Work with Knowledge for Support of e-Government

Tomáš Sabol, Ján Paralic
Technical University of Košice, Slovakia
Tomas.Sabol|Jan.Paralic@tuke.sk

http://esprit.ekf.tuke.sk/webocracy/

Overview of the presentation

PART I

- e-Government – basic terms, services for 21st century, aims
- Application areas of e-Government
- Critical success factors
- Good practise in e-Government
- Roadmap for implementing knowledge management in e-Government
- Webocracy project – basic data,consortium, expected achievements

PART II

- *WEBOCRAT* system functional overview and basic architecture
- Current status of the Webocracy project, pilot applications
- *WEBOCRAT* core technology - knowledge modelling and its exploitation for customisation support and retrieval
- Knowledge discovery in texts (KDT), text data mining
- Text data mining tasks - Clustering/visualization, Association rules, Classification models
- Exploitation of KDT in Webocracy project

2

PART I

3

Glossary of Basic Terms

e-Government: The interaction between government and citizens over the Internet

(Evolving from: merely publishing or disseminating government information electronically ? online interactions, and ? transactions between government and citizens)

e-Democracy: Increased participation of citizens in democratic processes using ICT
(? deliberative democracy – opinion polling, e-voting, discussion, on-line meetings etc.)

4

e-Government services for the 21st century

E-government should lead to a new / transformed relationship between government and citizens enabled by computer networks

- Electronic service delivery (and using new digital channels) - a key source of innovation
- Not only doing the things better, but also doing new things (the re-invention of government)

Motto: „*It is better to be **on-line** than **in line**, in front of a government office*“

5

e-Government Aims

? **To improve public services** - for the benefit of citizens:

- bring PA closer to citizens
- available 24 hours a day
- faster, more convenient
- more personalised
- more joined-up (service provision across departmental boundaries) – effective citizen services are delivered independently of organizational structure
- more responsive – if ensured that government electronic service delivery is driven by the use that citizens make of it

6

e-Government (as designed by EC)

- **Interactive**
  - enhancing government, enhancing participation in democracy
  - efficiency & simplicity of services by joining-up administrations
- **Inclusive**
  - preventing digital divide, support the disadvantaged and people with special needs
  - promoting public Internet and multimedia access points
- **Entrepreneurial**
  - supplying online transactions for business
  - supporting public/private partnerships
- **Multilingual/Multicultural**
  - stimulating cross-border use of services and information
  - encouraging public services in more than one language

7

Application Areas of e-Government:

- Tax returns - electronic filing of tax returns and online contact centres
- VAT – online VAT registration and returns, trade statistics
- Online public procurement
- Benefits (e.g. pension) – enquiries, advice, benefit applications and payments online
- Online voting, online opinion polls (start with local polls / elections)
- Companies' registration – electronic registration of companies

8

Application Areas of e-Gov: (2)

- Online delivery of driving licence applications, car tax renewals, driving test applications,
- Electronic land registration (Land Registry)
- Health (advice of health and healthy living)
- Learning and Work bank - online services for citizens looking for jobs or training opportunities
- Culture online (galleries, museums, ...)

9

Application Areas of e-Gov: (3)

- Courts - legal advice and information, transactions between the public and the courts (e.g. civil claims)
- Environmental services
- Official information (index of documents)
- Online "policy panels" - involve citizen groups in policy formulation, create electronic forums, e-communities
- . . .

10

Critical Success Factors

- A clear **vision** of government what is trying to achieve
- "Top Level Champion", leadership
- Commitment to change
  - Technical
  - Cultural
- Genuine commitment from all parties involved
- Business Process Reengineering
- Understanding the benefits of the changes

11

Critical Success Factors (2)

- Everyone should have access to the Internet (fast, always on, everywhere, intelligent, easy, trusted)
- Creating a mixed economy in the electronic delivery of government services – i.e. "business-friendly government" (e.g. private and voluntary sector organisations able to access the information and databases that they need in order to deliver services)
- Legislation

12

Good Practice in e-Government

- User-centred design and involvement
  - Life-event approach (birth, going to school, marriage, business-start, moving, …
  - Target group orientation
    - differences in skills,
    - different user groups (citizens, professional intermediaries, companies , administrations, policy)
    - Different roles of individuals (citizens, migrants, tourists, permanent customer, job seeker etc.)
  - Usability studies, feedback function
- Navigation and search (easy navigation, detailed search function, …)

13

Good Practice in e-Government (2)

- Multi-channel delivery of services (Internet, WAP, Public access points, Call centres, …)
- Organisation, work & skills
  - Re-engineered organisation
  - Workflow systems
  - New skills – training, e-learning
  - Improving working condition, tele-working, …
- Public-private partnership – outsourcing, financing (EU funds, business models, …)

14

Good Practice in e-Government (3)

- Social inclusion
  - Public access points (incl. support)
  - Special programmes for target groups (people with disabilities, elderly, teenagers, …)
- Regional development
  - Cooperation between central & local government bodies
  - Economic development for remote areas
- Trust & Security – availability, integrity, authenticity, confidentiality, …

15

Priority areas (where to start?):

(Make the most difference to the citizen):

- Where the transaction volumes and user numbers are high
- Where there is interaction not just publication
- Where services can be joined-up

Important requirements:

- Technology integration (expensive!) – with the legacy system
- Process integration (from the perspective of your citizens')

16

Other sources / IST Projects

- IST-1999-29088
- Providing Innovative Service Models and Assessment - "PRISMA"
- http://www.prisma-eu.net

- IST-2000-26224
- Best eEurope Practices – "BEEP"
- http://www.beep-eu.org
- Knowledge base of case studies

17

Roadmap for Implementing Knowledge Management (in e-Government)

Adapted from Tiwana A.: The KnowledgeManagement Toolkit

1. Analysis of existing infrastructure
   - ✍ Groupware, intranet, extranet solutions
2. Aligning knowledge management / e-government and organisational strategy
   - ✍ Create a clearly articulated link between KM/e-Government and organisational strategy
3. Knowledge-based e-government architecture and design
   - ✍ Analyse components of the "info-structure", collaborative platform, knowledge sources, costs vs. added value

18

Roadmap for Implementing Knowledge Management (in e-Government) (2)

4. Knowledge analysis and audit
   - Create a clearly articulated link between KM/e-Government and organisational strategy
5. Design KM/e-government team
   - Identify key stakeholders, critical points of failure, create a balanced team
6. Create the system blueprint
   - Plan for building and incrementally improving KM e-gov system
7. Develop the e-government system

19

Roadmap for Implementing Knowledge Management (in e-Government) (3)

8. Pilot testing and deployment
   - Understand need/scope of the system deployment, identify failure points, use Result-Driven Incremental methodology
9. Reward structures, change management
   - Encourage use, gain employee support, training
10. Metrics to measure impact
   - Measure the impact

20

Basic Data

- Whole name: **Web Technologies Supporting Direct Participation in Democratic Processes**
- *IST 1999-20364*
- Duration: *39 months*
- Start date: *October 1, 2000*
- Total effort: *348 person-months*
- Consortium: *8 partners from 4 countries*
- Total costs: *1,773,079 EUR*
- EC funding: *1,499,915 EUR*

21

Consortium

- Technical University of Košice, SK – coordinator
- University of Wolverhampton, UK
- University of Essen, GE
- JUVIER s.r.o., SK
- CITEC Information Oy Ab, Finland
- Local authority of Košice - Tahanovce, SK (LATA)
- Local authority of Košice - Darg. hrdinov, SK (LAFU)
- Wolverhampton City Council, UK (WCC)

22

Project Aim

- Provide citizens with innovative communication, access, and opinion polling system in order to increase:
  - participation of citizens in democratic processes
  - transparency and accessibility of public administration

23

Expected Achievements (Technical)

- WEBOCRAT system:
  - Automatic routing of messages
  - Support for information publishing
  - Easy access to public administration information
  - Support for competitive tendering
  - Discussion forums
  - On-line opinion polling
  - Easy navigation and browsing through information

24

Expected Achievements (Methodological)

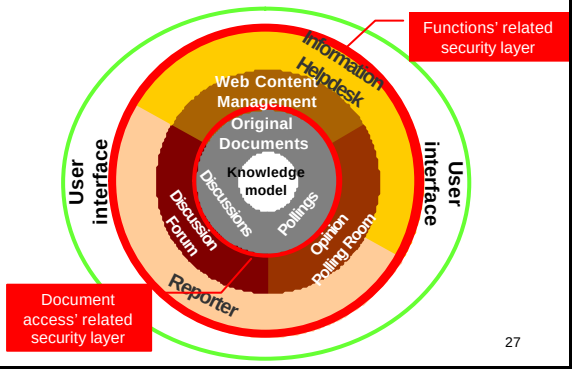
- To develop an overall methodological framework for using WEBOCRAT-like systems effectively in order to:
  - improve access to PA
  - increase quality of services provided of PA
  - support organizational learning

25

PART II

26

WEBOCRAT - Layered Structure



27

System Functional Overview (1)

- Knowledge Management module (1st layer)
  - Design and management of knowledge model
  - Querying the ontological knowledge model
- Document space (2nd layer)
  - Published documents expected to be read by different groups of users (Original documents)
  - Users' contributions to discussions on different topics (Discussions)
  - Records of users' opinions about different issues (Pollings)

28

System Functional Overview (2)

- Modules for management of particular types of documents (3rd layer)
  1. Web Content Management module
     - Linking of documents to elements of a knowledge model
     - Publishing of documents and access to them
  2. Discussion Forum module
     - Enables users to contribute to discussions they are interested in
     - Read contributions submitted by other users
  3. Opinion Polling Room module
     - Performing opinion polling on different topics

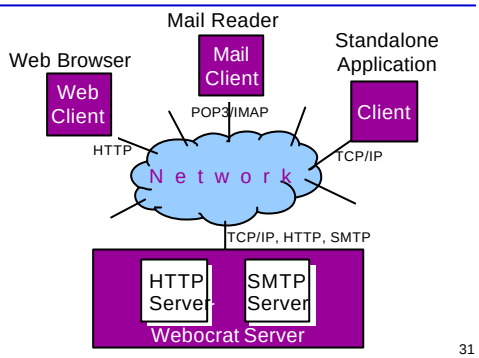
29

System Functional Overview (3)

- Information retrieval (4th layer)
  1. Citizens' Information Helpdesk module
     - A search engine based on the indexing of stored documents
     - Concepts from knowledge model and attributes of documents
  2. Reporter module
     - Definition and generation of different reports concerning information stored in the system
     - Semi-automatic linking of documents to knowledge model
     - Profile management, alerting services

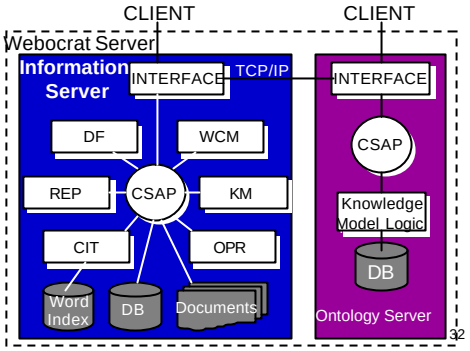
30

Basic Architecture of the System



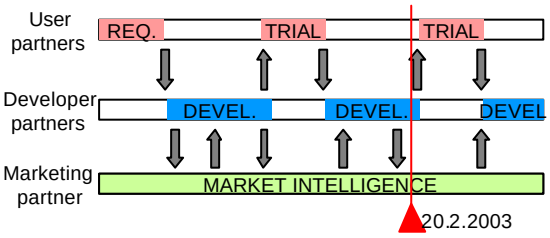
31

Coarse-Grain Server Structure



32

Current State of the Project (1)



33

Current State of the Project (2)

Finished:
- Specification and analysis of user requirements
- Webocrat architecture design
- Design of particular Webocrat modules
- Implementation of all Webocrat modules
- First proposal of methodology for introduction of Webocrat-like systems
- First pilot application has been performed and evaluated

Runs:
- Resulted suggestions for revisions and changes in tested modules are being implemented
- Second trial is being prepared

34

Pilot Applications

- User partners are responsible for realisation of pilot applications
  - Wolverhampton City Council
  - Mestská cast Košice - Dargovských hrdinov
  - Mestská cast Košice - Tahanovce
- Pilot application run in two phases:
  1. In first pilot application DF, WCM and OPR modules have been tested (***May – July 2002***)
  2. In the second phase the whole Webocrat system will be tested, but mainly KM, CSAP, CIH and REP modules (***March – May 2003***)

35

Pilot Application in Wolverhampton

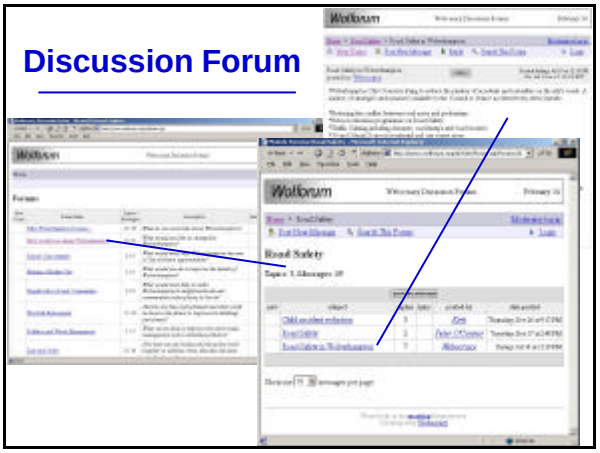
www.wolforum.org

Web Content Management

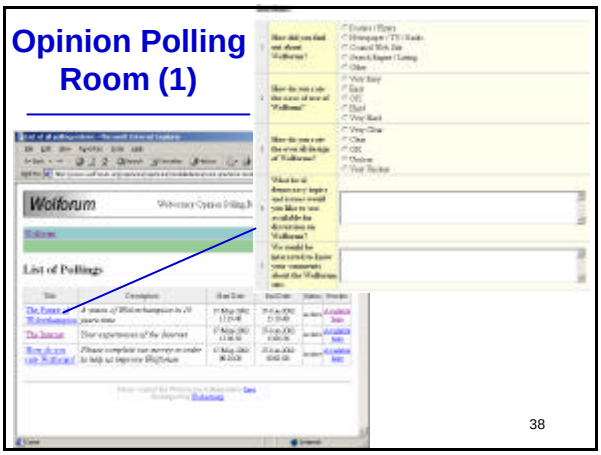
Discussion Forums

Opinion Polling Room

Discussion Forum

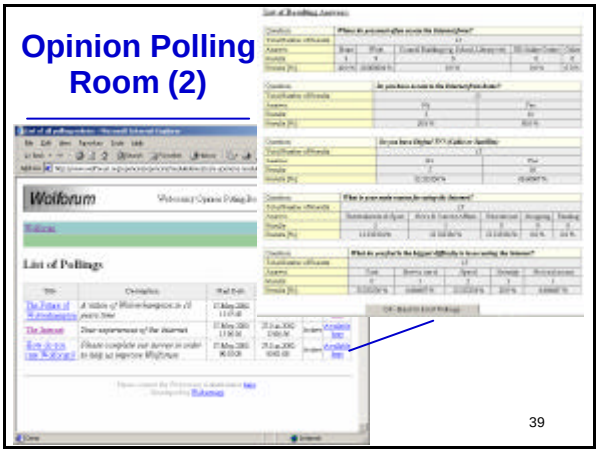


Opinion Polling Room (1)



38

Opinion Polling Room (2)



39

Pilot Application in Košice – Darg. hrdinov

Opinion Polling Room

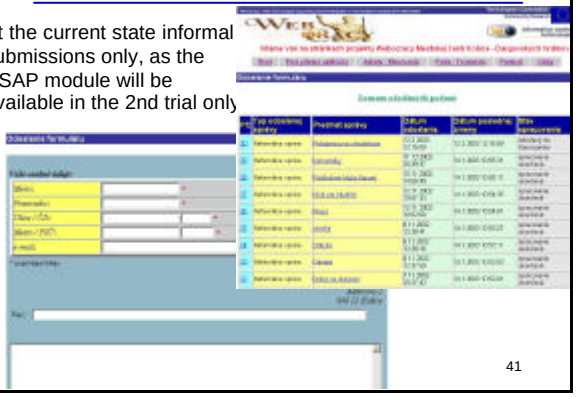
www.kosice-dh.sk

Electronic submissions from citizens and the possibility to track their processing status

40

Electronic Submissions and their Tracking

At the current state informal submissions only, as the CSAP module will be available in the 2nd trial only



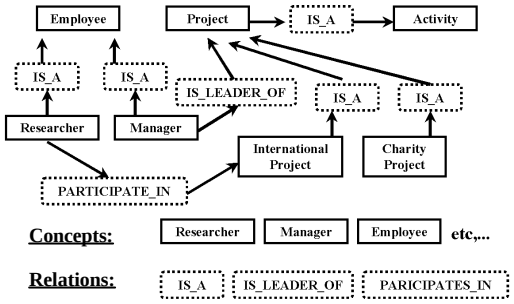
41

Ontology

- Definition by *Gruber*: "A set of definitions of content-specific knowledge representation primitives consisting of domain-dependent classes, relations, functions, and object constants"
- *Chandrasekaran* understands ontology as a representation vocabulary typically specialized to some domain. He suggests two purposes of it:
  - To define most commonly used terms in a specific domain, thus building a skeleton
  - To enable knowledge sharing and re-using both spatially and temporally

42

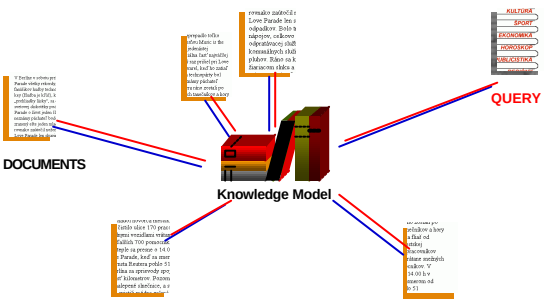
Ontological knowledge model



WEBOCRAT Core Technology

- **Ontological knowledge modeling** represents technological background on which **documents** manipulation in the *WEBOCRAT* system is built
- **Document**: any piece of information which can be stored as a computer file
- **Ontological knowledge model** plays the role of a context for documents.
- **Context of the documents** is defined using links between a document and relevant parts of the knowledge model

Conceptualisation and Retrieval



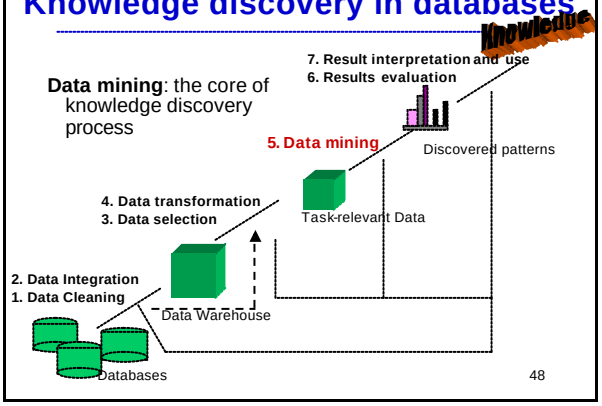
Customisation Support (1)

- **Personal profile definition**
  1. Domain(s) of interest
     - By means of elements from a knowledge model (subparts of this model)
     - These represent for the user interesting topics
  2. Alerting policy
     - About what type events in the system the user whishes to be alerted (e.g. submission of a discussion contribution, publishing a document, etc.)
     - In which way the alert should be delivered to the user

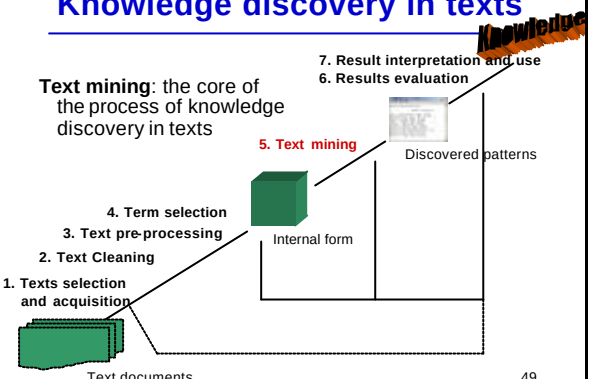
Customisation Support (2)

- **Forms of alerting**
  1. Notification using e-mail services
     - on event-per-event basis, i.e. he/she receives an e-mail message for each event he/she is alerted on
     - an e-mail digest format – based on time intervals and/or the size of e-mail messages
  2. Personal newsletter
     - The personal newsletter has the form of a document published in the publishing space
     - This document is generated by the system and contains links to all those documents, which may be of interest for the user
     - The document is generated when user logs in

Knowledge discovery in databases



Knowledge discovery in texts



Representation of textual documents

- **Index term** t_j is a document word whose semantic helps in remembering documents' main theme
- **Document** d_i is represented by a set of weights w_{ij} ? 0 associated to each pair (d_i, t_j)
- **Boolean model**: w_{ij} ? {0,1}; query is a subset of index terms linked by logical connectives *not*, *and*, *or*
- **Vector model**: $w_{ij} = tfidf(d_i, t_j) = N^{dj,ti}.\log(C/Nt_j)$; query is represented like a document, i.e. set of weights
- **Probabilistic model**: w_{ij} ? {0,1}; query is a subset of index terms (probabilistic description of ideal answer set)

Preprocessing of textual documents

- **Tokenisation** (identification of lexical units)
- Elimination of **stopwords**
- **Stemming** (easy for English, not trivial for e.g. Slavic languages)
- **Term selection**:
  - Unsupervised methods, e.g. document frequency threshold, or Latent Semantic Indexing - see also: Kostial I.: Using Latent Semantic Indexing for intelligent information retrieval (short paper at this conference)
  - Supervised methods, e.g. information gain

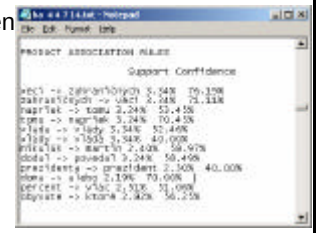
Clustering/visualization

Association rules

- Associations between concepts in the form B => H [sup., conf.] where

  sup.(B=>H) = P(B? H)

  conf.(B=>H) = P(H|B)



E.g.:

A. Term1 AND Term2 => Term3
B. Concept1 (AND Concept2) => Concept3
C. Word1 AND Word2 => Concept3

Classification models

- There are currently implemented 2 different approaches:
  - Bayesian classifier
  - Rule-based approach
- More about this and other approaches in Bednár P. - Hudák S.: Využitie kategorizácie textov pre linkovanie na ontológiu (short paper at this conference)

Exploitation of KDT in Webocracy

- **Clustering/visualization** – supporting tool within initial phase of knowledge model creation (large number of documents available)
- **Association rules** – automatic support for management of the knowledge model
- **Classification models**:
  - Tool to guide user at annotating new documents (semi-automatic linking of documents to knowledge model)
  - Retrieval of documents relevant to user query

55

Thank you for your attention!

http://esprit.ekf.tuke.sk/webocracy/

56