence has been broken on the international level due to the EU - US controversy about the transfer of passenger data to the US within the framework of the CAPPs initiative and on the national level due to legal judgements refusing to use intelligence as evidence and subsequent policy measures to alter the rules of criminal law to make possible the use of intelligence in the criminal procedure in the future. The silent implementation of intelligence led policing is not desirable. It ignores fundamental questions about the applicability of the concept in the legal system, especially within the traditional criminal law system, questions about privacy and privacy regulations and questions about the added value of this new kind of police strategy in comparison with older strategies of enforcement.

Answering these questions will without any doubt benefit policy makers who have, whenever confronted with upheaval about new policing strategies, often the impression that neither the

privacy principles nor the principles of criminal law allow for any discretion when defining new answers to threats of fraud and threats to security.

The problem definition is twofold: (a) How does intelligence led policing work?, and (b) Is intelligence led policing in contradiction with the principles of criminal law and of privacy regulation? The purposes of this research are to investigate these questions and to analyze the dilemma's created by the introduction of intelligence led policing, from a legal point of view as well as a computer science point of view. How far are we in the Netherlands and Europe with this strategy? How does intelligence led policing work? Is it really new? What about its added value compared to traditional enforcement strategies? How should we judge the validity of the methods and the reliability of the results? Is intelligence led policing in contradiction with the principles of criminal law and of privacy regulation? If yes, are these principles flexible or immovable? Should we not review these principles in the light of current threats to security? How is this balancing exercise done in the United States and how far is intelligence led policing realized in this country and in other countries, such as the United Kingdom and Germany, that are said to have older experiences with it? What ICT provisions should be available if and when intelligence led policing is to be introduced? From a computer science point of view the research entails, *e.g.*, the development of methods that enable measuring, monitoring, and / or regulating (1) the quality of data sources, (2) the intraceability of personal data during data fusion, (3) the quality of data fusion, and (4) the balance between sensitivity and specificity of probabilistic profiles.

- **BEST: Intelligent disclosure of case-law using Semantic Web technology to determine BATNAs in damages disputes**

PI: Dr. A.R. Lodder, Vrije Universiteit Amsterdam (Email: lodder@rechten.vu.nl)

The judiciary is faced with enormous case loads. Therefore, alternative dispute resolution mechanisms such as mediation are very welcome. The legal domain of this project is damages disputes, which concerns a substantial number of the court cases. A decision to either go to court or to mediate is ideally based on a well-informed choice. Currently the necessary information to make such a decision is often lacking.

One of the aims of this project is to provide litigants with information about the expected outcome of a court proceedings, in (legal) negotiation literature called a BATNA (Best Alternative To a Negotiated Agreement). This information will be provided through intelligent disclosure of case-law using Semantic Web technology. The problem of finding relevant case-law in order to determine a BATNA can be reduced to the problem of finding closely related documents in a large, semi-structured collection.

ANITA: Law, technology and practice: exchanging crime-related information within a new legal regime

Mr. H.H. Kielman (Email: hugo@kielman.net) and Mr. W. Koelewijn (Email: wouterkoelewijn@tiscali.nl), Leiden University, Faculty of Law

Penal law is strongly influenced by catastrophes such as the terrorist attacks in Madrid and the murder of Theo van Gogh. New laws are introduced to help prevent terrorist attacks. One of the measures introduced is the further facilitation of the exchange of data among police institutions. A proposal for a new act, called the 'Wet Politiegegevens' ('Police data act') introduces a new regime for the use, storage and exchange of crime-related information. The new act provides more opportunities to exchange police data. However, to establish an effective and efficient regime, legislation simply does not suffice: lack of knowledge about valid legal norms, differences in ICT-systems, and the danger of information leaks may hamper information exchange among police regions.

The ANITA project proposes an agent-based solution for these problems. By taking advantage of such a distributed architecture, the above-mentioned problems can be solved. First, knowledge about valid legislation may be propagated through its implementation in agents, who may, *e.g.*, check whether certain queries in a police register are legitimate. Second, a distributed framework may interface effectively with different systems present in police practice. Third, sophisticated knowledge sharing protocols can prevent information leaks under certain circumstances.

In this lecture, we will indicate what consequences the changes in the legal regime will have for maintaining a balance among four interests related to the use of crime-related information: the proper execution of police tasks, the protection of specific police investigations, the protection of informants and the protection of citizens' privacy. An appropriate legal regime in combination with technological innovations can be helpful in establishing a more legitimate and verifiable practice of information exchange. We will explain the way in which the interaction between law, technology and the practice of data exchange in this specific domain can be examined empirically, and how this empirical work will feed the functional specification of the ANITA system.



MR. HUGO KIELMAND AND MR. WOUTER KOELEWIJN

I-CATCHER: Intensive-Care Access to Terminology & Course of Health Exploration and Retrieval

Dr. A. Abu-Hanna, Academic Medical Center (AMC) at the University of Amsterdam, Department of Medical Informatics (Email: a.abu-hanna@amc.uva.nl); Dr. M.C.A. Klein, Vrije Universiteit Amsterdam, Artificial Intelligence Section (Email: michel.klein@cs.vu.nl)

Intensive Care (IC) is a complex, expensive form of care. Electronic Patient Records (EPR) containing information about patients and outcomes are now available that can partially support primary care and the evaluation of its quality. In the I-Catcher research project we propose two urgent related extensions to the EPR functionality: 1. the efficient accessibility of IC terminological knowledge using emerging semantic web technology, in particular mapping one ontology onto another, and 2. the extraction of temporal knowledge, in particular learning predictive models based on temporal data collected daily in the IC unit. In this talk we will report on the progress of I-Catcher in these two research lines.

In the terminology accessibility line we investigate how patient data can be related to other data sources by aligning local terminologies to other (standardized) terminologies or ontologies. Even when the terminologies cover the same domain, this often is a non-trivial task. The task is even more complicated when the terminology does not contain much structure. We are developing an approach in which we use knowledge from one terminology to improve the structure in another terminology. This is done in a two-step process: in the first step lexical analysis techniques are used to find out to what category a term belongs, which is then used to create structure-rich concept descriptions in the second step. The first results show that, when compared to approaches based on syntactic analysis, the recall can be greatly improved without sacrificing the precision.

In the temporal extraction line we investigate how to exploit the sequential organ failure assessment (SOFA) scores in predicting patient death. The SOFA scores, an information instrument only recently introduced in the IC, quantify the degree of derangement of organs for each patient on each day of the IC stay. We investigate two approaches. In the first approach we abstract from the temporal dimension by aggregating temporal information in summary statistics in order to use off-the-shelf machine learning methods. In the second approach we discover frequent temporal patterns, assess their suitability for prediction, and suggest a method for the integration of temporal patterns within the logistic regression models that are in use today.



DR. AMEEN ABU-HANNA

AUTHENTIC: Bridging two worlds

Drs. I.E. Berezhnoy, Universiteit Maastricht, Institute for Knowledge and Agent Technology
(Email: I.Berezhnoy@cs.unimaas.nl)

The assessment of paintings is largely performed by human visual analysis. Undoubtedly, the assessments of skilled art experts are of great value to the domain of the visual arts. However, inevitably human judgments are subjective and prone to error. Throughout the history of visual arts, experts judging the authenticity of paintings have made mistakes and revised their opinions.

We claim that recent advances in artificial intelligence (in particular in image recognition and machine learning) can be used to support the art expert while analyzing visual art. Quantitative and objective analysis enhances the quality and reliability of the visual assessment.

The main goal of the Authentic project is to take steps towards the development of a full-fledged toolbox for the automatic analysis of paintings. The development of such a toolbox helps to bridge the gap between two worlds - the world of the cultural heritage and the world of computer science.

In the presentation the latest developments in the Authentic project that focuses on the digital analysis of the paintings of Vincent van Gogh will be outlined.



DRS. IGOR BEREZHNOY

NARRATOR: Nominalizations; an example of detailed semantic analysis for information retrieval in Narrator

Drs. H.G.B. Reckman, Leiden University, The Leiden Centre for Linguistics
(Email: h.g.b.reckman@let.leidenuniv.nl)

The Narrator project aims at the development of an information system that discloses personal illness narratives to facilitate patients in finding relevant experiences of their fellow patients, taking into account their personal profiles and information need. The first prototypes of the system will be for breast cancer patients.

This talk will be about the language-technological component of the project, where the intention is to combine different methods of information retrieval. In addition to statistics-based techniques, such as Latent Semantic Indexing (LSI), detailed semantic analysis is used. The semantic representations of the narratives are generated by a parser (Delilah) and searched in by an inference tool.

We will argue that detailed semantic analysis gives possibilities for very precise information retrieval. This is expected to be crucial in finding the narratives that are most relevant for a particular patient, among many narratives about roughly the same subject. To illustrate this point we will discuss in more detail a study of nominalizations, as an example of how detailed semantic analysis can be used to extract valuable information that would otherwise go unnoticed.

A nominalization is a noun derived from a verb, e.g. *operatie* 'surgery/operation' from *opereren* 'to operate'. Both the verb and the noun refer to an event that normally involves an operating person and a person being operated on. In the present approach all these elements of the meaning and the relations between them are spelled out explicitly in the semantic representation. This makes it fall out rather easily, for example, that *een operatie ondergaan* 'to undergo an operation/surgery' and *geopereerd worden* 'to be operated on' are the same thing. Since in detailed semantic analysis negation can also be analyzed fairly adequately, an inference tool can select, on the basis of the semantic representations, exactly those narratives in which a patient tells she had surgery, or, of course, exactly those in which the patient tells she didn't have surgery, depending on what the user asks for. In statistical approaches the equivalence of different phrasings, as well as the scope of negation are very difficult to capture.

Finally we will shed some light on how we plan on integrating different methods. Since searching with an inference tool in a large amount of material is problematic, statistical methods will be used to narrow down the search space. The semantic representations, which will be available anyway, because we need them for the reasons discussed above, may then also be used as a basis for statistical information retrieval.



DRS. HILKE RECKMAN

EIDETIC: Context-based object detection for CIBR

Ir. N.H. Bergboer, Universiteit Maastricht, Institute for Knowledge and Agent Technology
(Email: N.Bergboer@cs.unimaas.nl)

In the current electronic era, natural images and images from cultural heritage abound. Content-based image retrieval (CIBR) provides a way to automatically search for images based on their content, without requiring each image to be labeled by a human. Automatically generating catalogues of objects in images ---object detection--- provides an efficient way to search. Within the EIDETIC project, we develop methods for object detection.

The visual detection and recognition of objects is facilitated by context. In this talk, we present two types of learning methods for realizing context-based object detection in paintings. The first method is called the gradient method; it learns to transform the spatial context into a gradient towards the object. The second method, the context-detection method, learns to detect image regions that are likely to contain objects. The accuracy and speed of both methods are evaluated on a face-detection task involving natural and painted faces in a wide variety of contexts. The experimental results show that the gradient method enhances accuracy at the cost of computational speed, whereas the context-detection method optimises speed at the cost of accuracy.

The different results of both methods are argued to arise from the different ways in which the methods trade off accuracy and speed. We conclude that both the gradient method and the context-detection method can be applied to reliable and fast object detection in paintings and that the choice for either method depends on the application and user constraints.



IR. NIEK BERGBOER

MIA: Towards Better Health Care with Agents

Dr. F.J. Wiesman, Academic Medical Center (AMC) at the University of Amsterdam, Department of Medical Informatics (Email: f.j.wiesman@amc.uva.nl)

Hospital workers are faced with a multitude of decisions to be made and resulting actions to be performed based on a variety of information sources. The quality of care may be improved if hospital workers are freed from routine actions and get information that is geared to the medical context.

In the *MIA project* we investigate how software agents can aid hospital workers. Our goal is to provide agent support for three important directions: (a) automatic feedback to actions that are not in line with clinical-practice guidelines, (b) automatic retrieval of medical literature, and (c) scheduling patient treatment. These directions are addressed in three subprojects, each with its own agent. Each of the three agents may improve its behaviour by taking into account results of the other two agents.

In the *feedback* subproject MI/UM investigates how guideline systems can take the intentions of their steps into account. In medicine the use of guidelines (i.e., protocols) becomes increasingly widespread. Guidelines support physicians in managing categories of patients to spur best practices in medical care on the basis of evidence. Studies have shown the positive effects of computerized support. When the support consists of feedback, that is, a reaction of the computer on a physician's action that conflicts with the guidelines, it is essential to allow the physician to deviate from the literal protocol as long as he or she follows the *intention* of the protocol. Hence, MI/UM focuses in the representation of guidelines on intentions and goals rather than actions.

The *retrieval* subproject (IKAT/UM) focuses on available information that is stored in free text. The goal of this subproject is to retrieve medical literature that is highly relevant with respect to a specific patient at hand. Research has shown that especially the part of formulating the information needs is too time consuming for physicians. Therefore, the main goal of this subproject is to investigate whether it is possible to support physicians in formulating these implicit information needs by making them explicit automatically. Current results include an inventory of generic information needs, in the form of information-need templates. The templates can be instantiated with data from the electronic patient record.

CWI's *scheduling* subproject investigates the problem of patient treatment scheduling. Scheduling the complex treatment plans of patients requires coordination between all the different autonomous departments involved. Due to the dynamic nature of a hospital, any approach must be efficient, online, and flexible. In cooperation with medical experts CWI investigates the use of autonomous software agents negotiating with each other in order to make appointments (or reschedule them) for patient treatment. An initial model has been specified which distinguishes two kinds of agents: (a) the department agent, which manages the resources of one department, and (b) the treatment agent, which schedules the treatment of one patient as well as possible by making appointments with department agents and exchanging appointments with other treatment agents.



DR. FLORIS WIESMAN

Appendix 3: ToKeN projects and related posters

EIDETIC: Intelligent Content-based Image Retrieval

PI: Prof. dr. E.O. Postma (Email: postma@cs.unimaas.nl)

- 1 *Contextually Validated Object Detection*
N.H. Bergboer, E.O. Postma, H.J. van den Herik
- 2 *Method to detect a gist change*
X.A.N.D.R.A. van Montfort, H.P. de Greef, D.G. Bouwhuis
- 3 *Human-Centered Content-Based Image Retrieval*
E.L. van den Broek, E.M. van Rikxoort, P.M.F. Kisters, T.E. Schouten, L.G. Vuurpijl
- 4 *Enhancing Digitised X-ray Images of Etchings for Authentication Purposes*
M. van Staalduinen, J.C.A. van der Lubbe, E. Backer

I2RP: Intelligent Contact Retrieval and Presentation in Public Historical Multimedia Databases

PI: Prof. dr. L.R.B. Schomaker (Email: schomaker@ai.rug.nl)

- 5 *Rhetorical Video Editing*
S. Bocconi, L. Hardman, B. Arsenijević, C. Cremers, G. Kempen

DUMPERS: Distributed user modeling and Exploration in Personalized Recommender Systems

PI: Dr. M.W. van Someren (Email: maarten@swi.psy.uva.nl)

- 6 *Discovering Stages in Web Navigation*
V. Hollink, M. van Someren

CHIME: Cultural Heritage in an Interactive Multimedia Environment

PI: Prof. dr. H.L. Hardman (Email: Lynda.Hardman@cwi.nl)

- 7 *Multimedia Authoring Support in Four Phases: harmonizing domain and discourse knowledge*
L. Hardman, F. Nack, K. Falkovych
- 8 *Enabling Museum Data Integration with RDF/OWL*
G. Schreiber, M. van Assem
- 9 *Discovery Service for User Models in a Multi-Application Context*
P. De Bra, L. Aroyo, V. Chepegin

AUTHENTIC: Knowledge discovery and disclosure for visual art: authentication and dating of graphic art and paintings

PI: Dr.ir. J.C.A. van der Lubbe (Email: J.C.A.vanderLubbe@its.tudelft.nl)

- 10 *Quantitative Analysis of Brushstrokes and Spatial Characteristics in Visual Art*
I.E. Berezhnoy, E.O. Postma, H.J. van den Herik
- 11 *Decision Trees for Paintings Classification*
A.I. Deac, J.C.A. van der Lubbe, E. Backer

ANITA: Administrative Normative Contact Transaction Agents

PI: Dr. mr. C.N.J. de Vey Mestdagh (Email: C.N.J.de.Vey.Mestdagh@rechten.rug.nl)

- 12 a) *The Story: Gossiping* and b) *The Protocols*
W. Teepe

VINDIT: Combining visual and textual Contact for Contact retrieval

PI: Prof. dr. E.O. Postma (Email: postma@cs.unimaas.nl)

- 13 *Sensory-motor Coordination in Gaze Control*
G. de Croon, E.O. Postma, H.J. van den Herik

I-CATCHER: Intensive-Care Acces to Terminology & Course of Health Exploration & Retrieval

PI: Dr. A. Abu-Hanna (Email: a.abu-hanna@amc.uva.nl)

No poster submitted.

MIA: Medical Information Agent

PI: Prof. dr. ir. A. Hasman (Email: hasman@mi.unimaas.nl)

14 *The Information-Need Digestive System*

L.M.M. Braun, F. Wiesman, H.J. van den Herik, A. Hasman

15 *Efficient Patient Treatment*

I.B. Vermeulen, S.M. Bohte, D.J.A. Somfun, J.A. La Poutré

16 *The intentions of guidelines and decision support*

A. Latoszek-Berendsen, J. Talmon, A. Hasman, H.J. van den Herik

DIME: Distributed Interactive Medical Exploratory for 3D Medical Images

PI: Prof. dr. ir. J.H.C. Reiber (Email: j.h.c.reiber@lumc.nl)

17 *A Problem Solving Environment for Image-Based Computational Hemodynamics*

L. Abrahamyan, J.A. Schaap, A.G. Hoekstra, D. Shamonin, F.M.A. Box, R.J. van der Geest, J.H.C. Reiber, P.M.A. Slaat

TIMEBAYES: Building and Using Temporal Bayesian Models in a CPR Setting

PI: Dr. P.J.F. Lucas (Email: peterl@cs.kun.nl)

18 *Exploiting Causal Independence in Large Bayesian Networks*

R. Jurgelenaite, P. Lucas

NARRATOR: Narrative Disclosure of Health Care Knowledge

PI: Dr. P.J. Toussaint (Email: p.j.toussaint@lumc.nl)

19 *Narrator: Patients' requirements regarding internet applications that disclose fellow patients' illness stories*

R.I. Overberg, P.J. Toussaint, L.F. Wolf, J.H.M. Zwetsloot-Schonk

20 *Access for all by cognitive engineering*

H.H. Nap, H.P. de Greef, D.G. Bouwhuis

EVIDENCE: Making sense of evidence: software support for crime investigations

PI: Prof. dr. mr. H. Prakken (Email: henry@cs.uu.nl)

No poster submitted.

DALE: Data Assistance for Law Enforcement

PI: Dr. W.A. Kusters (Email: kusters@liacs.nl)

No poster submitted.

ACCESS: Agent-based Criminal Court Electronic Support Systems

PI: Prof. dr. F.M.T. Brazier (Email: frances@cs.vu.nl)

No poster submitted.

CASSANDRA: Context-Aware SenSing for AggressioN Detection and Risk Assessment

PI: Prof. dr. D.M. Gavrilu (Email: gavrilu@science.uva.nl)

21 *CASSANDRA Context-Aware SenSing for AggressioN Detection and Risk Assessment*

D.M. Gavrilu, T. Andringa

LAYLA, the Laymans Legal Assistant: Assisting consumers in business-consumer conflicts by diagnosing, advising and mediating

PI: Dr. R.E. Leenes (Email: r.e.leenes@uvt.nl)

No poster submitted.

TRIGRAPH: tri-modal writer identification

PI: Prof. dr. L.R.B. Schomaker (Email: l.schomaker@ai.rug.nl)

No poster submitted.

IPOL: Intelligence led policing. An effective law enforcement strategy or a threat to civil liberties and the principles of criminal law?

PI: Prof. mr. Th.A. de Roos (Email: T.A.deRoos@uvt.nl)

No poster submitted.

BEST: Intelligent disclosure of case-law using Semantic Web technology to determine BATNAs in damages disputes

PI: Dr. A.R. Lodder, Vrije Universiteit Amsterdam (Email: lodder@rechten.vu.nl)

No poster submitted.