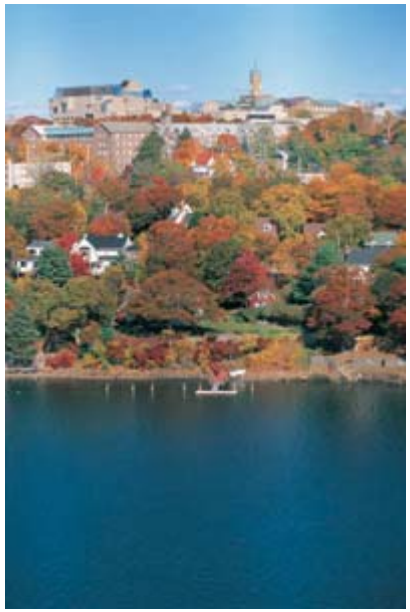
Graduate Studies in Computer Science at Dalhousie University

Evangelos Milios
Faculty of Computer Science
Dalhousie University
www.cs.dal.ca/~eem

Dalhousie U. Facts

- Founded in 1818
- The smallest Medical/Doctoral university in Canada
  - Medical school
  - Law and Business schools
  - Engineering
- World class
  - Oceanography
  - Biology
  - Medicine
  - Sciences



- Member of the G-13 research intensive universities in Canada
- Regional Research Hub for Atlantic Canada

Faculty of Computer Science

Faculty of Computer Science

- Established in 1997
- Strengths in:
  - Information retrieval, text mining
  - Health informatics & Knowledge management
  - Bioinformatics
  - Human-computer interaction, visual computing
  - Computer networks, network management, intrusion detection
  - Algorithms, graph theory, parallel computation

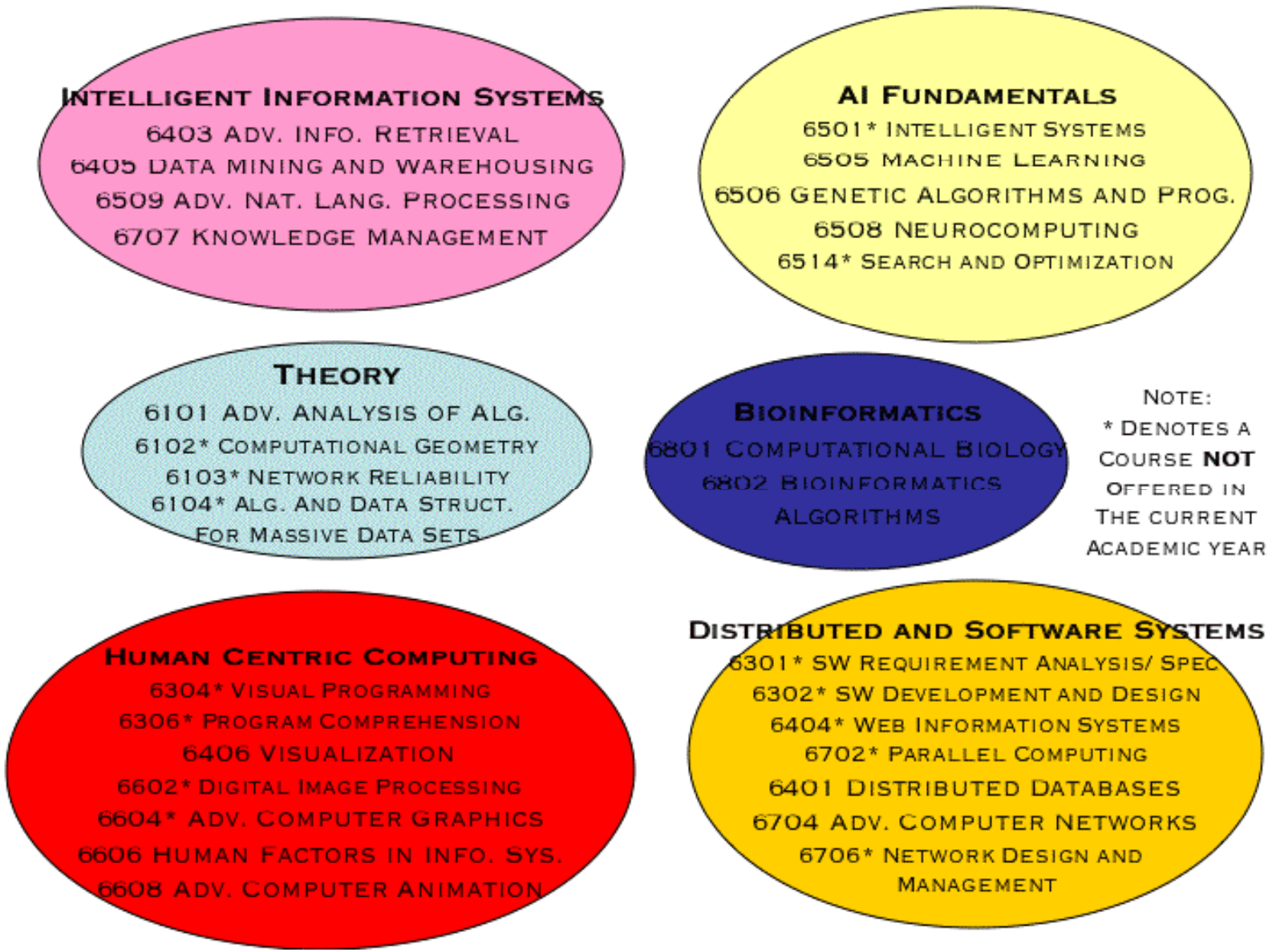
Interdisciplinary outlook

- Master's degrees in:
  - Computer Science
  - Health informatics (with Medicine)
  - Electronic commerce (with Business and Law)
  - Bioinformatics (with Biology)
- Joint research projects with
  - Mathematics
  - Engineering
  - Medicine
  - Business
  - Biology

Coursework

- Number of courses depends on the degree program
- Breadth requirement must be satisfied by both Master's and PhD students
  - For PhD students, all courses taken for a Master's degree count

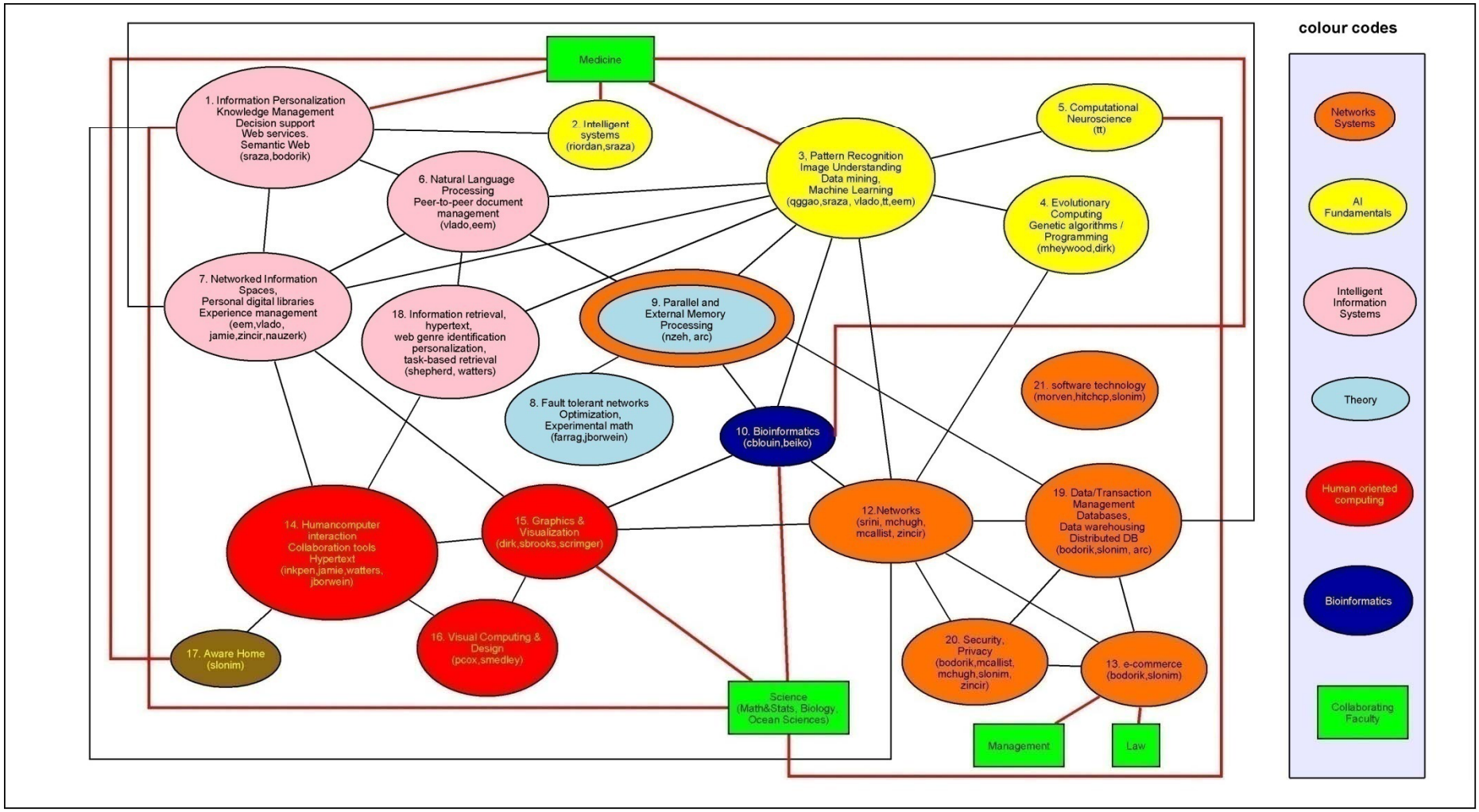
Breadth bubble diagram



Breadth Requirement

- ONE course from FOUR different research areas of the breadth bubble diagram
- Only courses with a CSCI number may contribute
- OUTSIDE of the above FOUR courses
  - Up to TWO grad courses from another discipline, with prior approval
  - # of 4^{th} year CSCI courses +
    # of grad courses from another discipline
    ≤ 2

Research overview



Research snippets

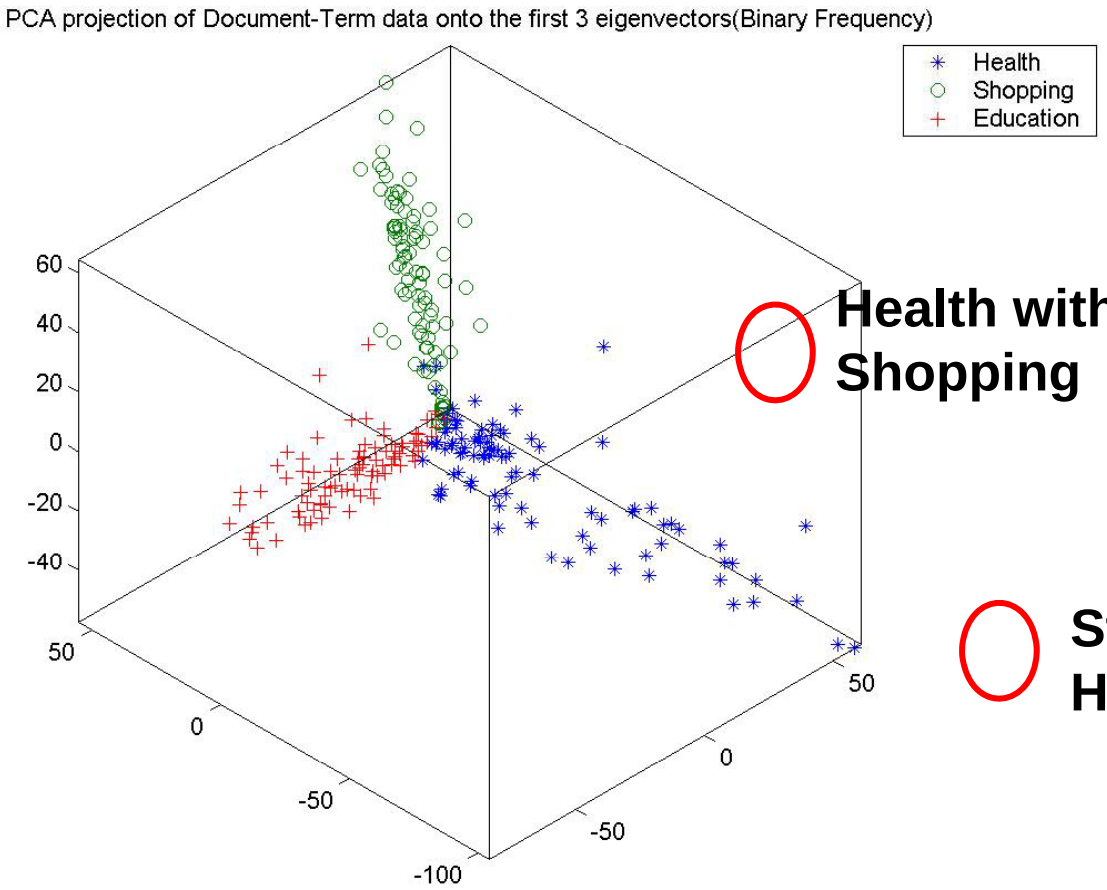
INTELLIGENT INFORMATION SYSTEMS

Web Page Categorization Using PCA

Michael Shepherd, Carolyn Watters, Jack Duffy ….

Web Information Filtering Lab (www.cs.dal.ca/wifl)

Recall and
Precision > 0.80



Health with Shopping

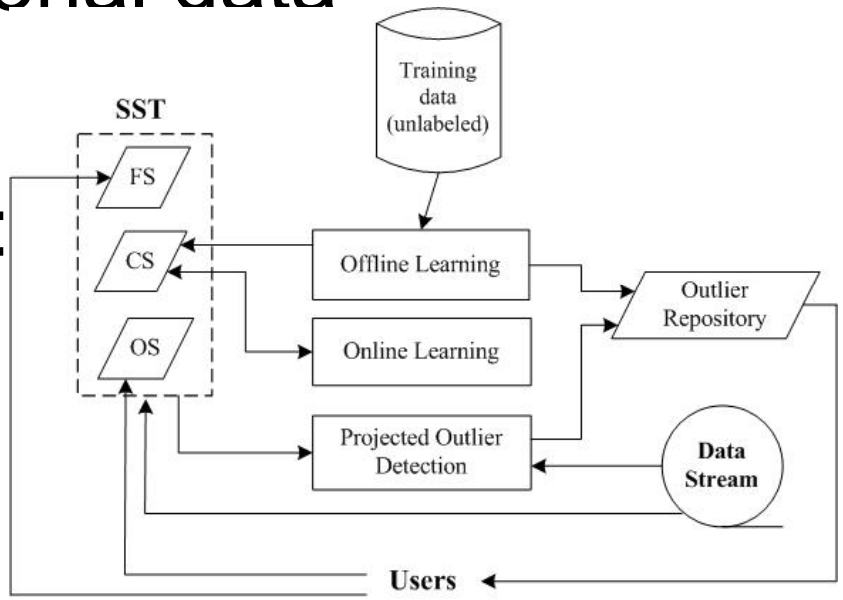
Strong Health

Data Mining on Outlier Detection (OD) for High-Dimensional Data Streams

Q. Gao, H. Wang

- Develop innovative OD solutions based on projected outlier subspace analysis
- OD for high-dimensional data
- OD for stream data
- Research group link:

http://flame.cs.dal.ca/~opami/



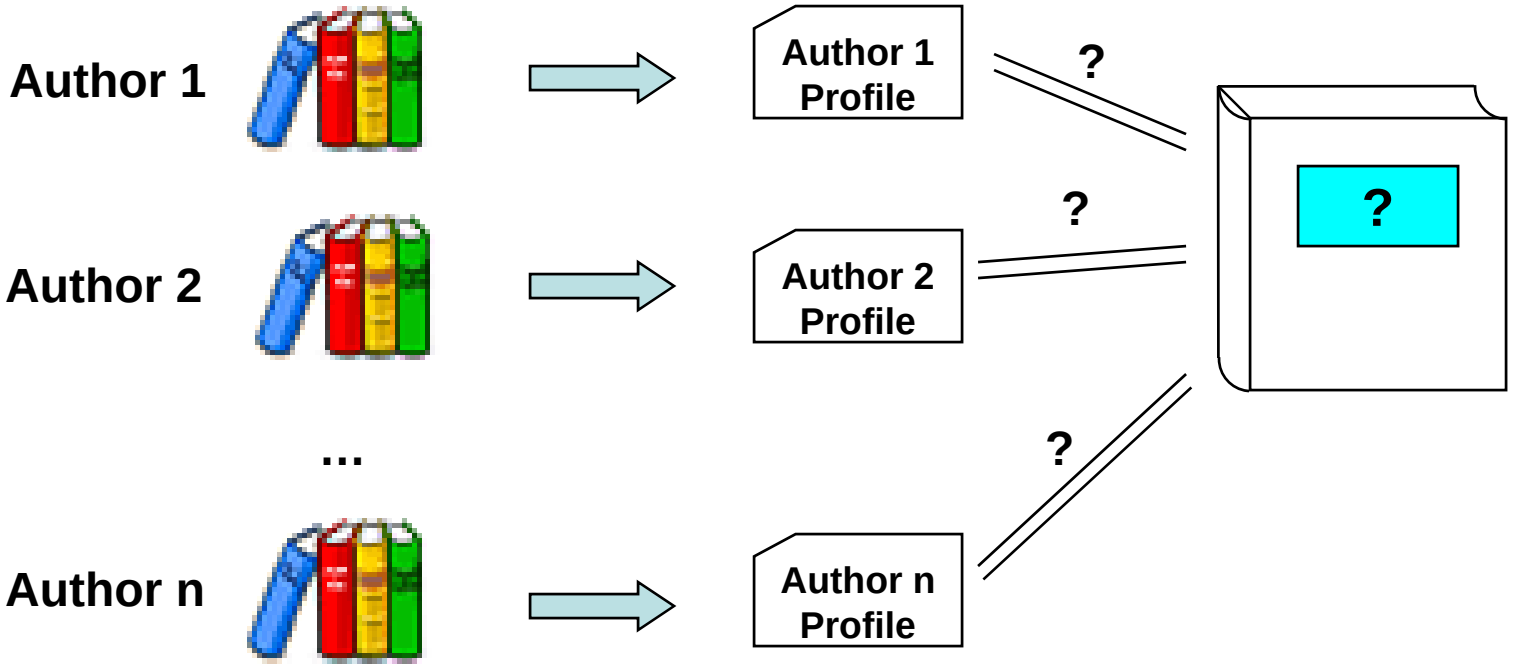
Visual Semantic Computation

Q. Gao, D. Gorodnichy

- Develop perceptual query language and interface toolkit for visual semantic computing
- Content based image/video retrieval
- Motion analysis for surveillance
- Generic image segmentation for supporting semantic interpretation
- Research group link: *http://flame.cs.dal.ca/~ipami/*

Authorship Attribution using Character N-grams

Vlado Keselj

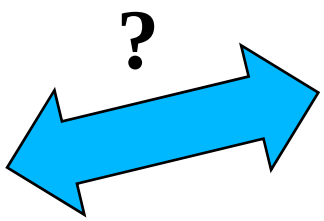


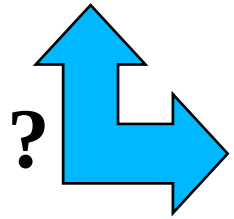
Dickens: A Tale of Two Cities

```
_th 0.016
the 0.014
he_ 0.012
and 0.007
nd_ 0.007
```

Dickens: Christmas Carol

```
_th 0.015
___ 0.013
the 0.013
he_ 0.011
and 0.007
```





Carroll: Alice's adventures in wonderland

```
_th 0.017
___ 0.017
the 0.014
he_ 0.014
ing 0.007
```

NICHE Research Group

(k**N**owledge **I**ntensive **C**omputing for **H**ealthcare **E**nterprises)

Raza Abidi

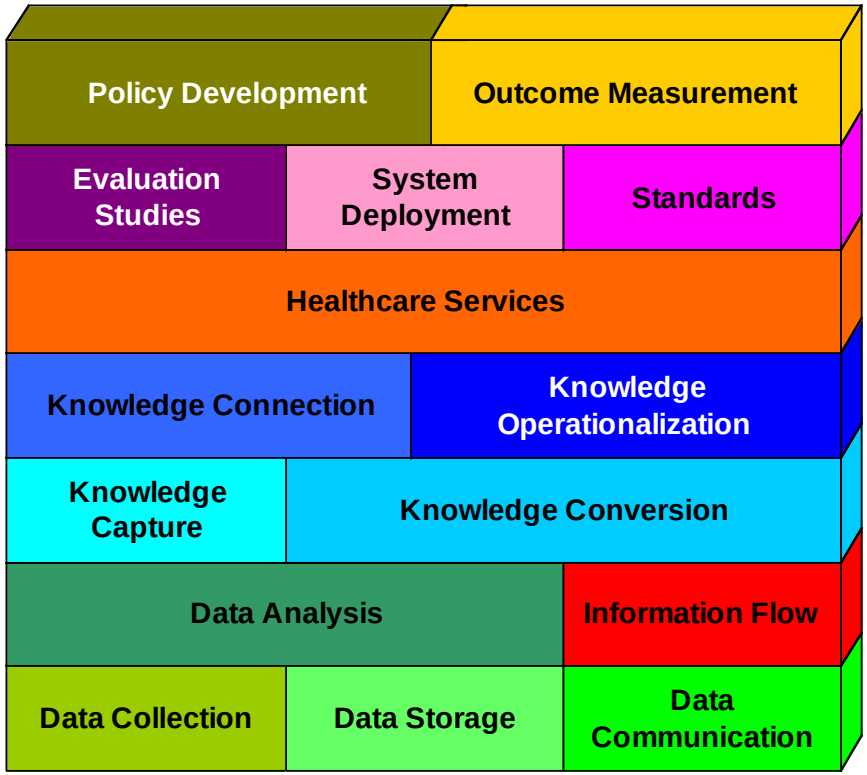
Research Focus is Interdisciplinary

- Computer Science
  - Knowledge management
    - Semantic Web & Ontologies
  - Intelligent personalization
    - Semantic web service composition
    - Dynamic context-sensitive information (content) personalization
- Health Informatics
  - Clinical decision support systems
  - Health knowledge modeling
    - Clinical practice guidelines
    - Clinical pathways
  - Knowledge translation
  - Health data mining

Key Health Informatics Projects

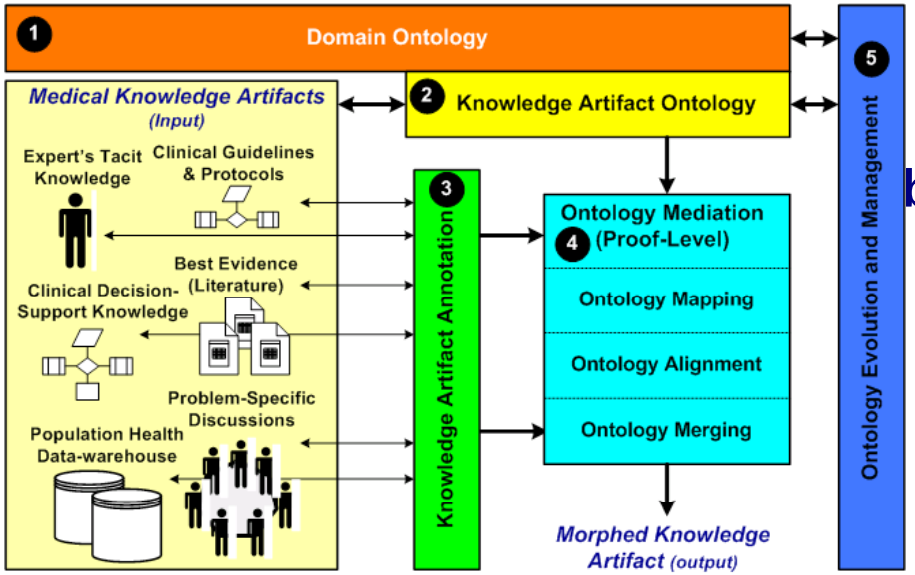
- Knowledge translation in pediatric pain
  - Web 2.0, Social network analysis
- Point-of-care decision-support system for breast-cancer follow-up
  - Semantic web, Reasoning engines
- Care planning for prostate cancer through Care Maps
  - Semantic web, planning systems
- Glaucoma detection from optic discs analysis
  - Data mining, Image analysis
- Knowledge sharing patterns in Emergency Department
  - Knowledge management
- Personalized patient educational program for cardiovascular diseases
  - Adaptive hypermedia, AI

Health Informatics Research Landscape

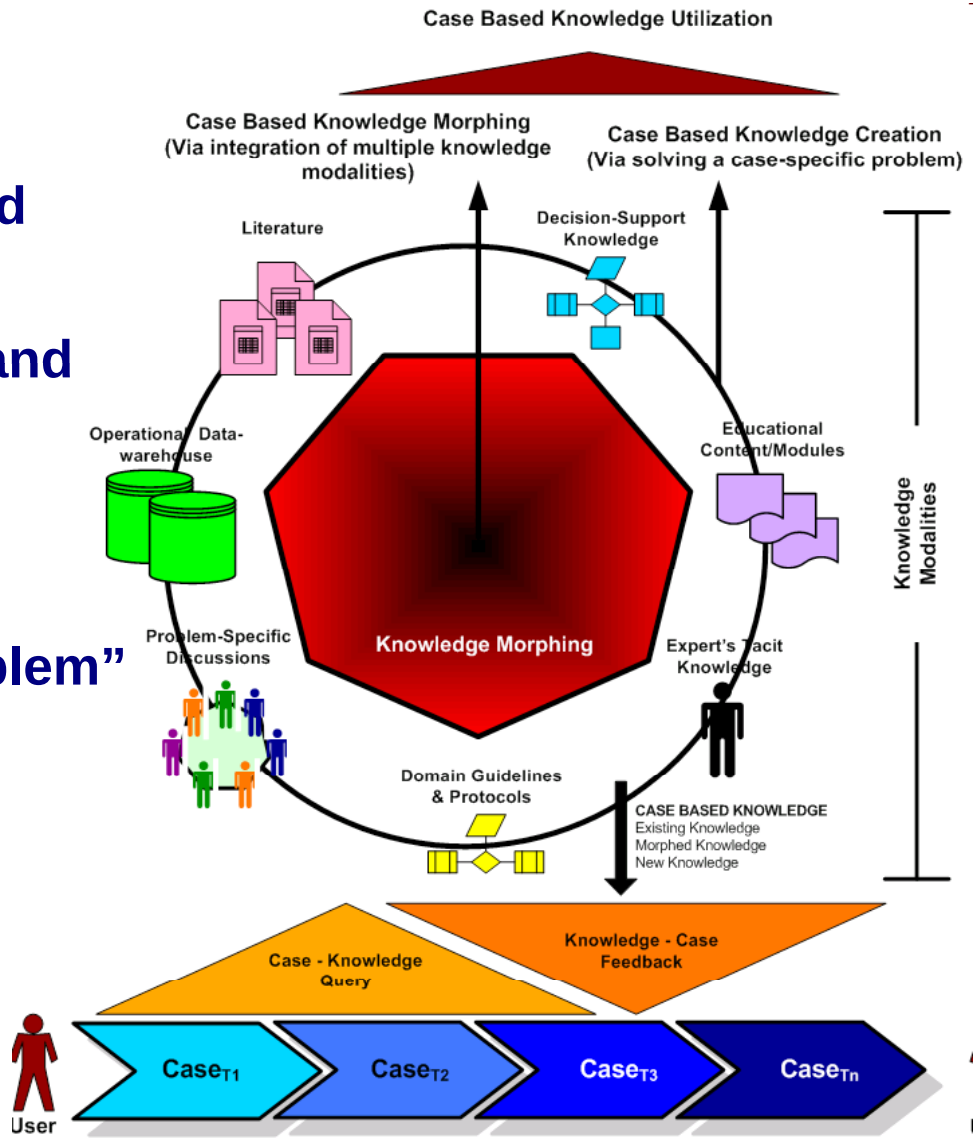


Knowledge Morphing

"The intelligent and autonomous fusion/integration of contextually, conceptually and functionally related knowledge objects **that may exist in different representation modalities and formalisms, in order to establish a**

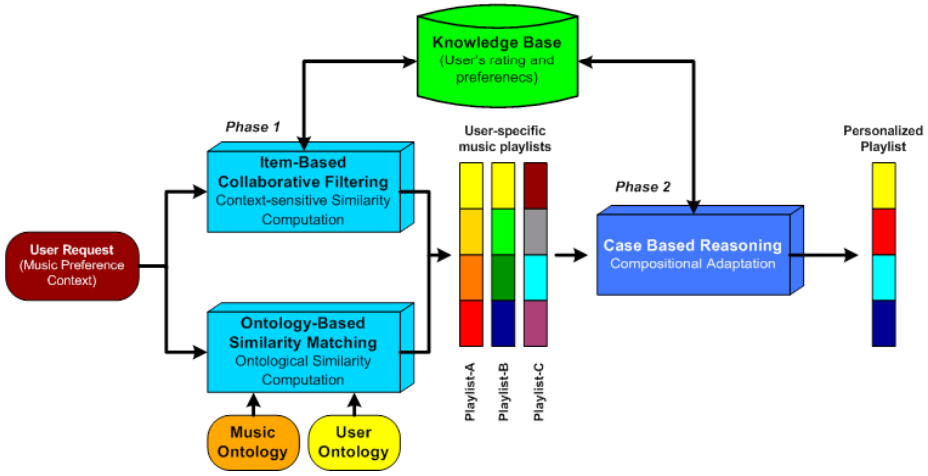


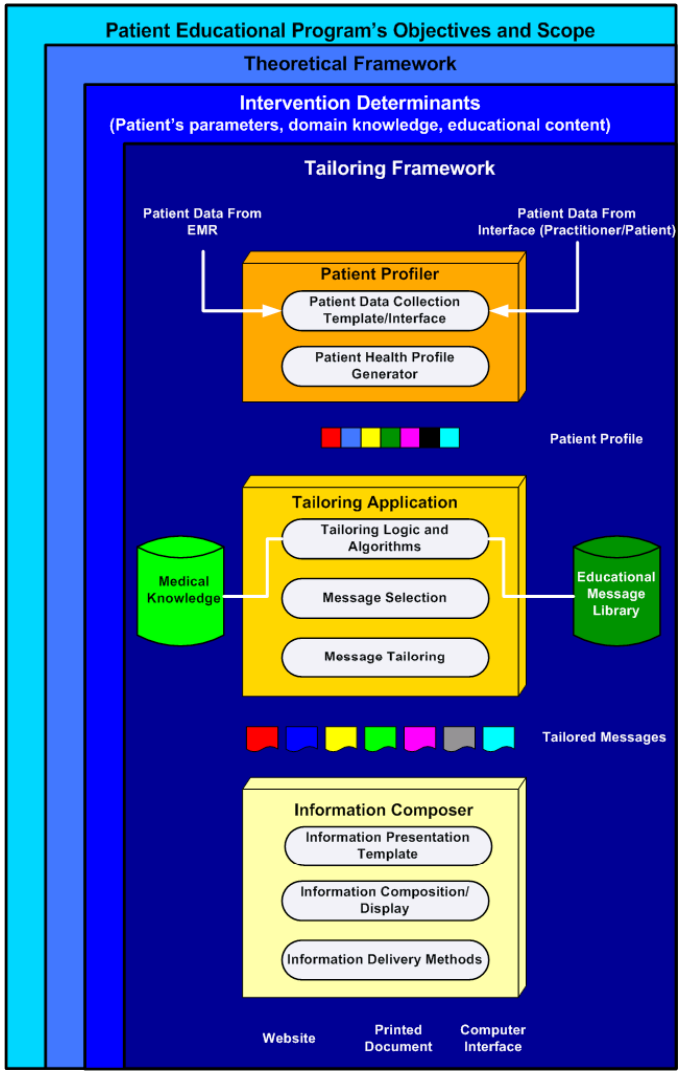
blem"



AdWISE: Adaptive Web Information and Services Environment

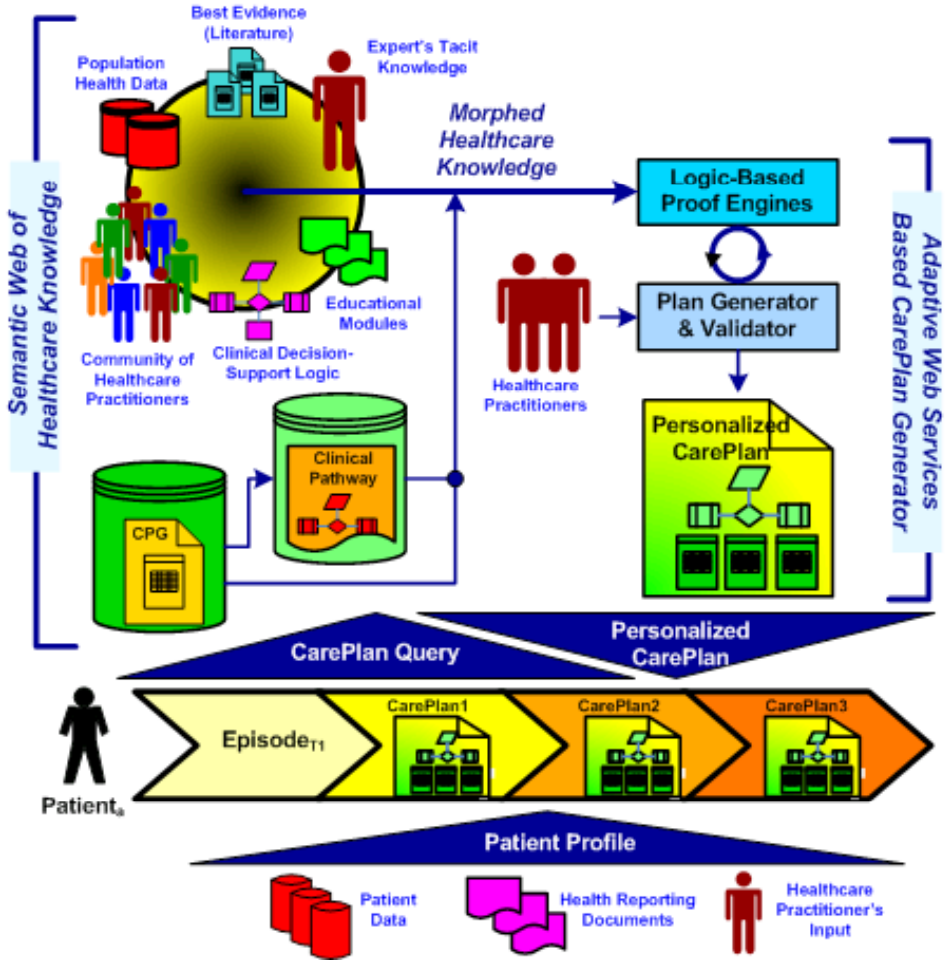
- Intelligent Content Personalization
  - AI Techniques
  - IR Techniques
- Applications
  - Personalized music playlists
  - Personalized news items
  - Personalized cardiovascular risk management recommendations





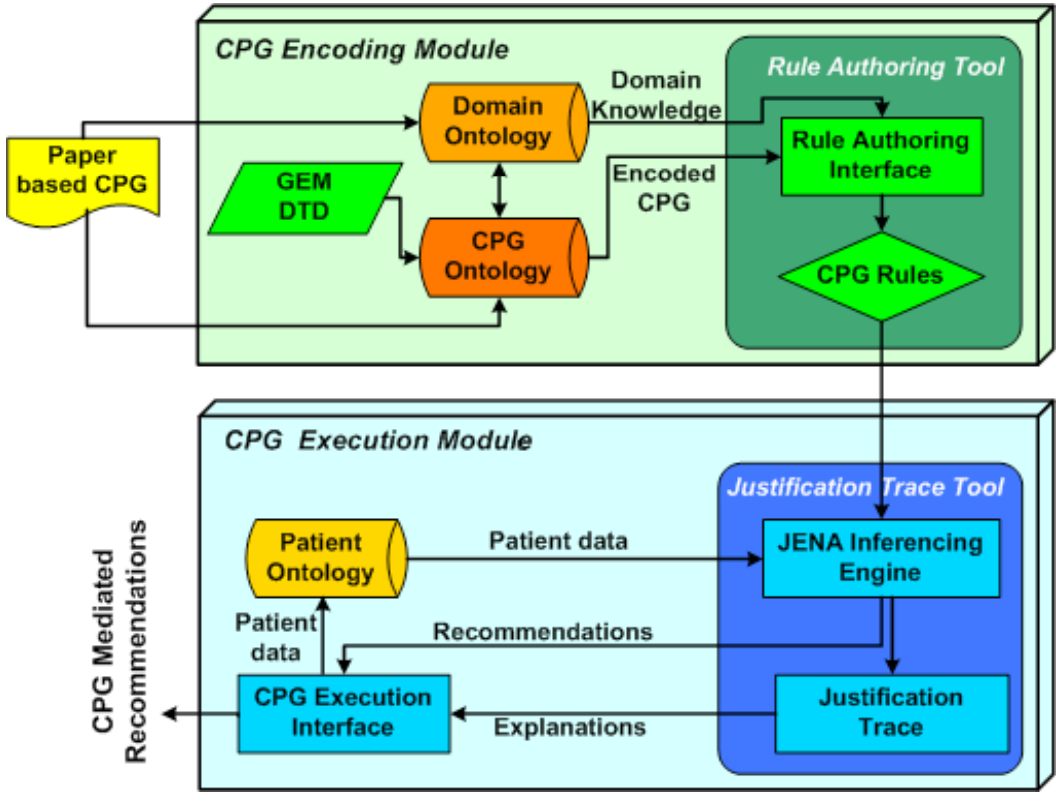
Adaptive Personalized Care Planning via a Semantic Web Framework

- CarePlan is a rich temporal, process-centric, patient-specific clinical pathway that manages the evolving dynamics of a patient to meet the patient's needs, institutional workflows and medical knowledge.



Decision Support Systems

- Semantic Web Approach
  - Knowledge Modeling
    - Ontologies
  - Knowledge Execution
    - Ontology based (logical) decision rules
    - Logic based proof engines
    - Trusted Solutions



Desktop of the future

E. Milios

Personal Document Library

Metadata extraction

Keyterm extraction

Taxonomies: Hierarchical clustering

Key sentence extraction - Summarization

Browsing

Query

Lucene Apache Search Engine Inverted Index

Results

Organization:
clustering
ranking
informative snippets
keywords, key sentences

Automatic Topic Extraction

E. Milios

topic 1	topic 2	topic 3	topic 4	topic 5	topic 6	topic 7	topic 8	topic 9
error	neuron	image	analog	data	control	function	rule	distribution
generalization	neurons	images	circuit	clustering	model	functions	rules	probability
learning	synaptic	object	current	principal	motor	basis	set	gaussian
training	firing	recognition	figure	cluster	forward	linear	step	data
optimal	spike	face	chip	pca	inverse	regression	form	parameters
order	time	objects	voltage	set	dynamics	kernel	fuzzy	model
large	activity	hand	vlsi	algorithm	controller	space	problem	bayesian
average	rate	pixel	circuits	points	feedback	gaussian	relative	mixture
small	synapses	system	digital	approach	system	approximation	extraction	density
examples	potential	view	implementation	clusters	position	rbf	expert	likelihood

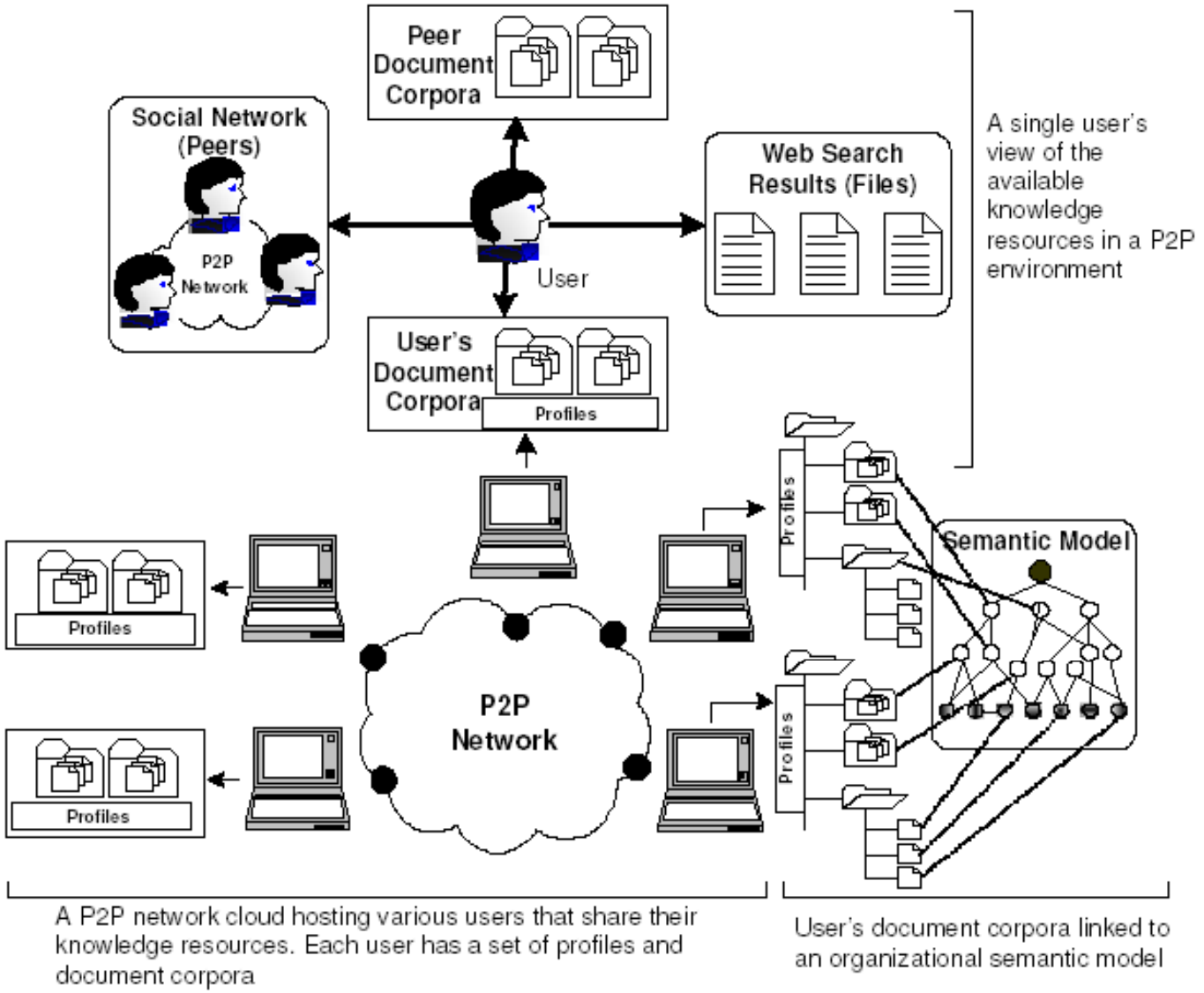
Figure 2. Example word-topics for the NIPS dataset

topic 1	topic 2	topic 3	topic 4	topic 5	topic 6	topic 7	topic 8	topic 9	topic 10
language	game	church	house	air	league	war	apollo	party	system
english	player	god	parliament	aircraft	football	german	earth	government	computer
greek	cards	christian	members	world	team	army	moon	president	game
languages	players	jesus	commons	force	world	soviet	lunar	political	games
word	games	christ	lords	military	club	battle	time	national	apple
russell	play	orthodox	bill	ship	home	germany	mission	minister	atari
century	card	baptism	act	gun	season	world	program	states	commodore
theory	hand	life	power	war	won	forces	module	united	home
words	round	catholic	chopin	ships	game	french	jpg	election	software
modern	played	roman	speaker	navy	major	union	crew	state	video

Figure 3. Example word-topics for the Wikipedia dataset

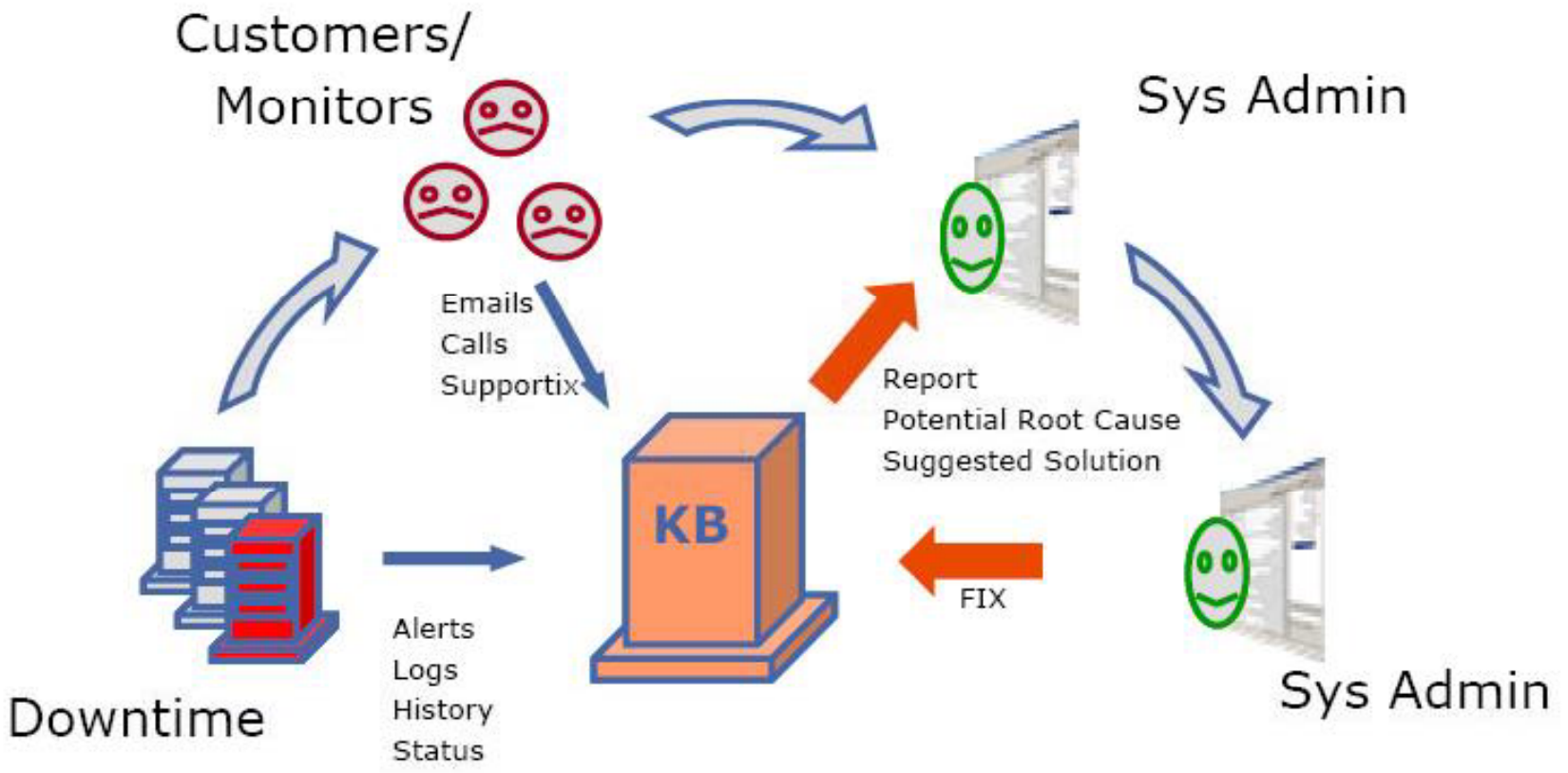
Peer-to-Peer Document Management

V. Keselj, E. Milios, S. Abidi



Experience Management

E. Milios, N. Zincir-Heywood

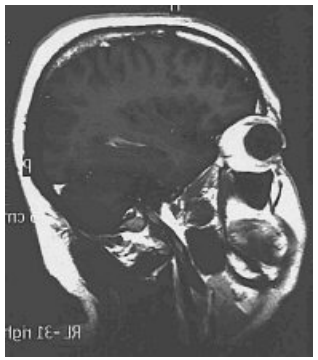


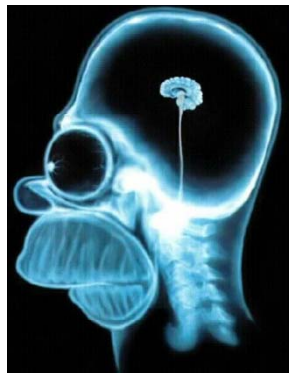
AI FUNDAMENTALS

DALHOUSIE

COMPUTATIONAL NEUROSCIENCE GROUP

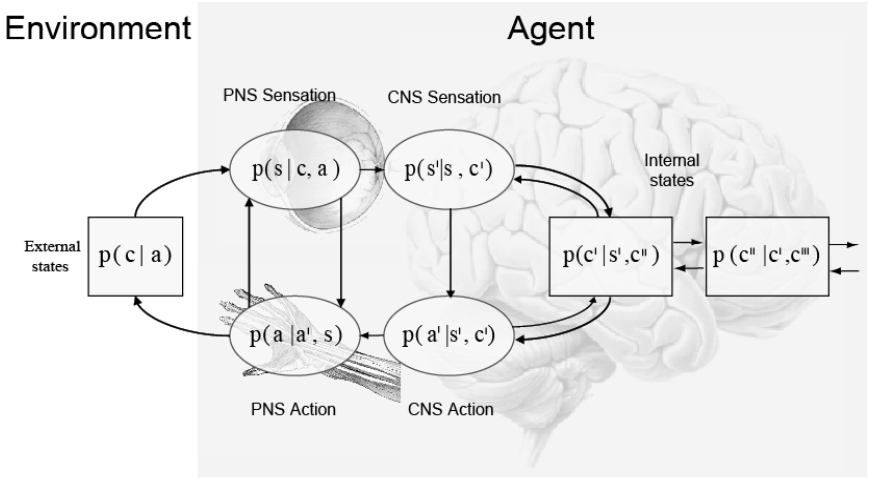
Studying Minds



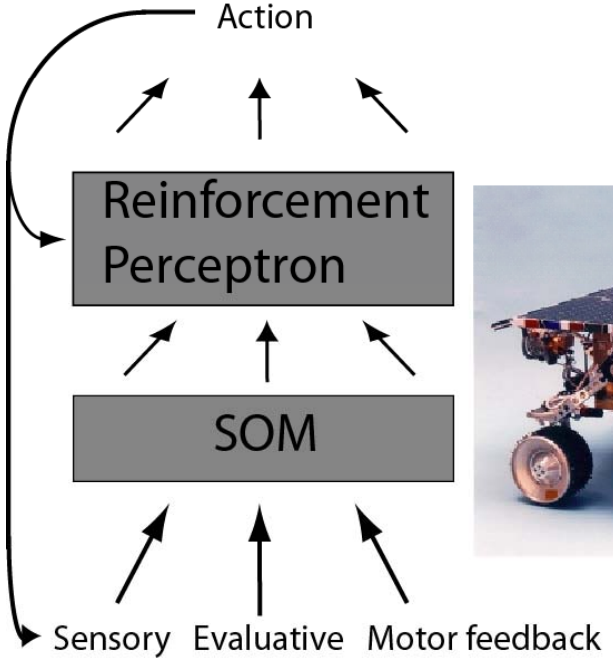


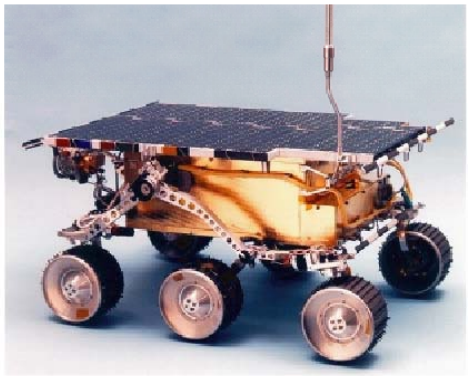
Dr. Thomas Trappenberg

Computational Neuroscience

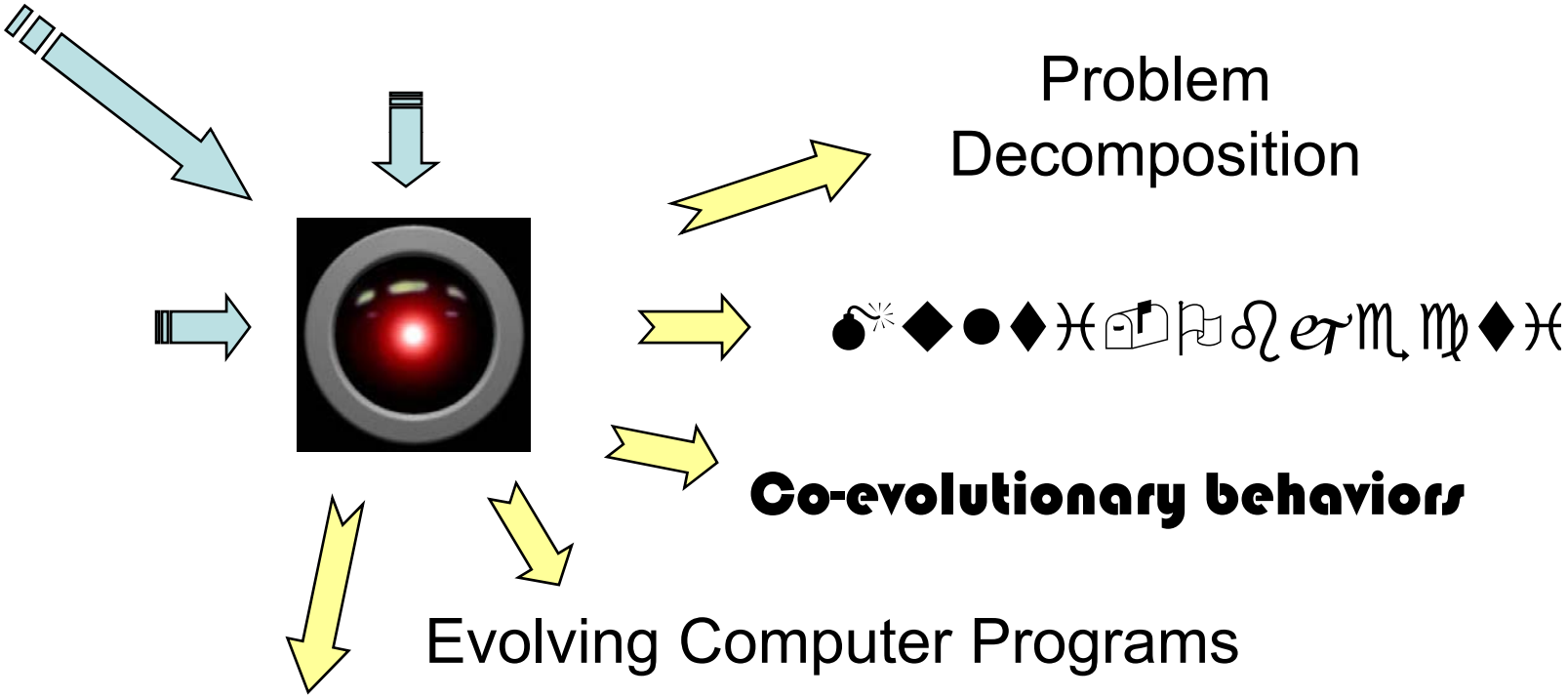


Machine Learning





Genetic Programming

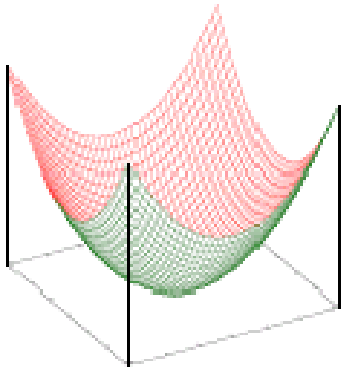


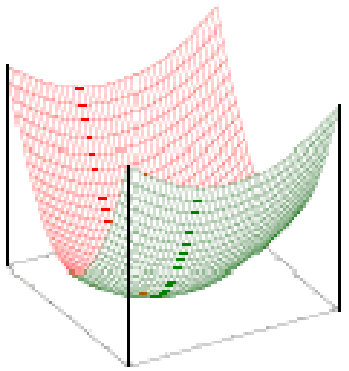
Malcolm
Heywood

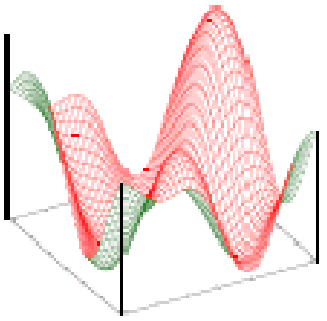
Evolutionary Computation


Dirk Arnold

- evolutionary algorithms are optimisation strategies “gleaned from nature”
- areas of application range from engineering design and control to financial forecasting and art
- research of Dalhousie’s Evolutionary Computation group focuses on understanding, improving, and developing adaptive strategies
- contact: Dr. Dirk Arnold (http://www.cs.dal.ca/~dirk)





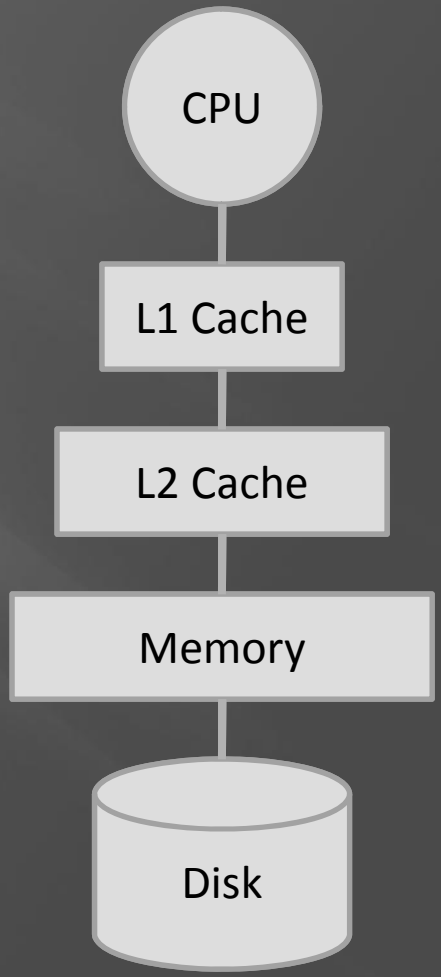


THEORY

Algorithms and Data Structures for Memory Hierarchies

Norbert Zeh
Canada Research Chair in Algorithms for Memory Hierarchies

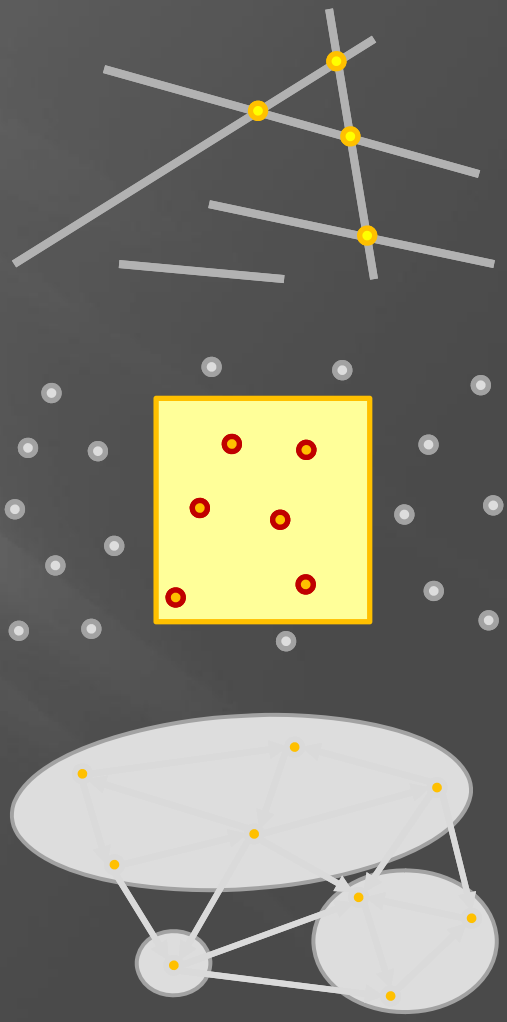
- Disk I/O bottleneck when processing massive datasets
- Low cache efficiency in traditional algorithms
- Need algorithms with high access locality to
  - Take advantage of caches
  - Take advantage of disk read-ahead
- **Techniques fundamentally different from traditional algorithms!**



Algorithms and Data Structures for Memory Hierarchies

Norbert Zeh
Canada Research Chair in Algorithms for Memory Hierarchies

- **Geometric problems**
  - Databases (range queries, etc)
  - GIS (map overlay, window queries, etc)
  - …
- **Graph problems**
  - Web modeling
  - GIS (route planning, logistics)
  - Bioinformatics (protein clustering, etc)
  - …



Fault-tolerant networks

Zizo Farrag

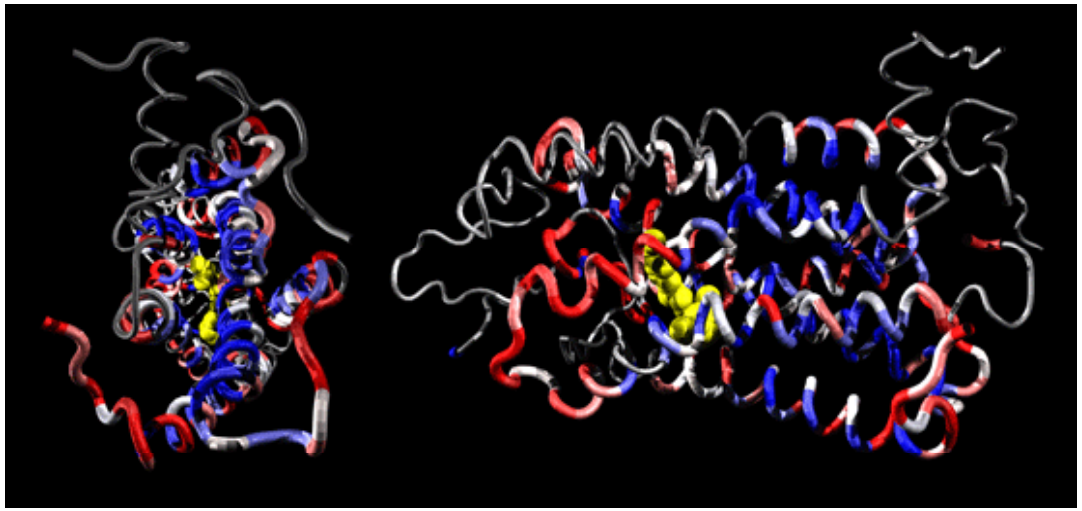
- Design and Reconfiguration of fault-tolerant networks.
- Objectives: construct a network that
  - Can continue to operate in the presence of certain faults,
  - Is optimal or near-optimal in cost,
- Cost will depend on the parameters to be optimized
- Efficiency of reconfiguration measured by the time needed to identify a healthy sub-graph of the network (that excludes the defective components).

BIOINFORMATICS

Bio-informatics

Dr. Christian Blouin

Optimizing confidence intervals in phylogeny
Parallel Computing in protein phylogeny
Sequence alignment curation using Artificial Intelligence
A C++ bioinformatics library
Interactive Phylogeny
Protein Biophysics and the substitution process
Structural Evolution
Folding of protein loops



HUMAN CENTRIC COMPUTING

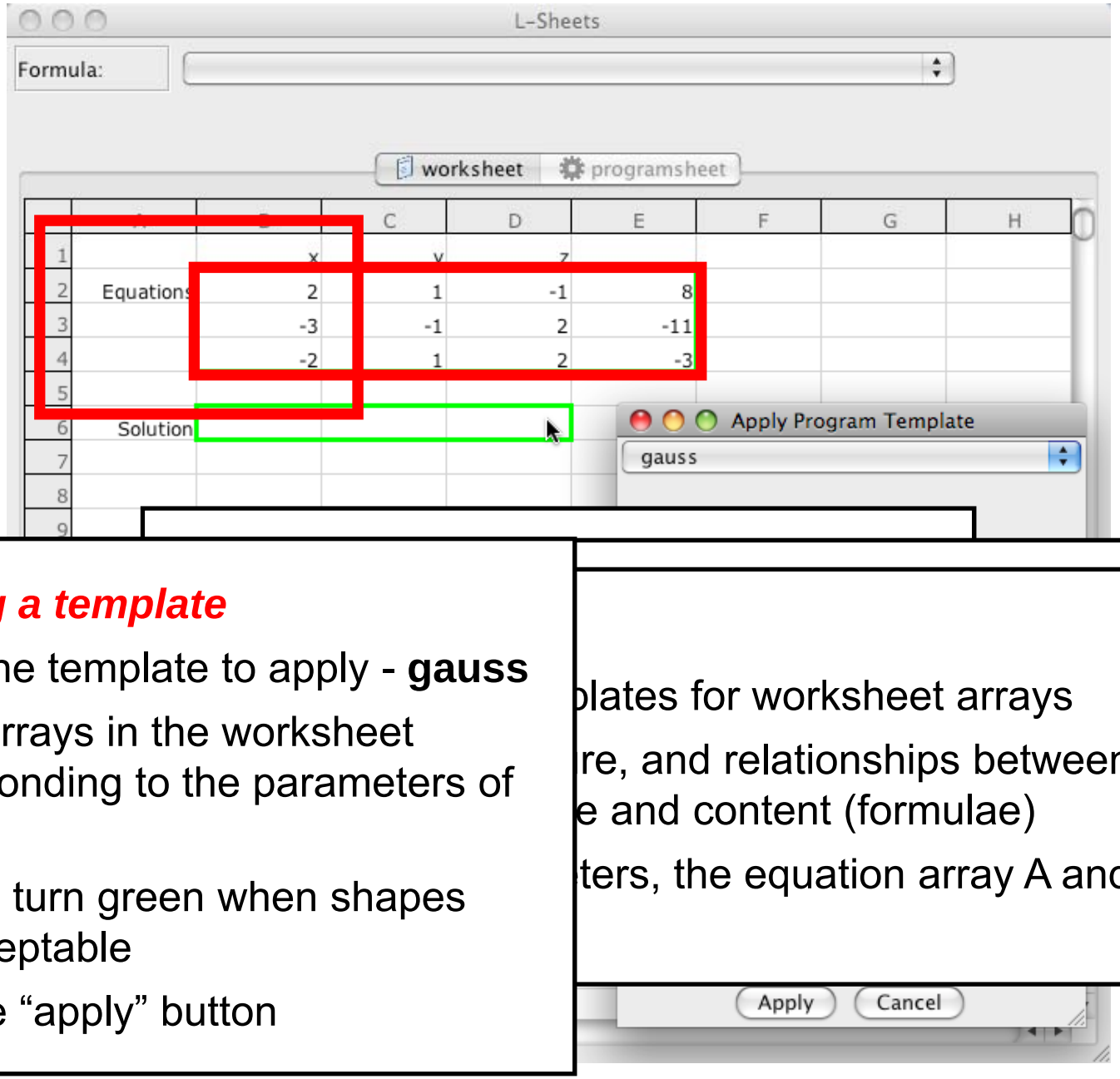
isual Languages and Computation
Phil Cox

Visualisation in software development

- Visual Languages (VL)
  - graphical notations that directly express the multidimensional structure of algorithms and data.
- Visualisation of execution
- End-user and domain-specific programming

Some current projects

- Design of structured objects
- Programming by demonstration
- VLs for industrial software development
- Spreadsheet programming and templating
  - Example: Gaussian elimination for solving sets of linear equations (not a typical usual end-user application!)...



Applying a template

- select the template to apply - **gauss**
- select arrays in the worksheet corresponding to the parameters of **gauss**
- outlines turn green when shapes are acceptable
- click the “apply” button

lates for worksheet arrays

re, and relationships between
e and content (formulae)

ters, the equation array A and

The Dalhousie Graphics and Visualization Lab

GVLAB

The Graphics and Visualization Lab

- The focus is on both:
  - the development of new graphical techniques, and
  - the application of those techniques, often in cross-disciplinary areas

- Our lab incorporates expertise in areas such as:
  - image processing
  - 3D computer graphics
  - physically-based rendering
  - visualization
  - and, traditional art

Graduate Courses & Faculty Members

- Visualization (6406)
  - focuses on graphical techniques for data visualization that assist in the extraction of meaning from datasets
- Advanced Computer Animation (6608)
  - covers topics in computer animation, including forward and inverse kinematics, motion capture, and physically based modelling
- Digital Image Processing (6602)
  - covers topics in digital picture processing such as visual perception, digitization, compression and enhancement

DISTRIBUTED AND SOFTWARE SYSTEMS

Network Information Management and Security

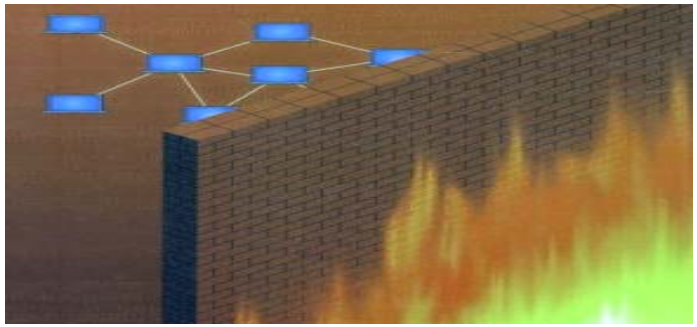
QuickTime™ and a TIFF (Uncompressed) decompressor are needed to see this picture.

Nur Zincir-Heywood

www.cs.dal.ca/~zincir

Network Analysis

Attack Modelling



Fault Management

Computer Security

Dawn Jutla
http://husky1.stmarys.ca/~djutla/
Dawn.Jutla@smu.ca

Collaborative User Services for Private Data Management (CUSP)

The CUSP (Collaborative User Services for Private Data Management) project intends to deliver sophisticated user privacy services over the Semantic Web. This Canadian project is a collaborative effort between faculty in the Sobey School of Business, Saint Mary's University and the Faculty of Computer Science, Dalhousie University.

Currently many knowledge-intensive privacy-related tasks are manual. Using Semantic Web technologies (OWL, RDF, XML, UDDI, SOAP, and WSDL), knowledge-base and database methodologies, and building on the P3P platform (XML vocabulary for privacy), the CUSP project automates human decision making processes with respect to online privacy.

Further information at http://users.cs.dal.ca/~bodorik/Cusp.htm

Peter Bodorik
www.cs.dal.ca/~bodorik

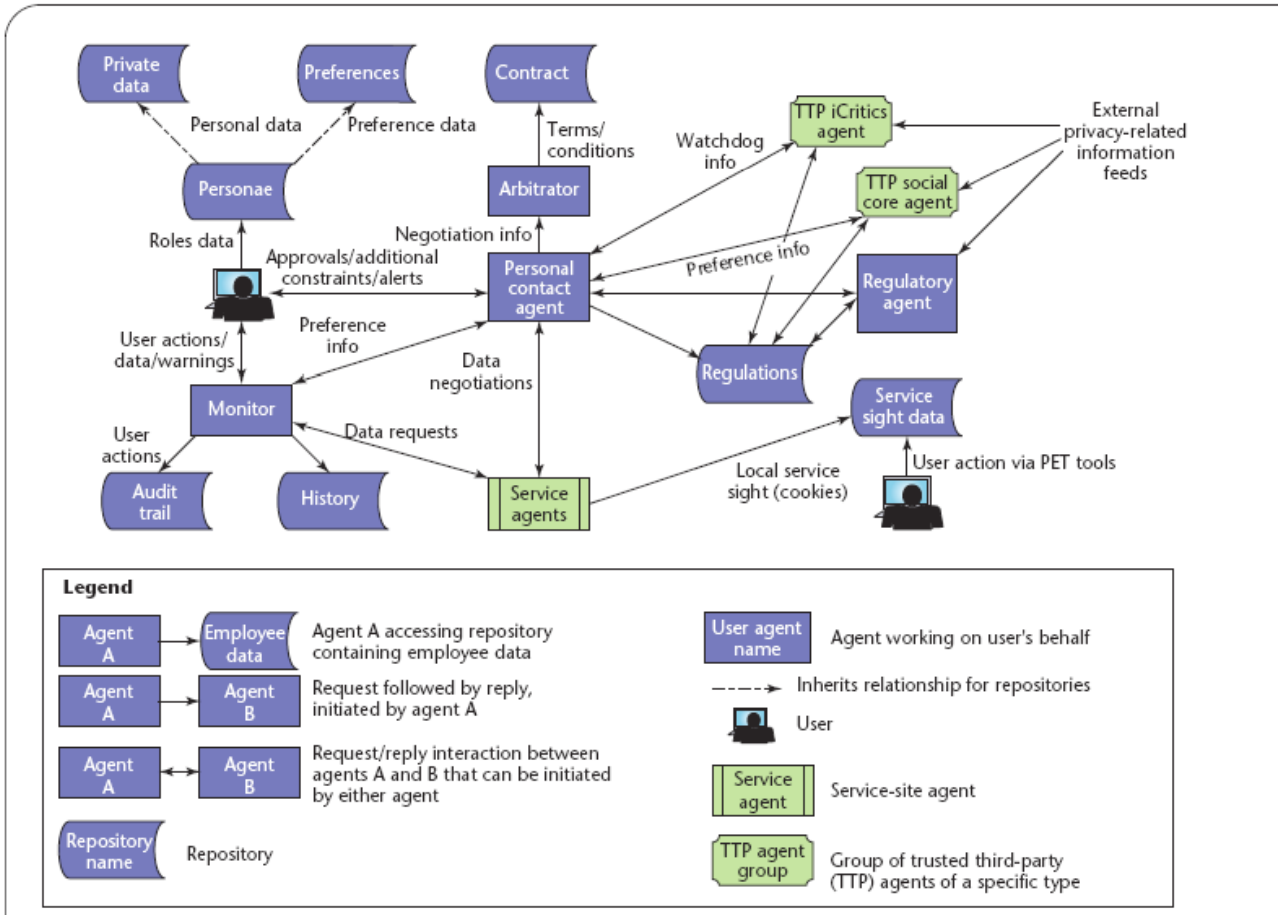


Figure 3. A multiple agent system and knowledge base architecture for user-controlled e-privacy. The monitor agent oversees the user's interaction with Web forms and other interaction mechanisms at Web service sites. The context agent manages dynamic changes of the user privacy context as the user interacts with sites on the Web, informs the user about the current privacy-related context for decision support within a Web transaction, and triggers revision of user privacy preferences either due to other agents in the architecture or actions by the user. The arbitrator agent allows users to negotiate on their PII's usage purposes, handling recipients and PII retention periods with an organization on the Web. The regulatory agent invokes privacy Web services and utilizes external service feeds and trusted third-party (TTP) agents to obtain knowledge on privacy regulations, guidelines, and service sites in multiple jurisdictions. The iCritics and social contract core agents are external agents proposed in other research that add further stakeholder value to our sociotechnical architecture.

From: Jutla D. and Bodorik P., "Socio-technical Architecture for User-Controlled Online Privacy," IEEE Security and Privacy, March/April 2005, pp. 24-34.

Privacy Policy Compliance in Web Services Architecture

Peter Bodorik
www.cs.dal.ca/~bodorik

Dawn Jutla
http://husky1.stmarys.ca/~djutla/
Dawn.Jutla@smu.ca

This project provides technologies to support compliance to privacy regulations in a Web Services Architecture. Automated agents examine messages exchanged when invoking web-services. The agents utilize a Privacy Knowledge Base to ensure that Private Information that is exchanged satisfies applicable privacy policies.

For further information contact Dr. Bodorik or Dr. Jutla.

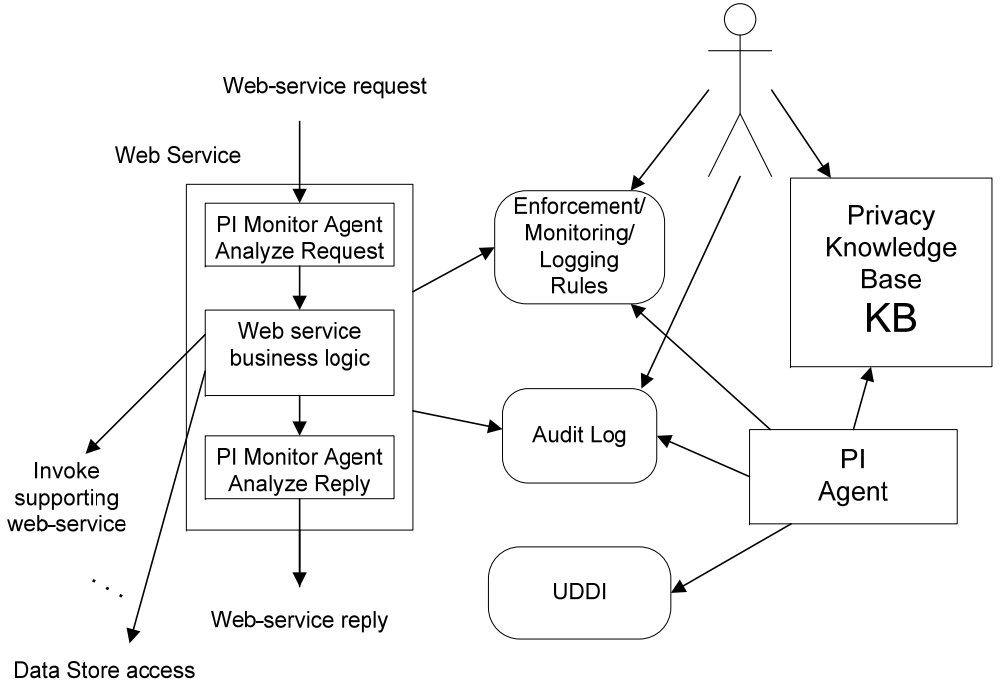


Figure 2 – Component View of Privacy Architecture

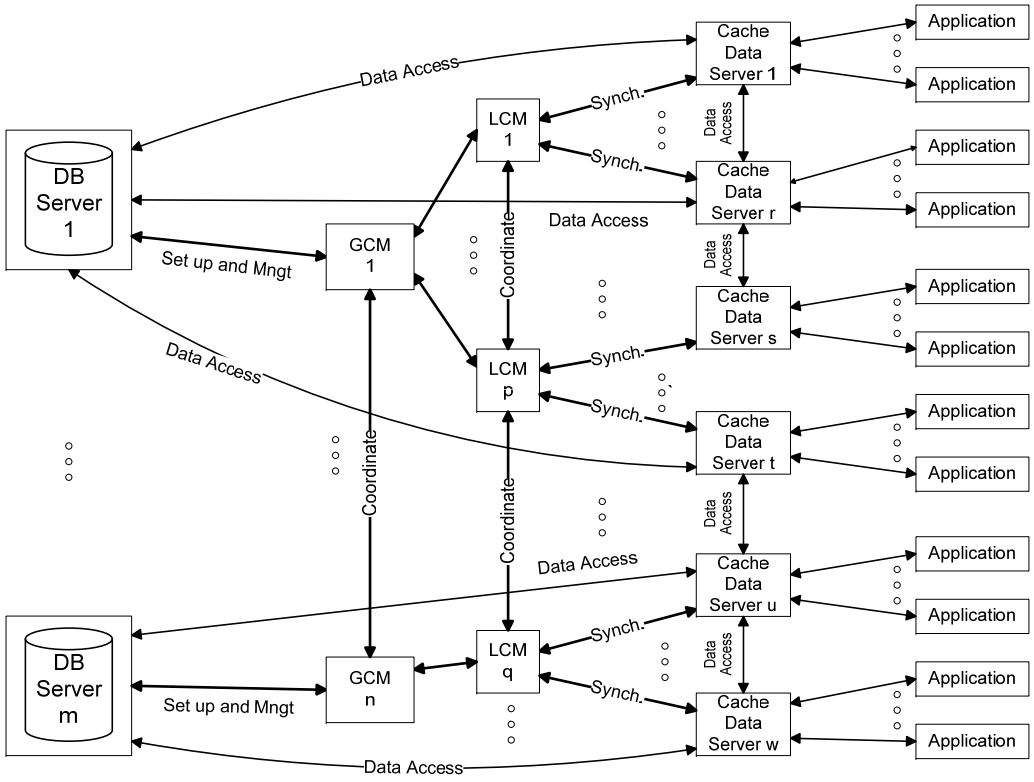
Highly Scalable High Performance Caching Architecture

Achieved by Interoperable Cache Managers and Data Servers

DB servers are becoming bottlenecks in enterprise caching architectures.
A highly scalable and high performance caching architecture is achieved by

- Offloading the caching responsibilities of a DB server to Global Cache Managers (GCMs)
- Local Cache Managers (LCMs) coordination with Cache Data Servers in caching protocols
- Interoperable caching protocols that support applications with different caching requirements

For further information contact Dr. Bodorik at www.cs.dal.ca/~bodorik



Peter Bodorik
www.cs.dal.ca/~bodorik

CSCI 6401 Distributed Databases

Instructor: Peter Bodorik

www.cs.dal.ca/~bodorik; email: bodorik@cs.dal.ca

Mondays, Wednesdays 11:05-12:25, Computer Science LAB-3

Objectives

The main objective of this course is to examine the issues arising in the design and implementation of distributed databases. Another objective is to examine current developments in the use of DBs and information systems in support of Enterprise Information Systems.

Course Organization

A portion of the course is devoted to the subject matter appearing in the textbook. Lectures are used to outline the problems and their solutions. You are expected to study the subject matter and pass assignments and tests.

You will investigate an assigned topic dealing with usage of DBs or systems accessing DBs, give a presentation on it and submit a report.

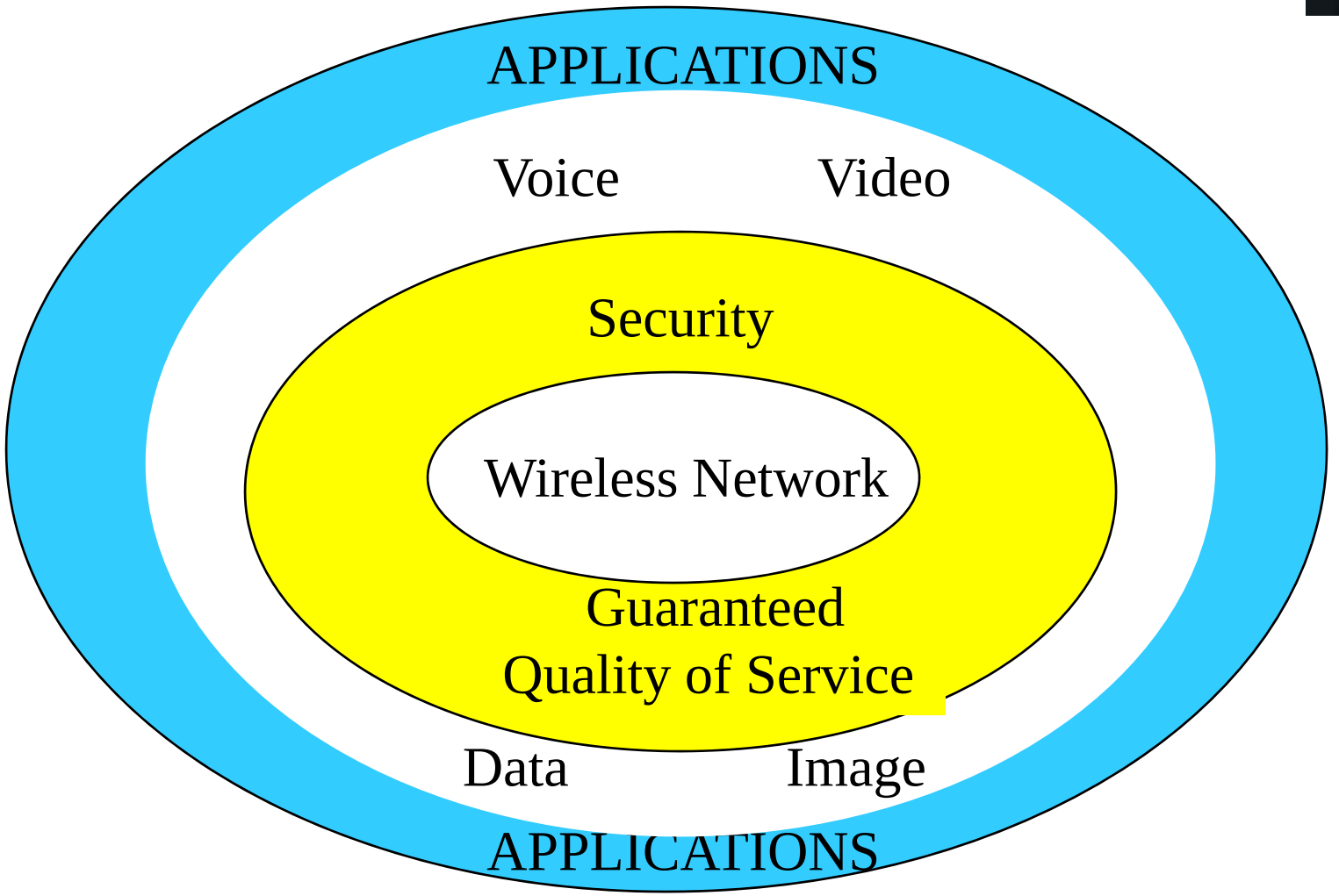
Dr. Srinivas Sampalli

WISE (Wireless Security) Group

- Investigate protocol vulnerabilities in wireless networks – WiFi, WiMAX and Ad Hoc Wireless
- Build a manual for best practice for wireless security.
- Design intrusion detection and prevention mechanisms for enhancing security.
- Implement prototypes and build a test bed for validating these detection and prevention mechanisms.
- Integrate security and quality of service in heterogeneous and hybrid networks.

Dr. Srinivas Sampalli

WISE (Wireless Security) Group



Graduate School Information

Choosing advisor & thesis topic

From: How to succeed in graduate school (by Marie deJardins, SRI International)

- a good thesis topic is interesting:
  - to you,
  - to your advisor, and
  - to the research community
- Professors may have
  - Well defined long-term research programs and expect their students to contribute directly
  - Much looser, but still related ongoing projects.
  - Tendency to take on anyone with an interesting idea (beware of advisor lack of commitment)

Scope of reading & topic

- Awareness & Reading
  - Be selective: you'll never be able to read everything that might be relevant
  - Become and stay aware of directly related research
- Topic options
  - Narrow, well defined topic.
    - Plus: finish fast
    - Minus: it may not be as exciting
  - Exotic topic
    - Plus: potentially exciting
    - Minus: difficulty convincing people it's worthwhile.

Good topic choices

- Solve a real problem, not a toy problem
- Choose:
  - a central problem that's solvable and acceptable
  - with extensions and additions that:
    - are successively riskier and that
    - will make the thesis more exciting.

Programme Form

- FGS is responsible for the program of all graduate students at Dalhousie.
- Coursework for a graduate student is approved by a faculty advisor
- Programme Form
  - Shows the list of approved courses for a student
  - A contract between the student and Dalhousie
  - List can be changed later (with approval)

DALHOUSIE University

Faculty of Graduate Studies

GRADUATE STUDENT PROGRAMME FORM

TO BE COMPLETED BY DEPARTMENT ONE MONTH AFTER REGISTRATION.

STUDENT NAME:	
ADDRESS:	
TELEPHONE:	E-MAIL:
STUDENT NUMBER:	START DATE:

YEARS OF PROGRAMME FEE REQUIREMENT: ☐ 1 YEAR ☐ 2 YEARS ☐ 3 YEARS *OR* ☐ FEE PER COURSE PROGRAMME

DEGREE PROGRAMME:	DEPARTMENT/SCHOOL:
REGISTRATION STATUS: ☐ FULL-TIME ☐ PART-TIME • PHD STUDENTS MUST REGISTER FULL-TIME	AREA OF SPECIALISATION:

TOTAL OF FULL-CREDIT CLASSES (WITH CREDIT VALUE) REQUIRED FOR DEGREE (EXCLUDING THESIS):

INDICATE ALL REQUIREMENTS; INCLUDE BOTH COURSE NUMBER AND CREDIT VALUE

REQUIRED COURSES FOR PROGRAMME

SUBJECT CODE	COURSE NUMBER	# CREDIT HOURS	SUBJECT CODE	COURSE NUMBER	# CREDIT HOURS

CREDITS TO BE ADDED LATER:

CREDITS AUDITED:

ANCILLARY CREDITS (E.G. UNDERGRADUATE COURSES NOT REQUIRED TO COMPLETE THE PROGRAMME):

ADDITIONAL REQUIREMENTS FOR DEGREE (GIVE CLASS NUMBER IF APPLICABLE):

☐ THESIS:	☐ PROJECT:
☐ PRACTICUM:	☐ LANGUAGE REQUIREMENT(S):

OTHER REQUIREMENTS (PLEASE SPECIFY):

NAME(S) OF SUPERVISOR(S):

NAMES OF COMMITTEE MEMBERS (IF KNOWN):

APPROVALS:

STUDENT:		DATE:	
SUPERVISOR:		DATE:	
GRADUATE COORDINATOR:		DATE:	
FACULTY OF GRADUATE STUDIES:		DATE:	

Revised September 2003

For more information

- WWW: http://www.cs.dal.ca/graduate/
- Email: grad@cs.dal.ca
- Resources about graduate school:
  - thesis writing
  - how to do research
  - how to give presentations
  - job interview preparation

http://users.cs.dal.ca/~eem/gradResources/gradResources.htm