BCNet Advanced Networks Conference 2004 keynote

Framing the requirements for research environments of the future

Brian R Gaines & Daniela Damian

There are now a number of Scientific Workplaces of the Future projects (SWOFs) providing collaboratories based on broadband networks, grid computing, multimedia workstations/ workrooms, tele-operation of laboratory facilities, and a range of computational tools including databases, digital libraries, numerical and logical data processing, simulation and visualisation. There is also growing use of all these facilities as declining costs have brought them within the budgets, initially of senior researchers, and now increasingly of nearly everyone. Much of the technology in use in these projects was not developed to support research processes, but the general-purpose nature of all information technology has meant that it can be repurposed to do so. If the development and application of the access grid, grid technologies and broadband computing in general, to the systematic acceleration of knowledge generation and utilisation is not to go beyond *ad hoc* usage we need a requirements framework that subsumes and structures the diversity of activities noted above. We have been developing a reference architecture for identifying weaknesses in the existing infrastructure, technologies that might be developed to improve it, and guidelines for its effective use

There is now a wealth of material available on the social and cognitive processes underlying research. Historians, anthropologists, sociologists and psychologists have studied how scientific, engineering and general scholarly communities operate, and there are significant theories and empirical data available. The impact of new instruments and new communication technologies have also been studied. One can use this material to develop metrics relevant to the impact of new technologies, such as the speed of information diffusion and the total corpus of relevant information and its accessibility, and plot the changes over time. There are also microeconomic models of the social processes of scientific communication taking into account the non-monetary nature of the reward system involved. From these one can begin to develop a requirements framework that should be useful in designing better systems and in avoiding the pitfalls that arise from the gap between the ideals of collaborative ‘pure’ science and the realities of the highly competitive real world of research and scholarship.

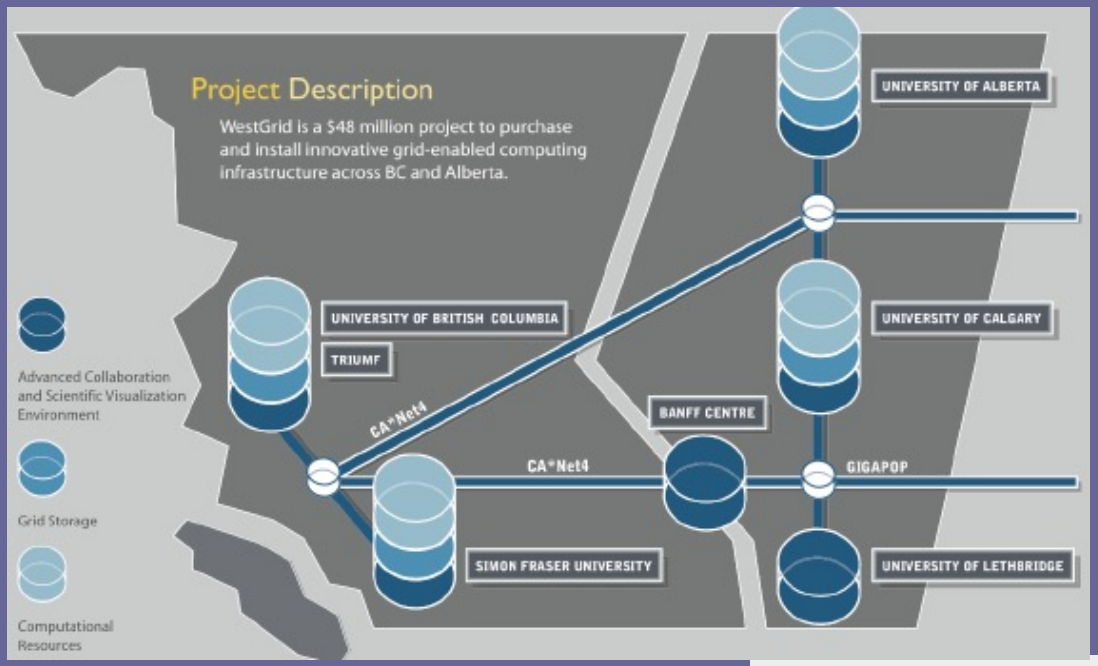
This presentation presents some of the models and data from other disciplines that we are using in order to begin to develop a reference model in which to frame the requirements for research environments of the future, and outlines some of the key issues that emerge in relation to the design and application of technologies aimed at systematically accelerating research.

Dr Gaines is Professor Emeritus, University of Calgary, and was previously Killam Research Chair and Director of the Knowledge Science Institute. He has authored over 500 papers in areas of information technology and human sciences, and developed a number of networked groupware systems supporting stock exchanges, hospitals and other communities. In recent years he has developed web services supporting knowledge acquisition and management being used by thousands of users world-wide. He was Canadian coordinator of the GNOSIS intelligent manufacturing systems consortium with participants world-wide, responsible for its support through the Internet, and studied the impact of the technology on the research. His current interests include the historical development of science and technology, particularly the role of social and technological processes. He has collaborated with Dr Damian in her research on industrial applications of collaborative software engineering in geographically distributed teams, in particular distributed requirements negotiations, and is currently collaborating with her on wider applications of the Access Grid technology for which she has received CFI funding. Dr. Damian is leading the Software Engineering Global interAction Laboratory (SEGAL) at the University of Victoria, where advanced state-of-the-art collaborative technologies are studied for support of global interaction in software engineering.

Framing the Requirements for Research Environments of the Future

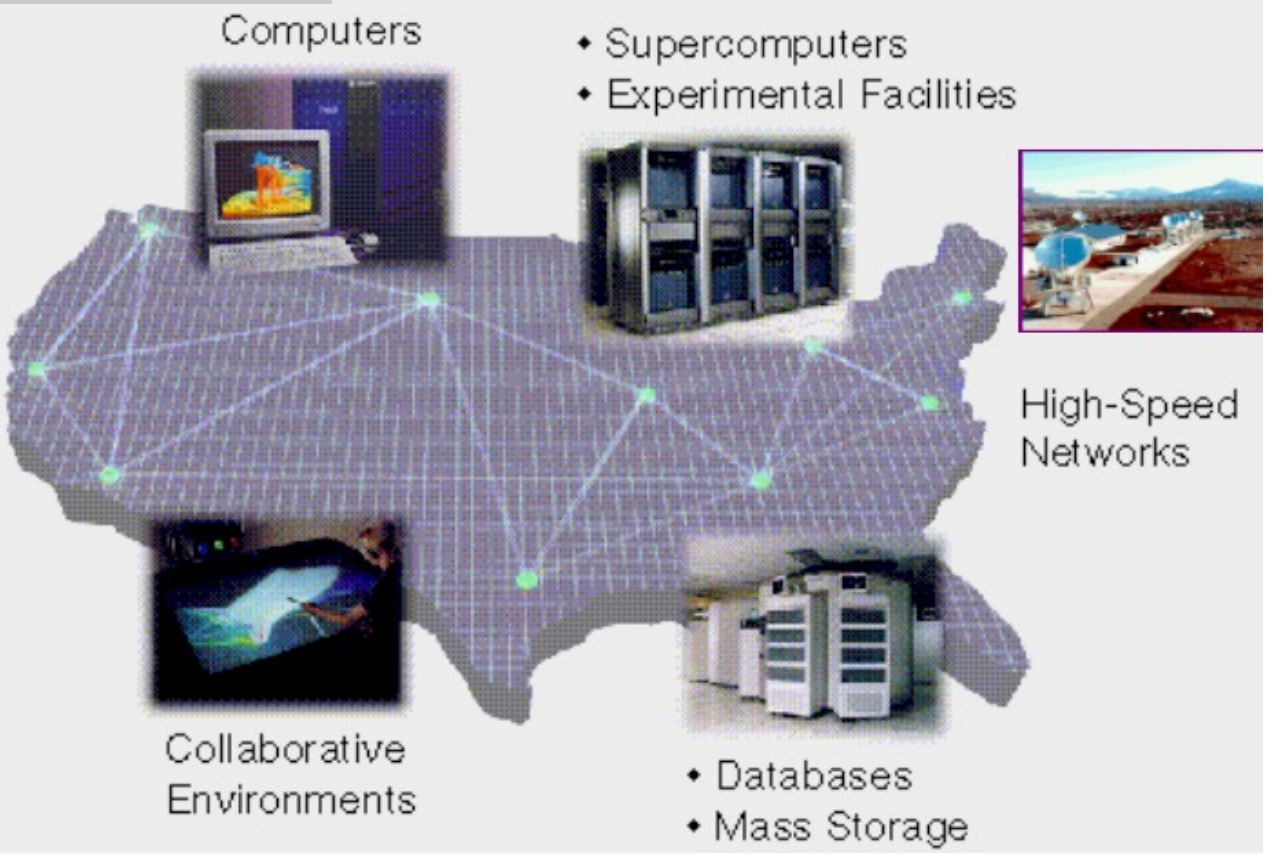
Brian R. Gaines
University of Calgary
gaines@ucalgary.ca

Daniela E.H. Damian
University of Victoria
danielad@cs.uvic.ca

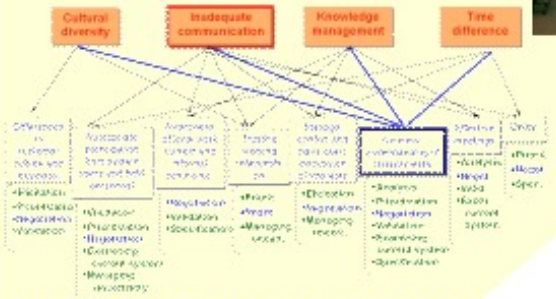


Knowledge Creation in a Grid World

How do we design scientific/scholarly workplace of the future (SWOF)?



Industrial case studies of requirements mgmt in global teams



Requirements engineering challenges due to geographical distribution of project stakeholders

GLOBAL SOFTWARE TEAMS: OPPORTUNITIES AND CHALLENGES

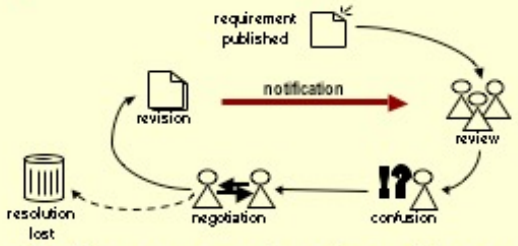
SEGAL Labs

(Software Engineering Global interAction Laboratories)

Dr. Daniela Damian, Director

segal.cs.uvic.ca

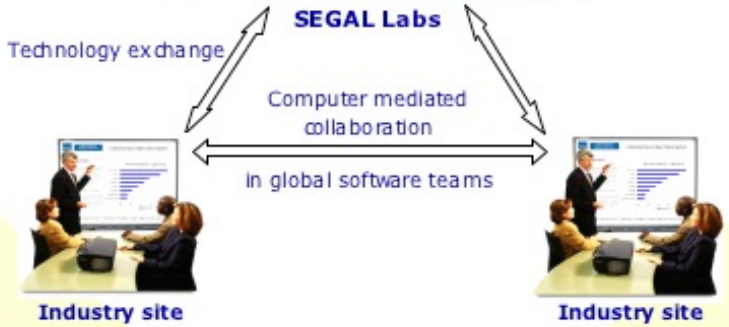
Requirements workspace awarenes



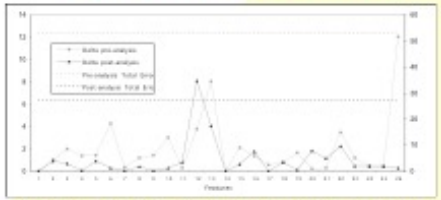
Providing awareness of requirement changes promotes complete and timely requirement communication and documentation

Requirements negotiations in laboratory

Remote customer negotiations conducive to agreements integrative of each negotiator's interests



Industrial case studies of softwar process improvement

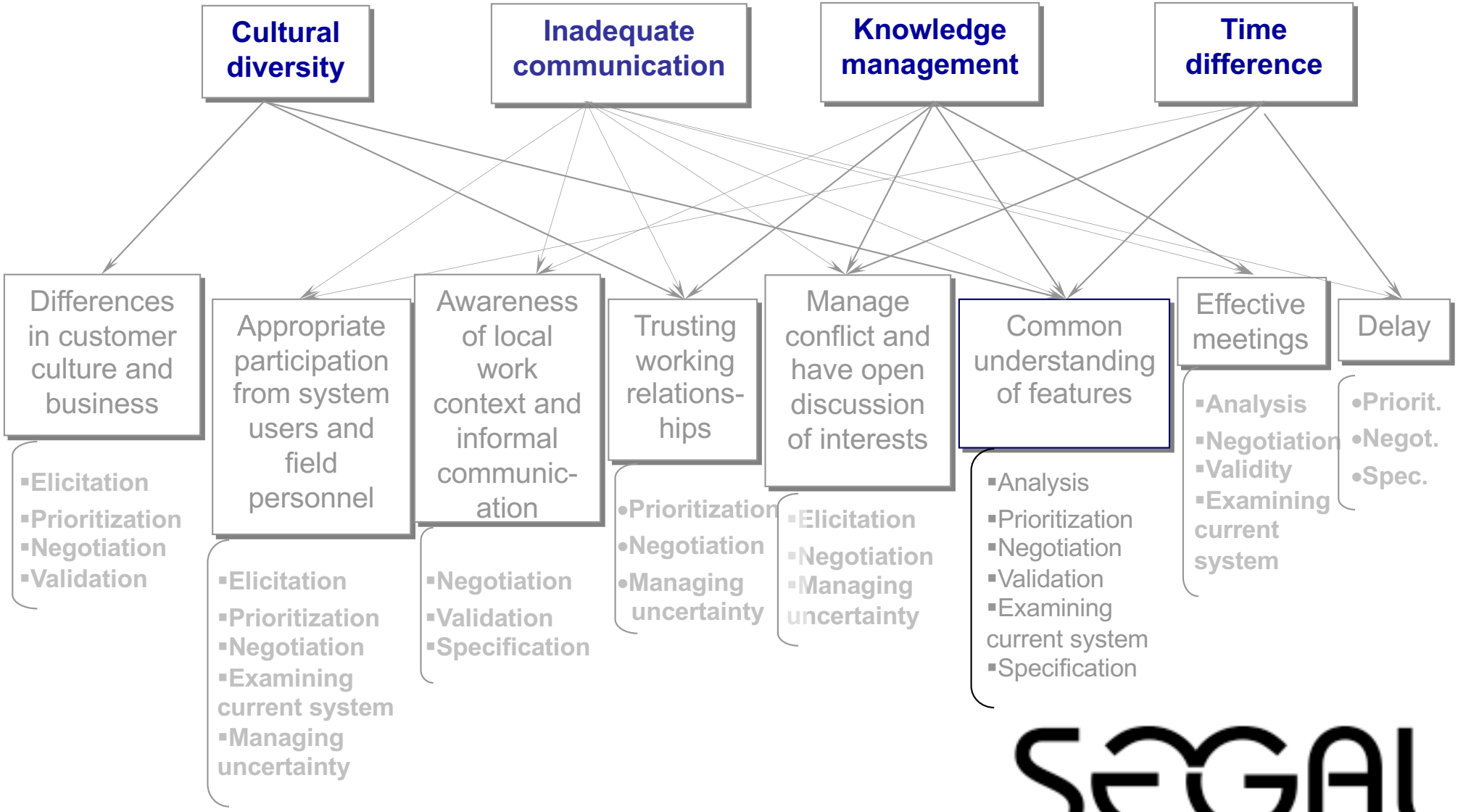


Project planning: significant relationship between requirements practice and estimations

Rich interactions between requirements

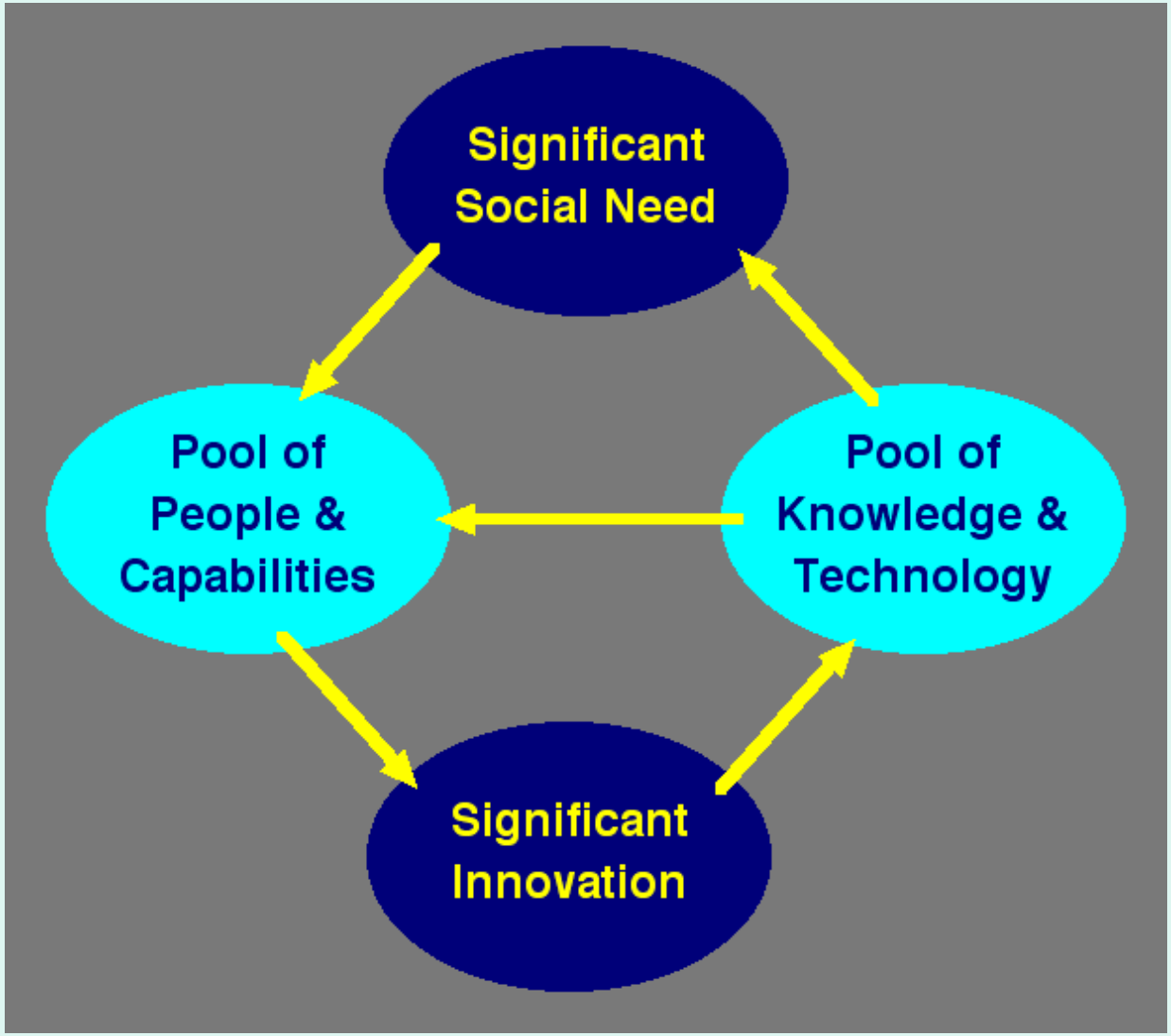
SEGAL Labs
•I Ozkaya E Hargreaves T Holt G Klessig A Quivey I Wang S Scanlan
Research Affiliates
•Dr M-Anne Storey Dr B R Gaines

Challenges in Remote Communication (Damian, 2004)



SEGAL
Software Engineering Global interAction Labs

The Need/Innovation Cycle

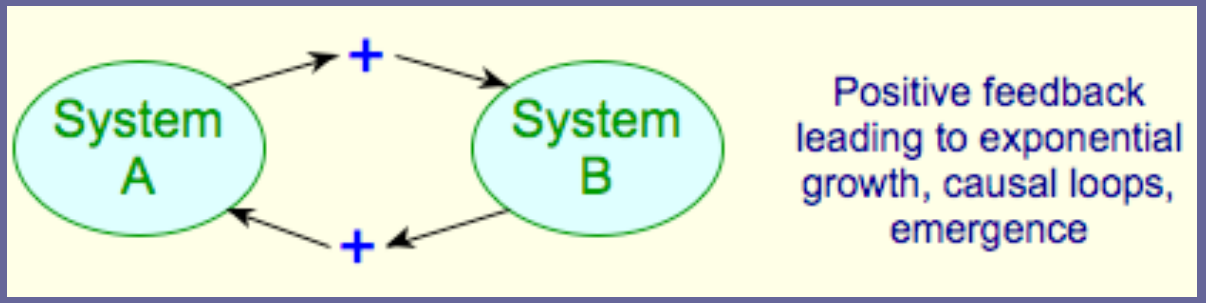


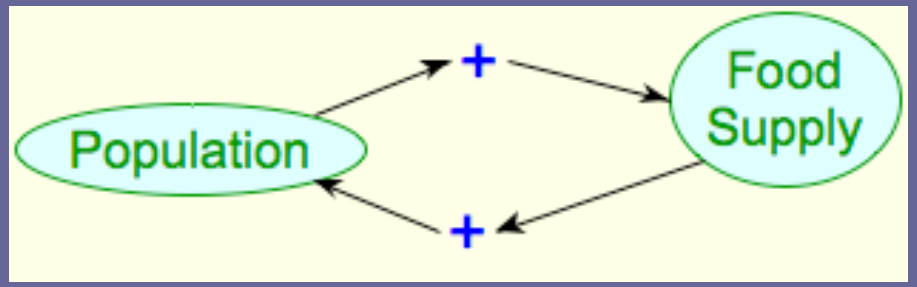
Jerzy Wojciechowski (1983)
"Knowledge is now the primary source of our problems, and more knowledge is required to solve them."

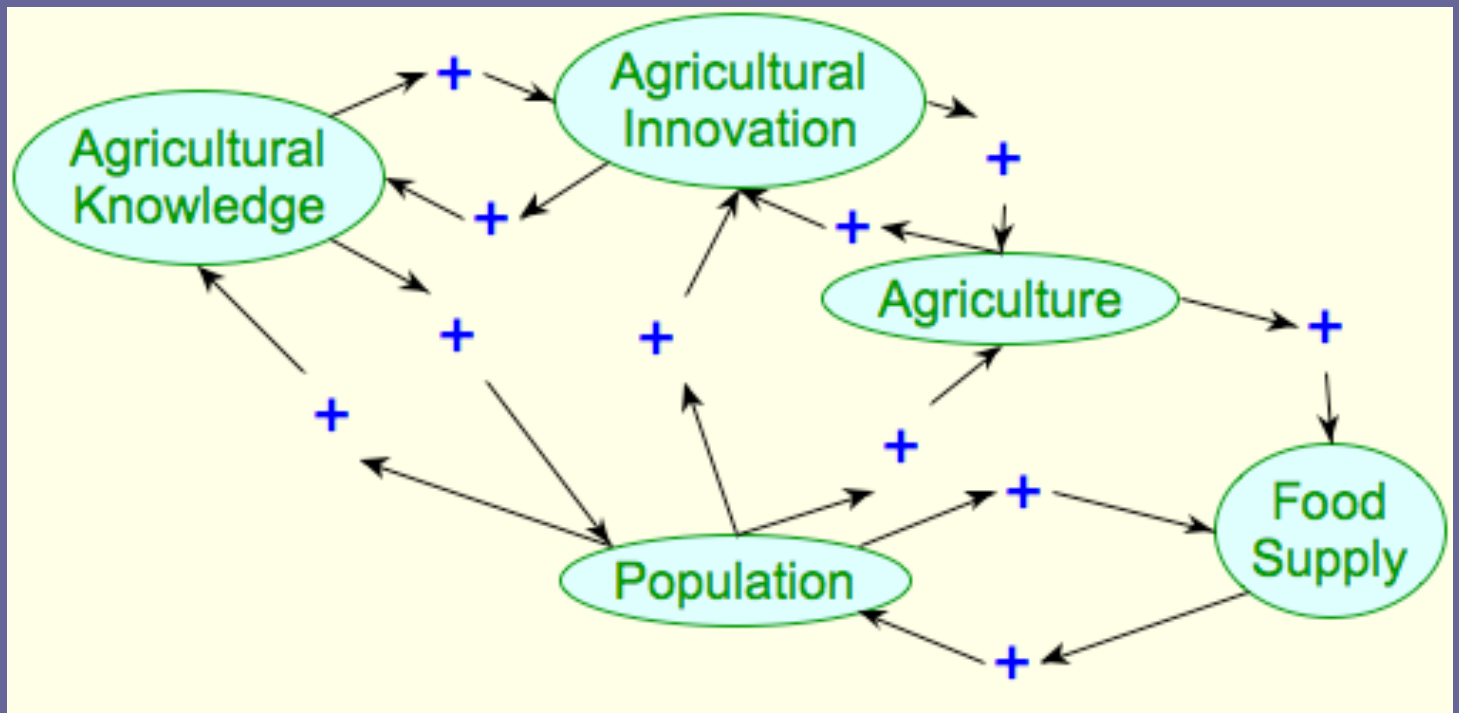
Ester Boserup (1999)
"Invention depends upon changes in motivation and is not related to the frequency of "geniuses" in one population or another."

Colum Gilfillan (1937)
"Inventions can be predicted from needs and possibilities"

Feedback in Influence Networks

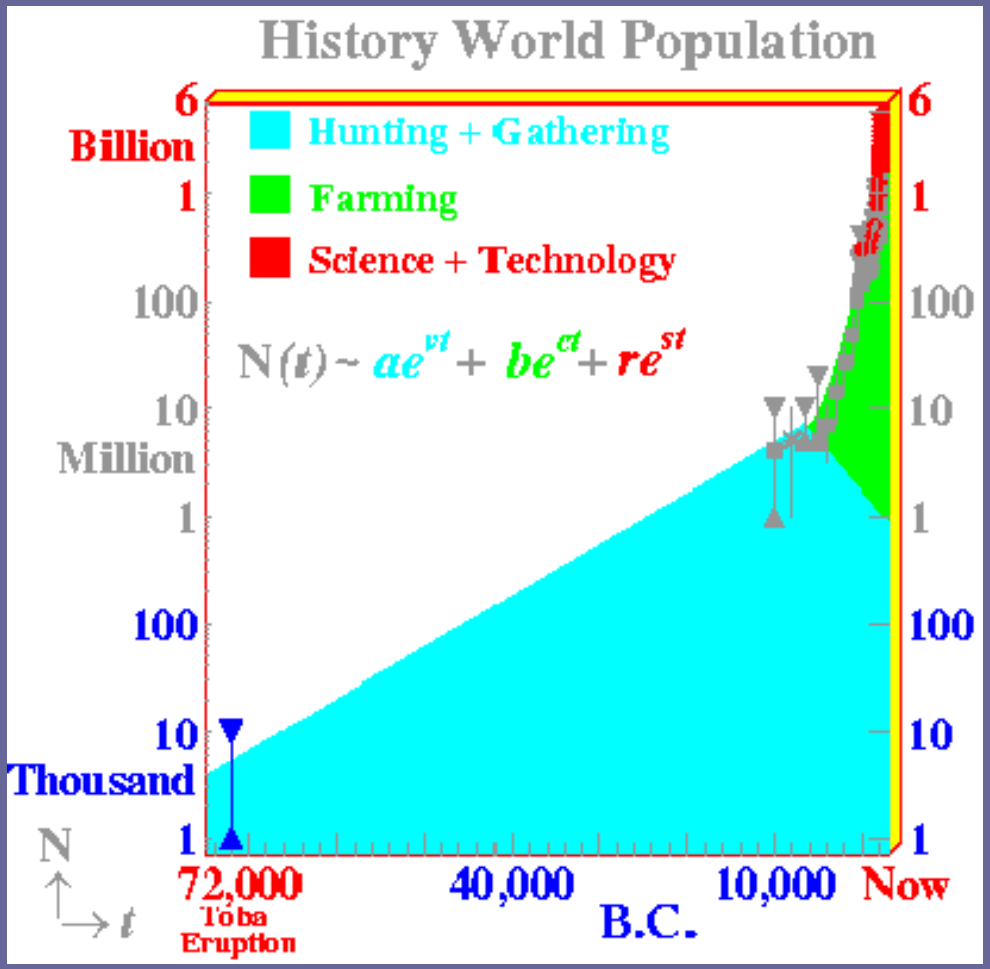


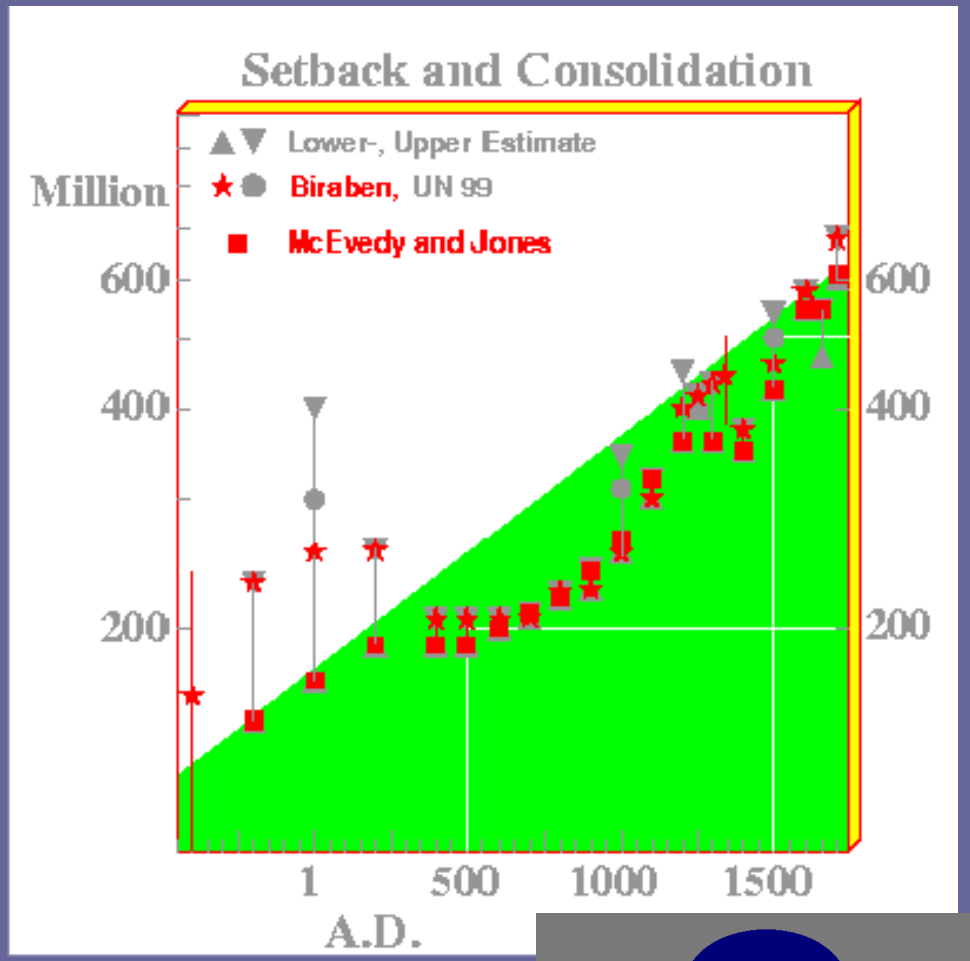


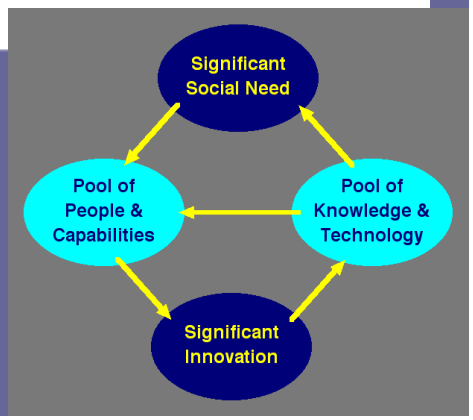


Water
Energy
Transport
Language
Writing
Printing
Trade
Money
Communications
Government
Law
Property rights

Population Growth (Pieter Duinker, 2002)

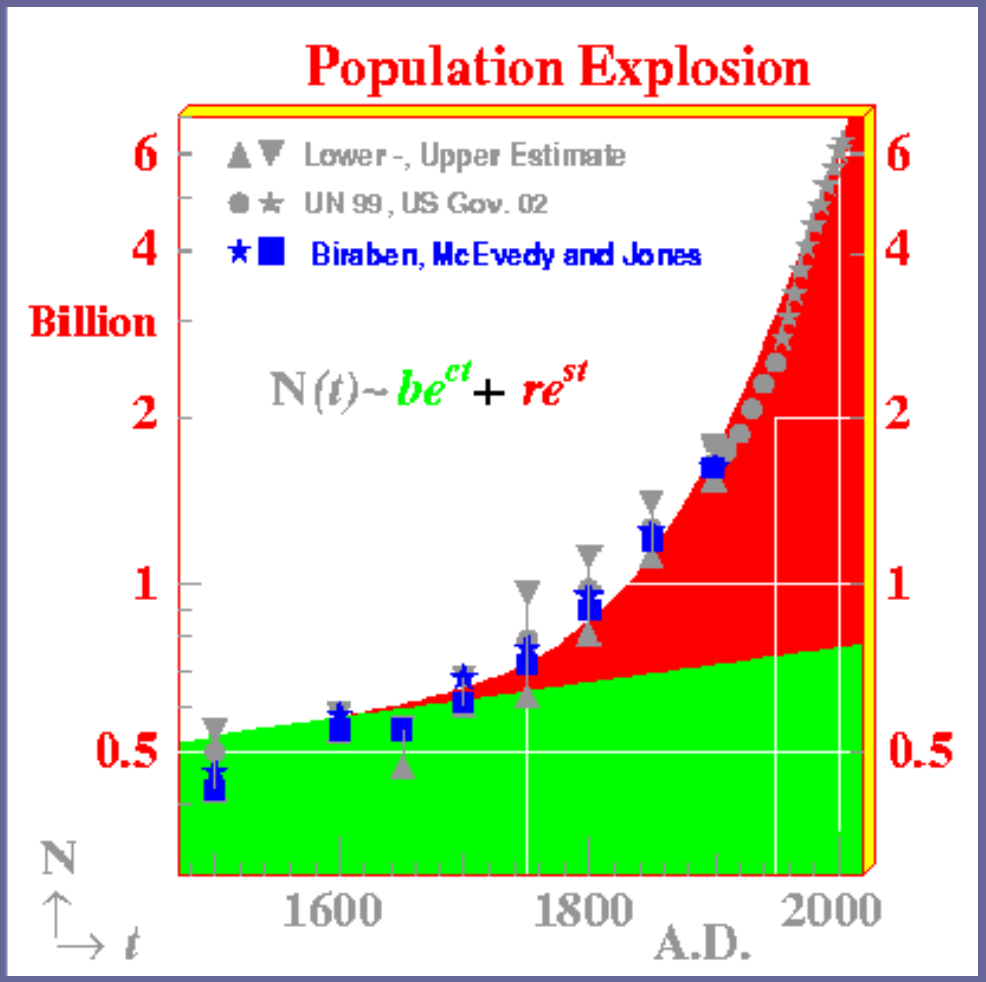


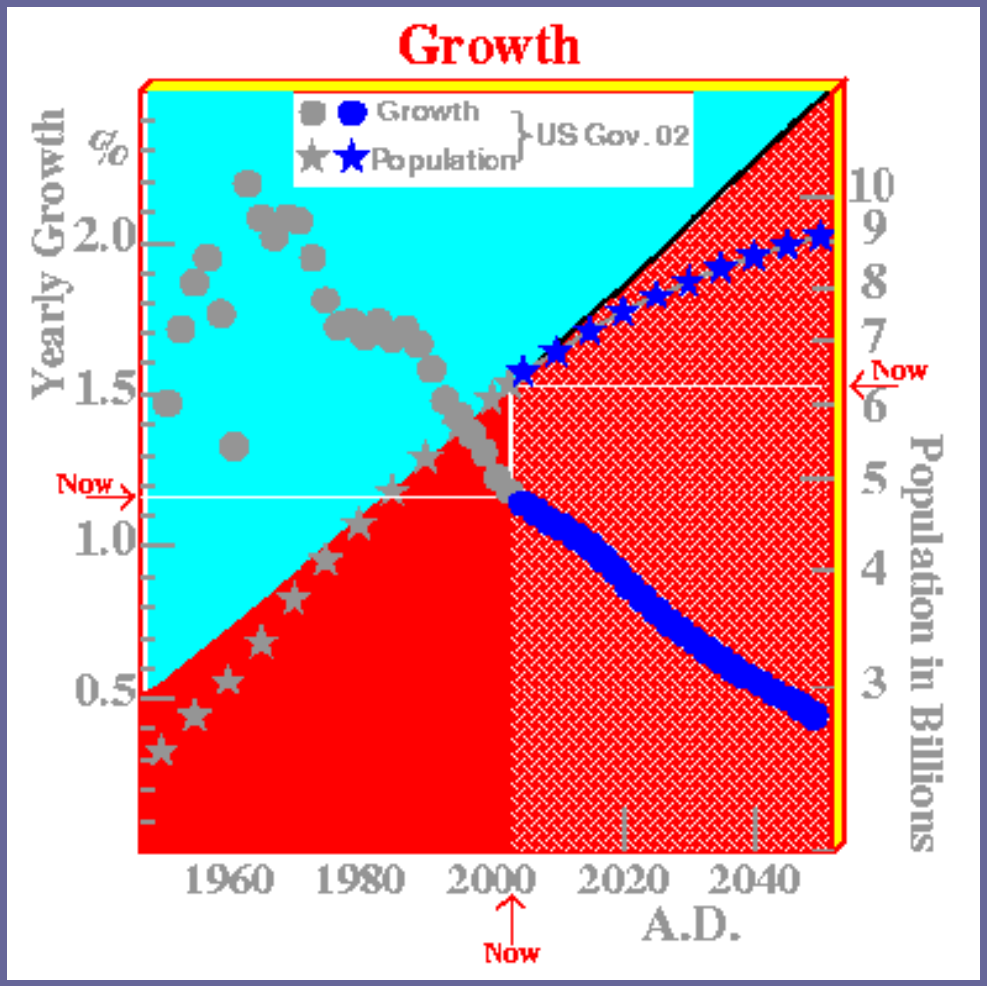




0.014%—0.07%—1.4%

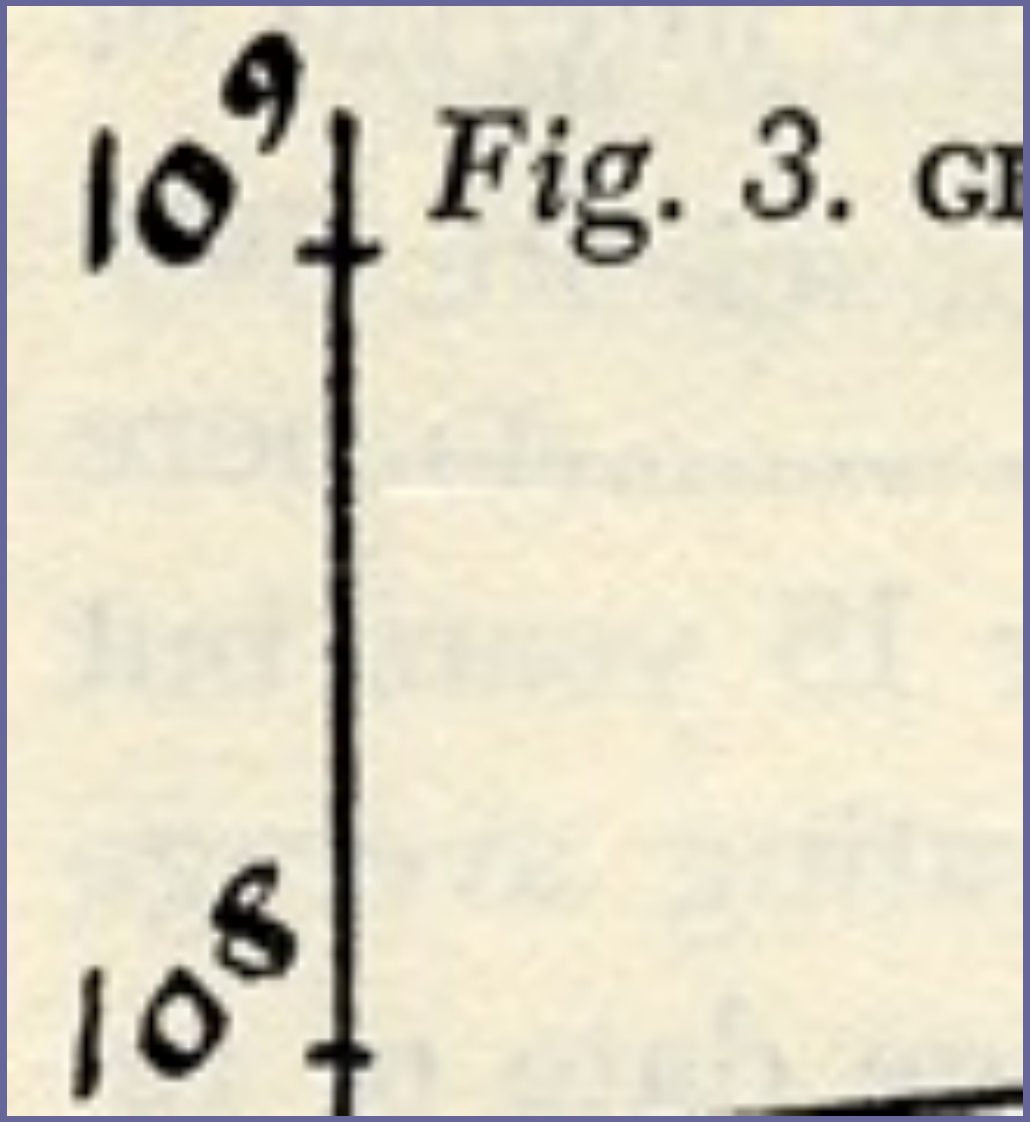
Population Growth (Pieter Duinker, 2002)





0.014%—0.07%—1.4%

Growth in numbers of scientists in the USA (Derek de Solla Price, 1963)



"if every school and college in the US were turned to the exclusive production of physicists…there would still necessarily be a manpower shortage in physics before the passage of another century"

Devices on chip (Cokki Versluis, 2002)

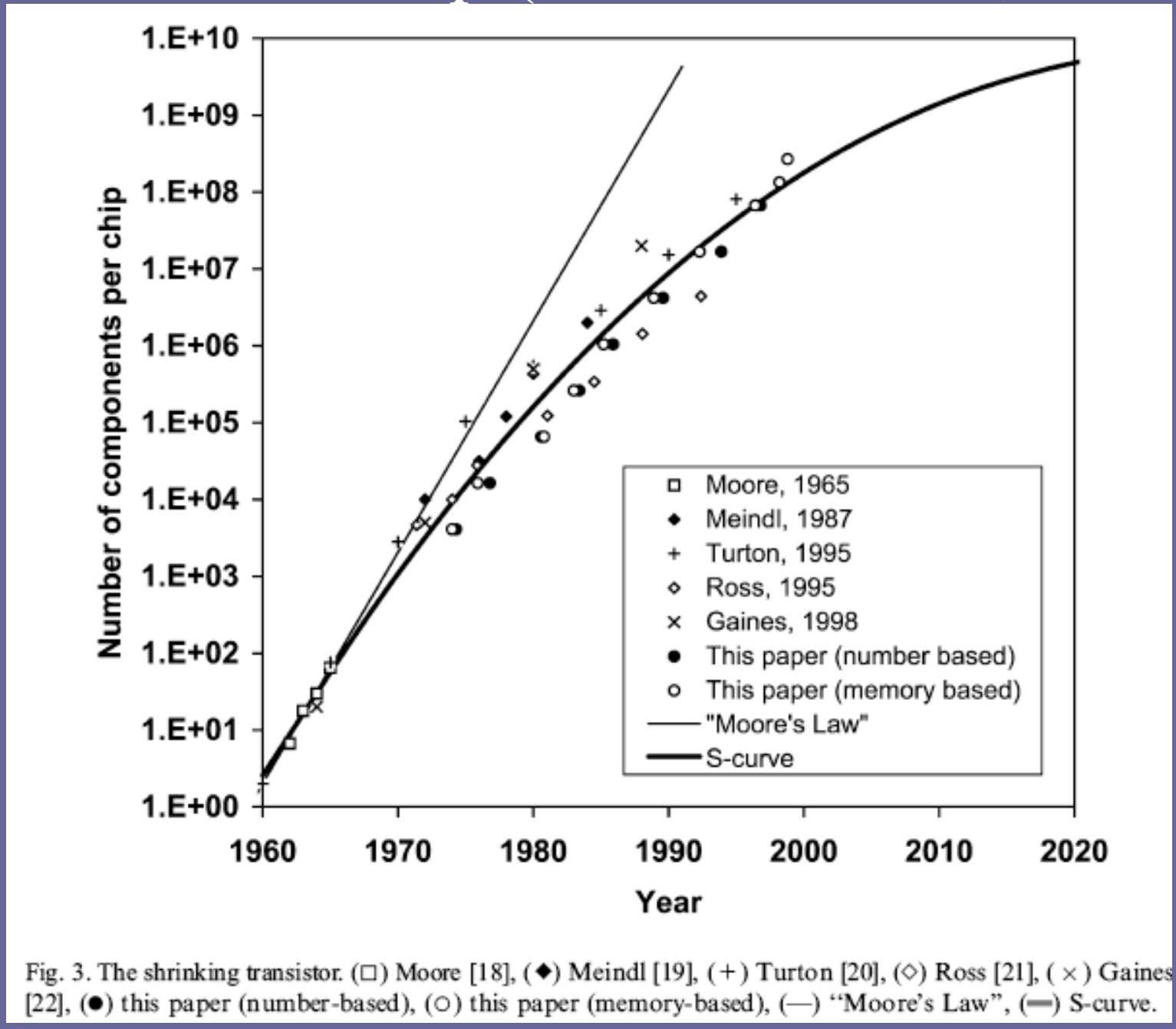


Fig. 3. The shrinking transistor. (□) Moore [18], (◆) Meindl [19], (+) Turton [20], (◇) Ross [21], (×) Gaines [22], (●) this paper (number-based), (○) this paper (memory-based), (—) "Moore's Law", (═) S-curve.

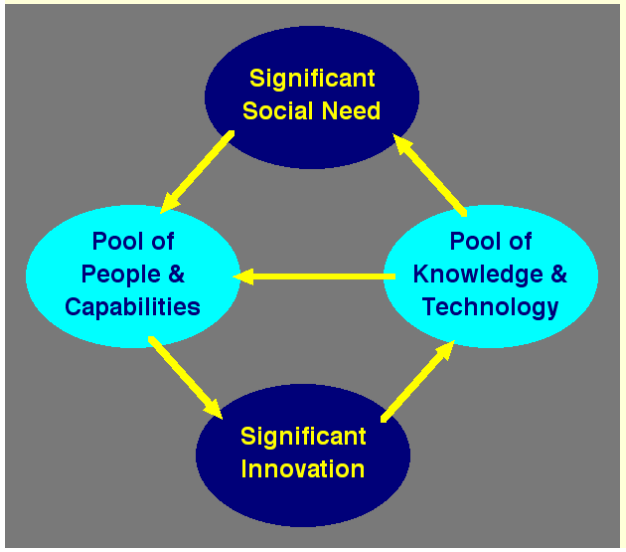
What is the Internet? (Ed Krol 1993)

What is the Internet?
RFC1462, 1993, Ed Krol

1.a **network of networks** based on the TCP/IP protocols,

2.a **community of people** who use and develop those networks,

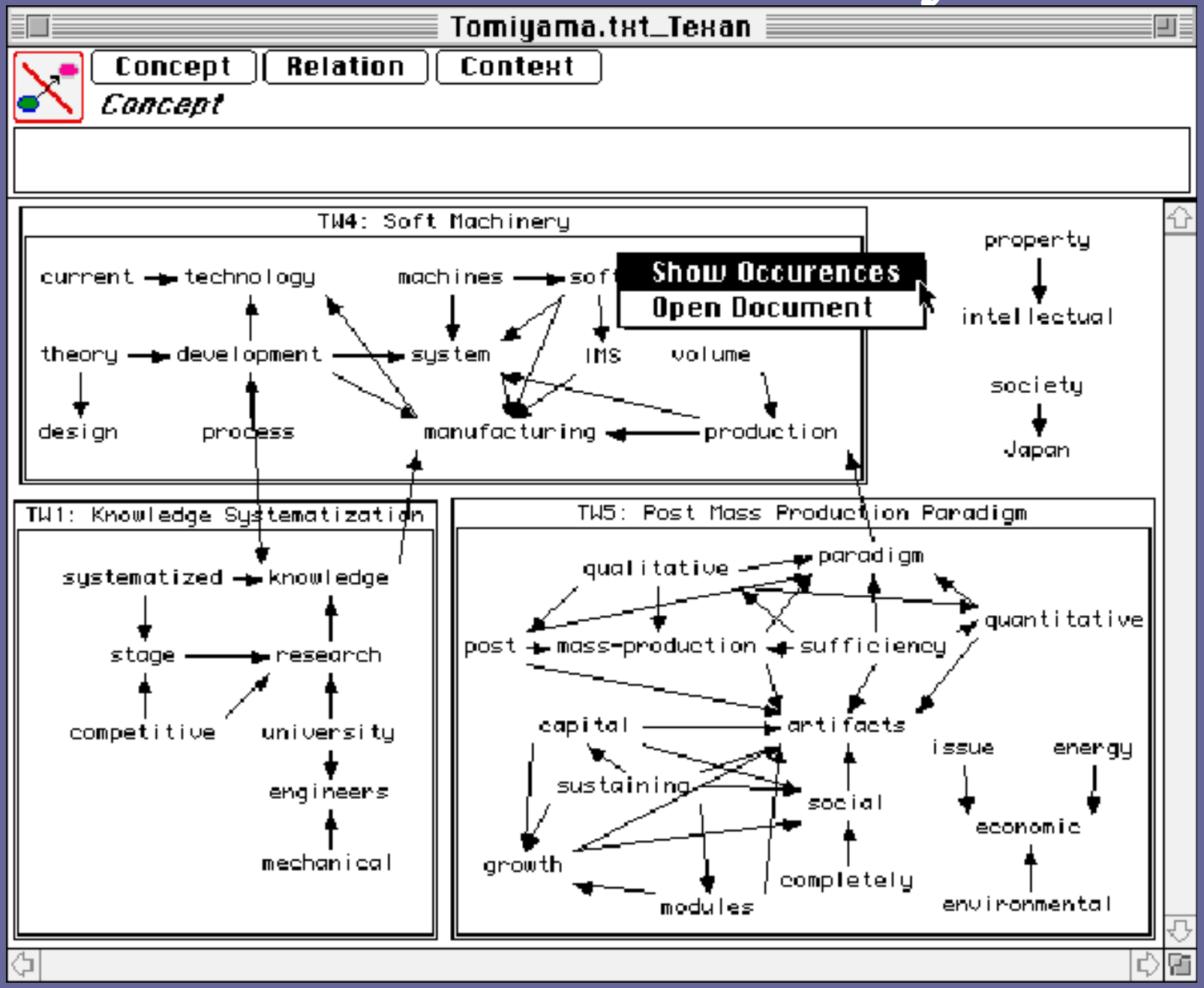
3.a **collection of resources** that can be reached from those networks.

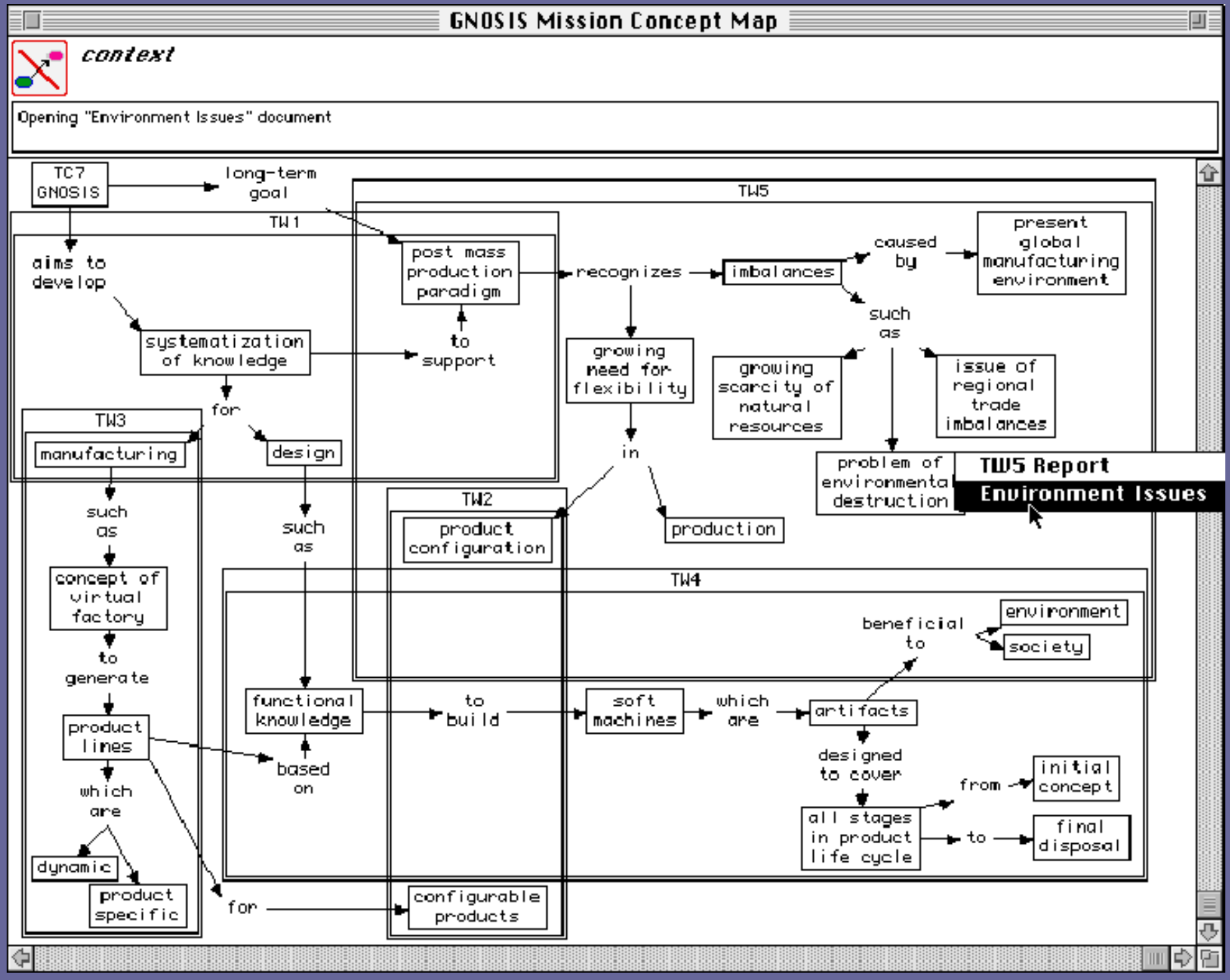


Supporting Research Communities

- GNOSIS IMS Consortium
- GNOSIS communications
- Automatic text analysis
- Elicited mental model
- Mediator—supporting a distributed community
- CHRONO—awareness of change
- Knowledge support system
- Impact of CMC usage on negotiation outcomes

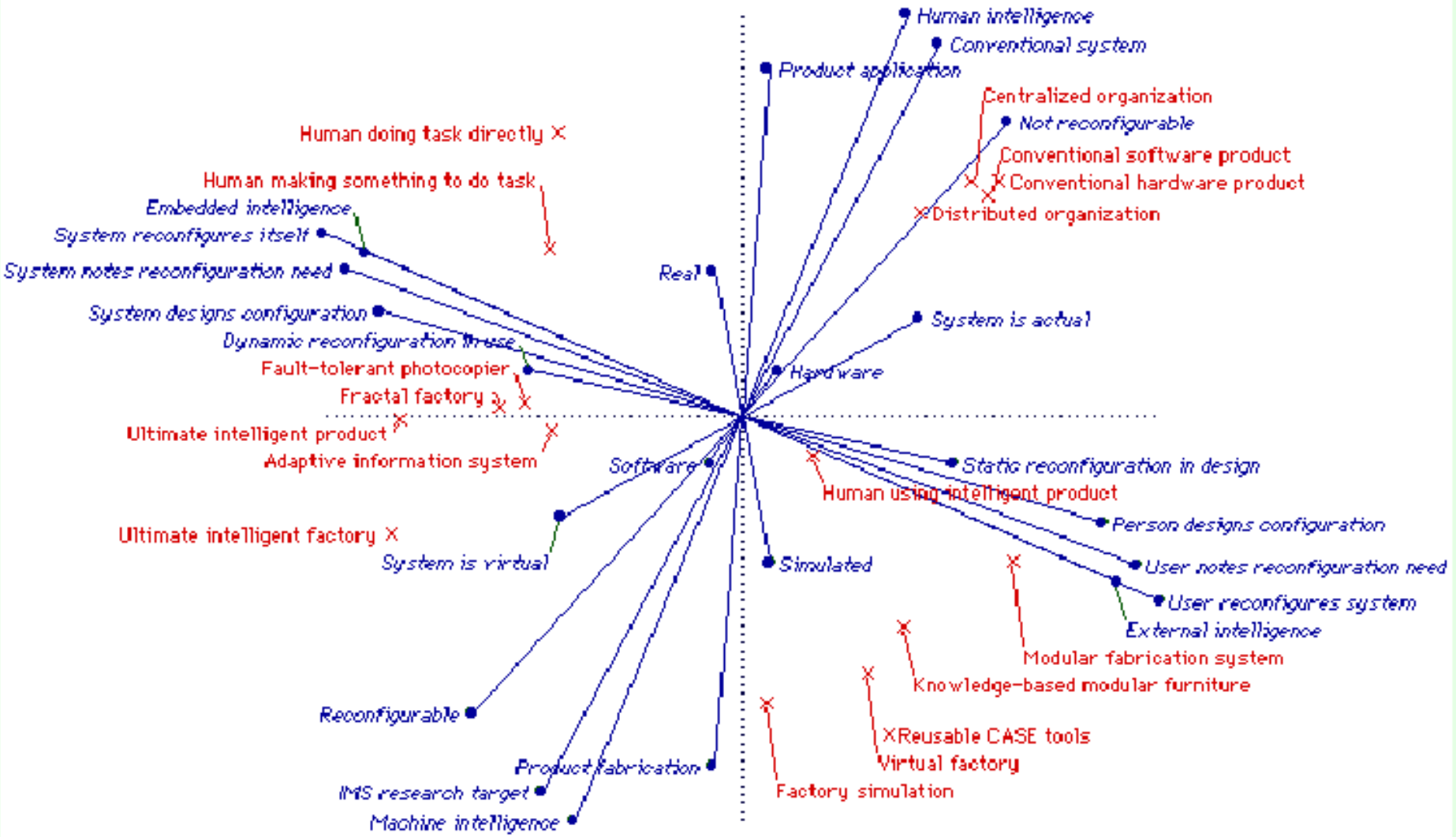
Automatic text analysis




GNOSIS Mission Concept Map
context
Opening "Environment Issues" document
TC7 GNOSIS
long-term goal
TW5
TW1
aims to develop
post mass production paradigm
recognizes
imbalances
caused by
present global manufacturing environment
such as
systematization of knowledge
to support
growing need for flexibility
growing scarcity of natural resources
issue of regional trade imbalances
for
TW3
manufacturing
design
problem of environmental destruction
TW5 Report
Environment Issues
in
TW2
product configuration
production
such as
such as
concept of virtual factory
TW4
beneficial to
environment
society
to generate
functional knowledge
to build
soft machines
which are
artifacts
product lines
based on
designed to cover
from
initial concept
which are
all stages in product life cycle
to
final disposal
dynamic
product specific
for
configurable products

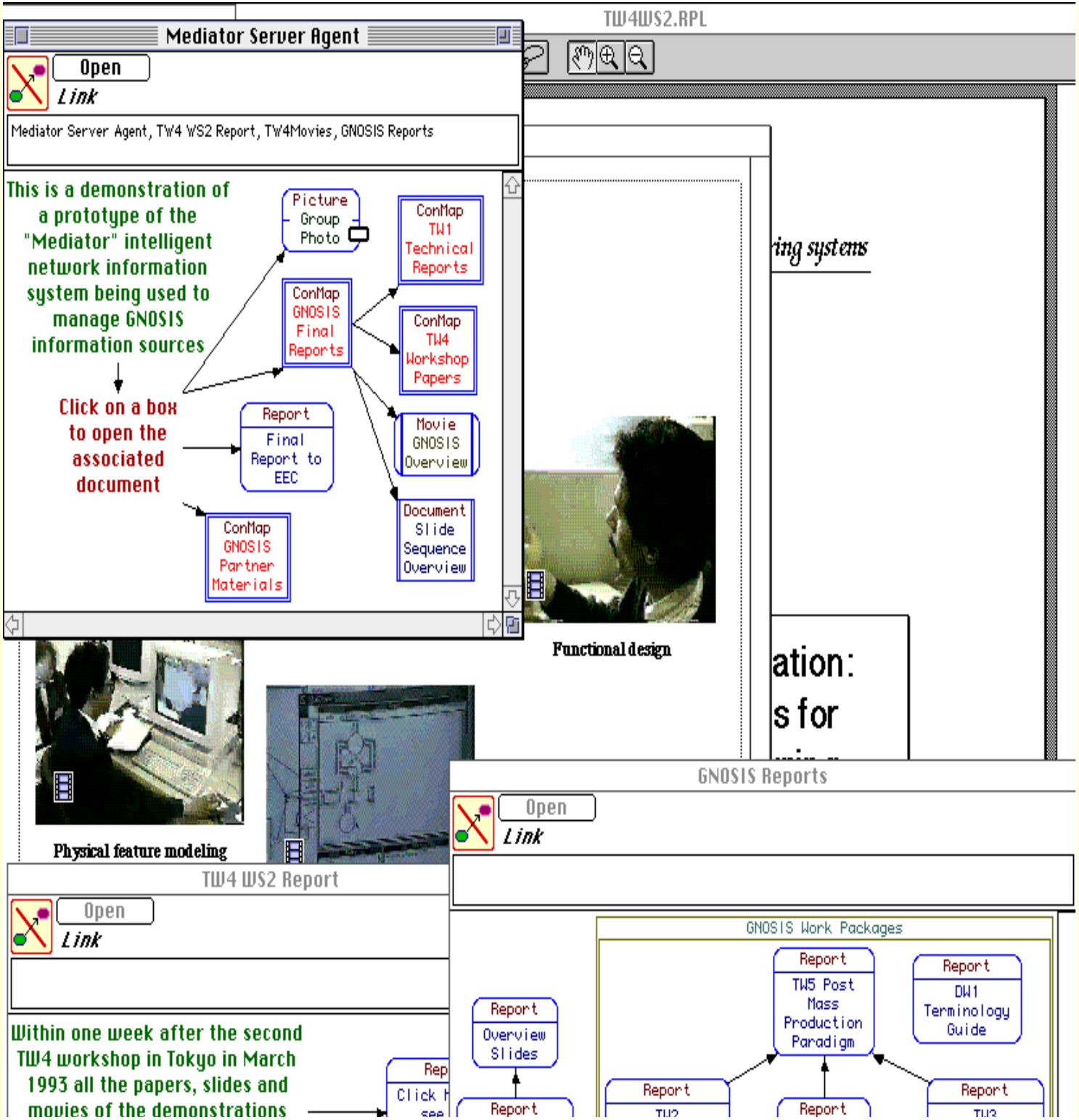
WebGrid-II *PrinCom Map*

PrinCom, Domain: Soft Machines, User: Group
Context: Dimensions of soft machines, 17 systems, 12 qualities

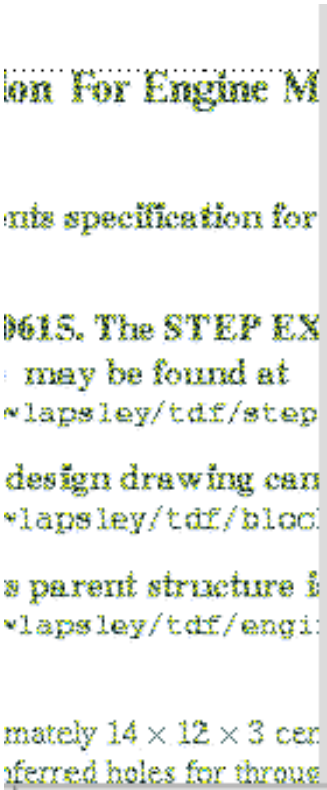


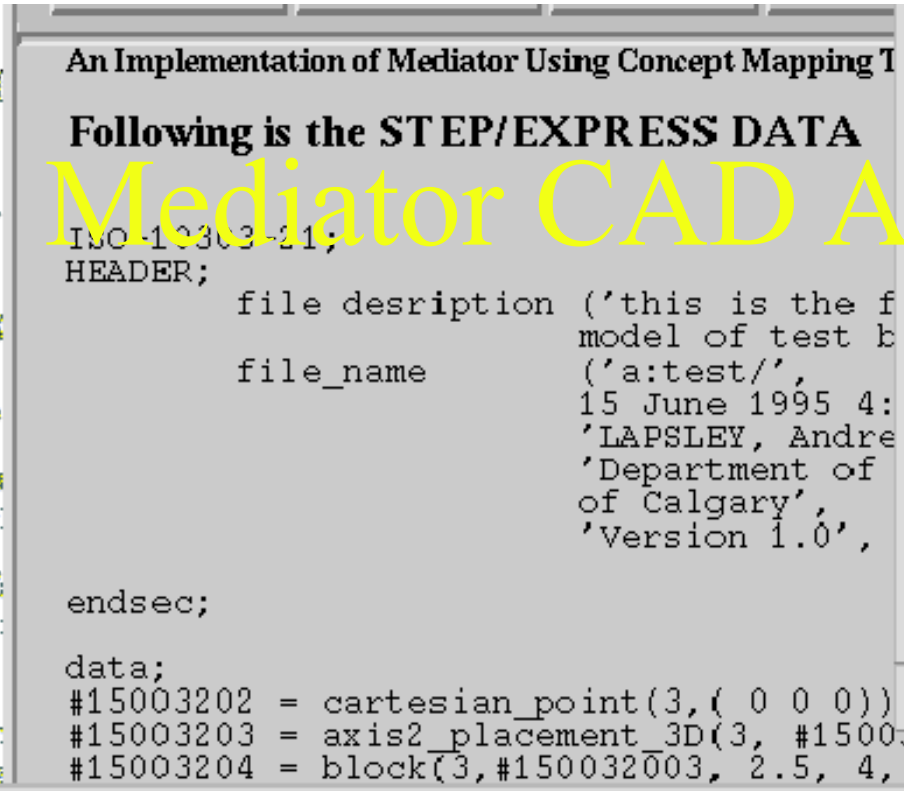
Click on title, systems or qualities to edit them

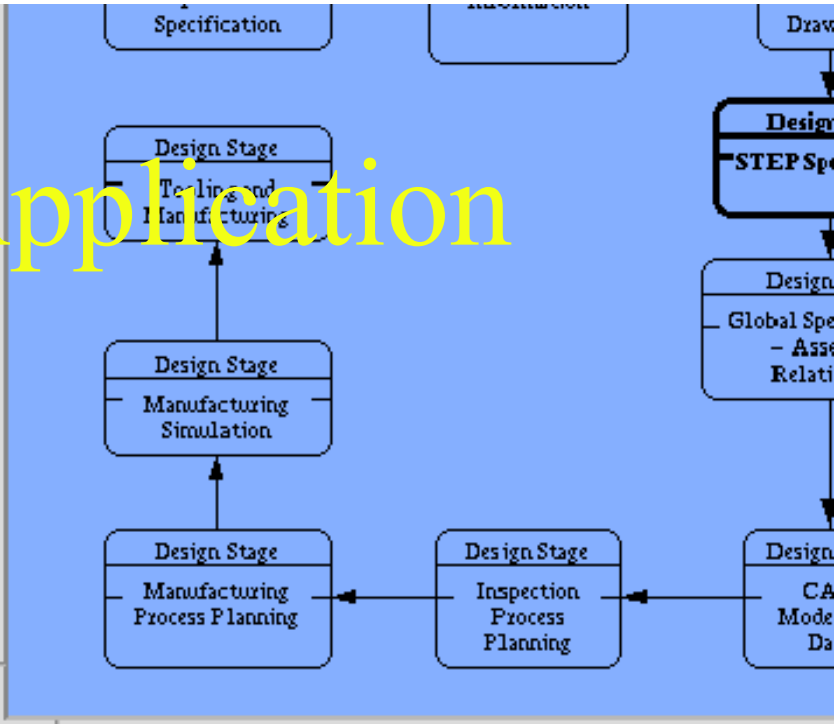
Mediator

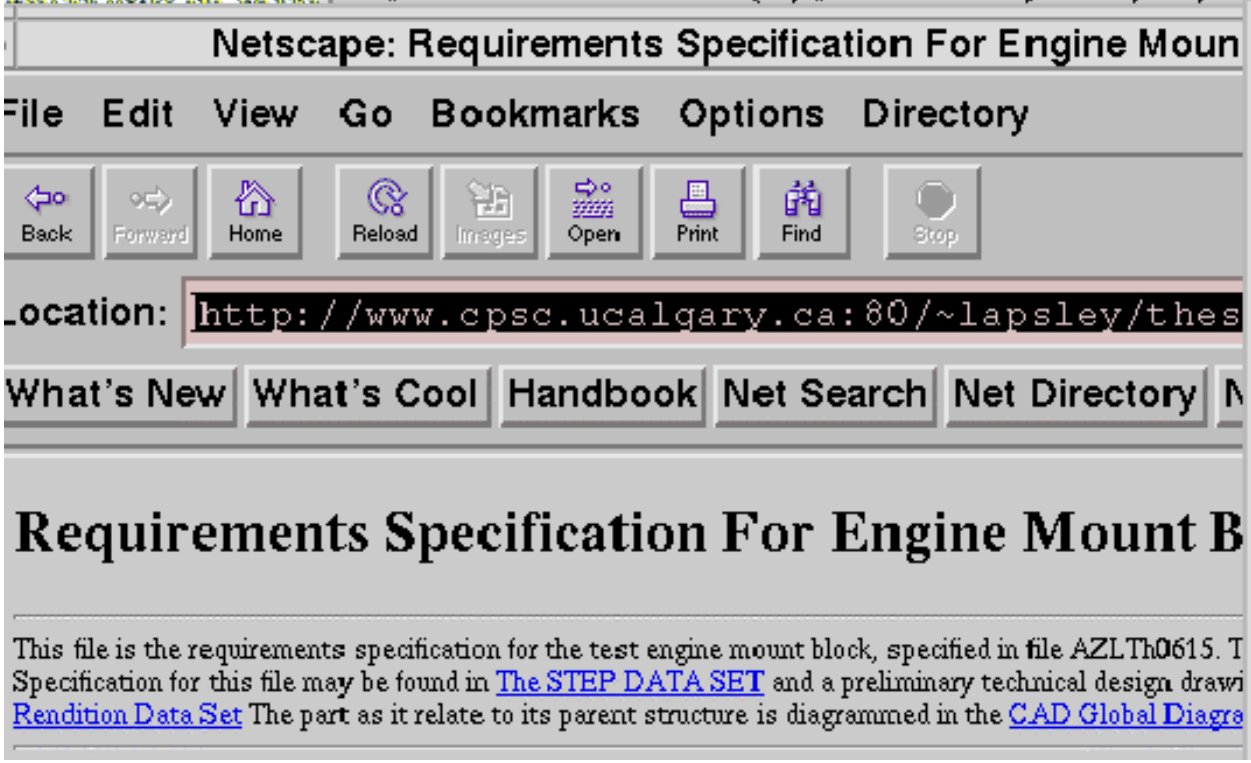

TW4WS2.RPL
Mediator Server Agent
Open
Link
Mediator Server Agent, TW4 WS2 Report, TW4Movies, GNOSIS Reports
This is a demonstration of a prototype of the "Mediator" intelligent network information system being used to manage GNOSIS information sources
Click on a box to open the associated document
Picture Group Photo
ConMap TW1 Technical Reports
ConMap GNOSIS Final Reports
ConMap TW4 Workshop Papers
Report Final Report to EEC
Movie GNOSIS Overview
Document Slide Sequence Overview
ConMap GNOSIS Partner Materials
Functional design
Physical feature modeling
GNOSIS Reports
Open
Link
TW4 WS2 Report
Open
Link
Within one week after the second TW4 workshop in Tokyo in March 1993 all the papers, slides and movies of the demonstrations
GNOSIS Work Packages
Report TW5 Post Mass Production Paradigm
Report DW1 Terminology Guide
Report Overview Slides

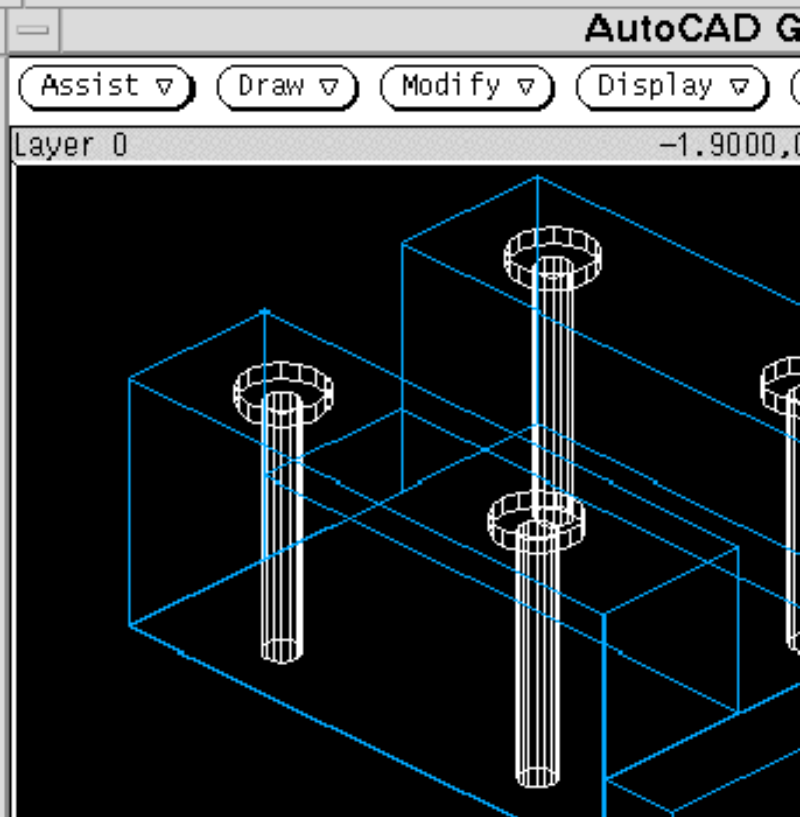
Mediator CAD Application



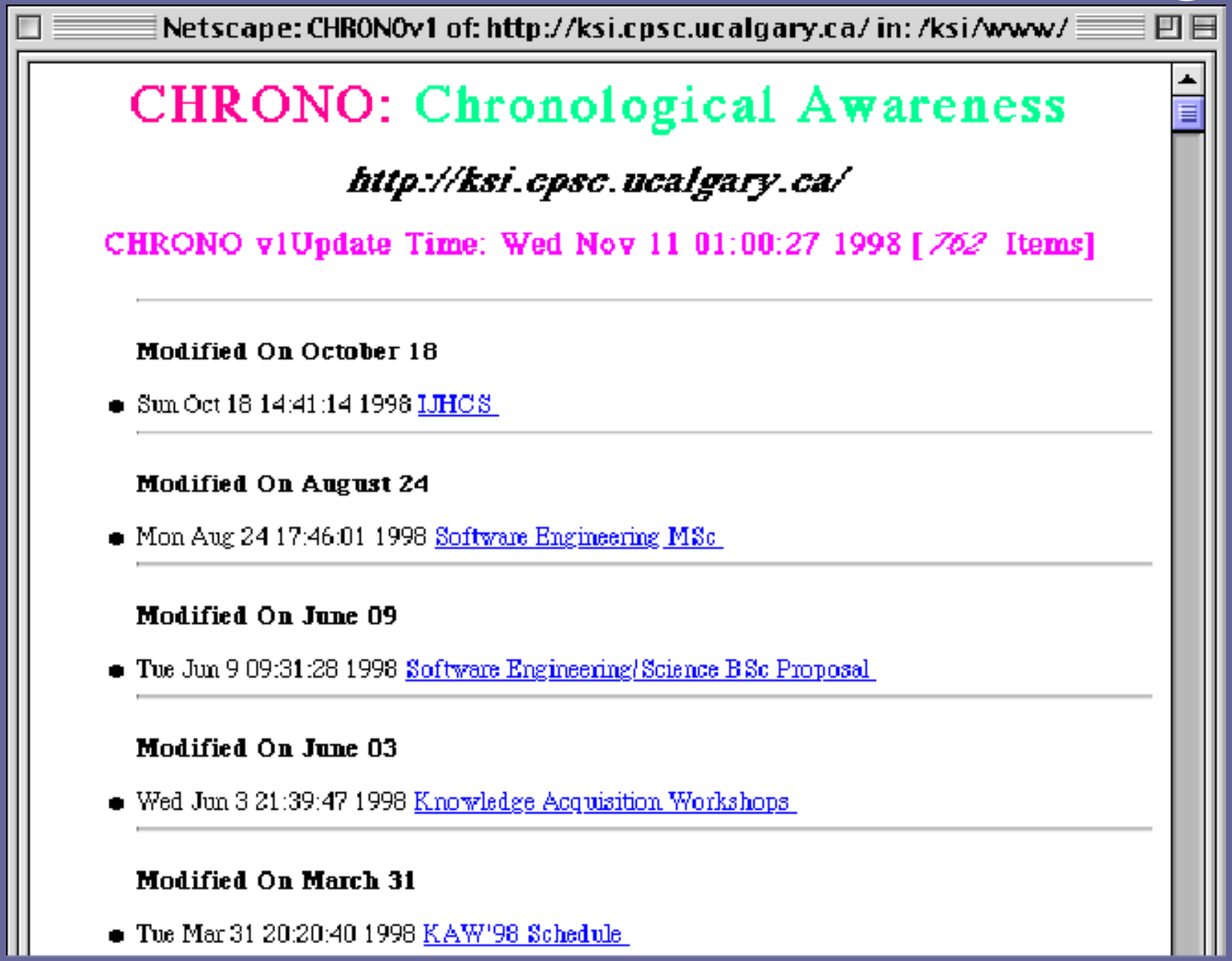




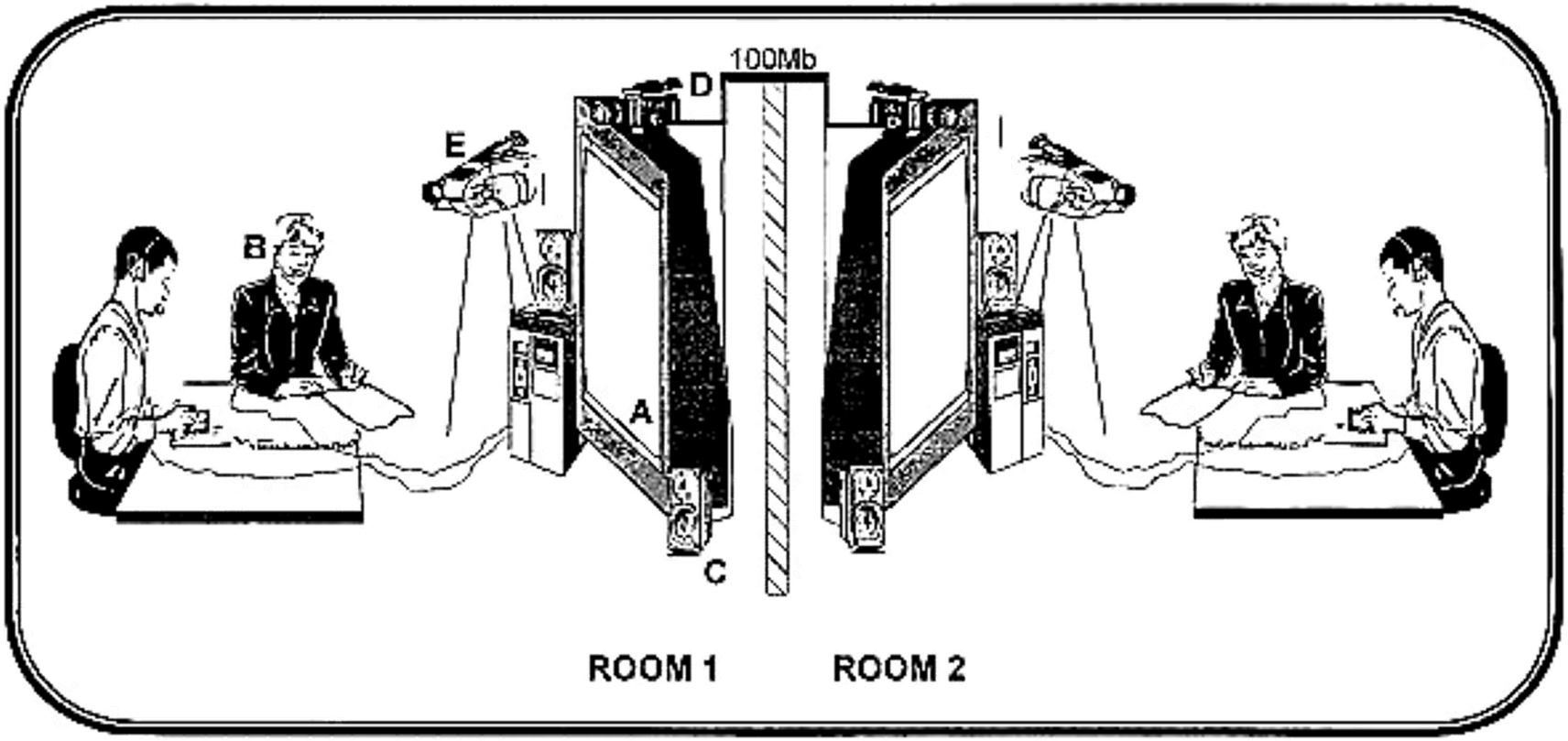




CHRONO—awareness of change



A study of requirements negotiations in laboratory setting



SEGAL
Software Engineering Global interAction Labs

Some relevant technologies

- Access Grid
- E-meetings: research issues
- Shared space — community document management
- Reproducible scientific computations
- Active documents
- CiteSeer
- Innovation through literature linkages
- Consequences of scientific definitions

Access Grid

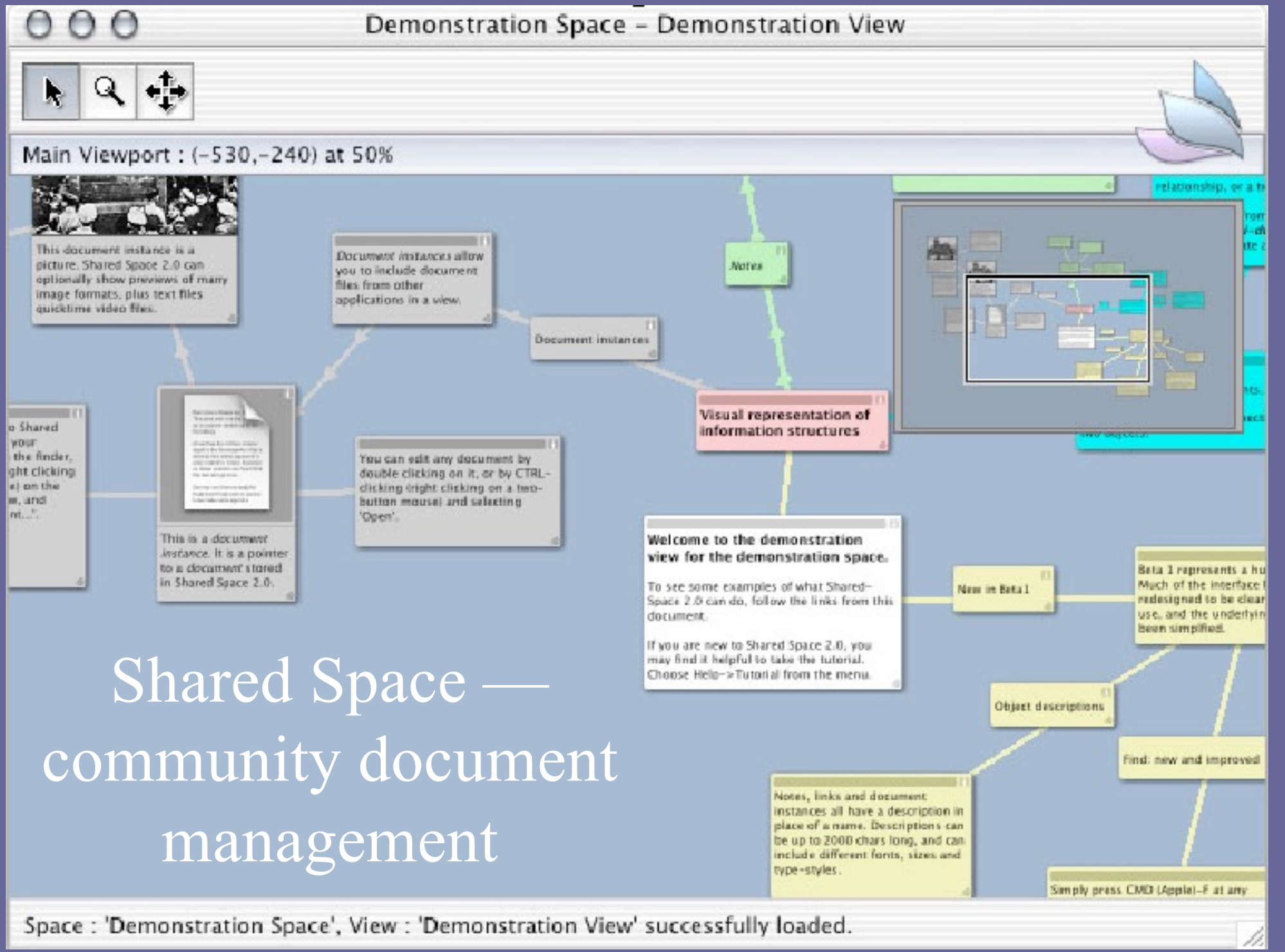
(Rick Stevens, *et al*)
how would it be if travel
was free and instantaneous?

- You would live where you want to
- You and your colleagues would meet in interesting places
- You would leave your things where you need or use them
- You would arrange to meet people in certain places at certain times
- You would need mechanisms to locate people and things
- Security might be problematic
- You would develop habitual hangout places

Slide derived from those of Rick Stevens, Argonne National Laboratory

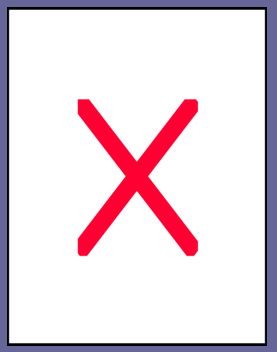
E-meetings: research issues

- What are the functions of scientific meetings?
- —to participate in presentations?
- —to hear an up-to-date, less formal account?
- —to understand more through critical discussion?
- —to assess the credibility of significant agents?
- —to have the opportunity to question significant agents?
- —to identify members of a social network & position oneself in it?
- —to draw attention to oneself and one's research?
- —to trade information?
- —to develop collaborations?
- —to mentor and support others?
- —to calibrate one's constructions through 'approbation' and 'imitation' (Adam Smith, 1776)
- —plus, plus, plus — what is missing?
- Does the technology support all these—what is needed?



Shared Space — community document management

Reproducible Scientific Computations (Jon Claerbout)



“ we noticed that a few months after completing a project, the researchers at our laboratory were usually unable to reproduce their own computational work without considerable agony.”
“we solved this problem by developing a concept of electronic documents that makes scientific computations reproducible.”

ReDoc rules:-

- `make burn` removes the result files such as the postscript a computed figure;

- `make build` recomputes the result files. It compiles the necessary programs, invokes a sequence of executables, and stores the results in predefined files;

- `make view` updates the result files & displays them using the appropriate viewer;

- `make clean` removes all the intermediate files that were created during the execution of the `build` command. It leaves the result and source files intact.

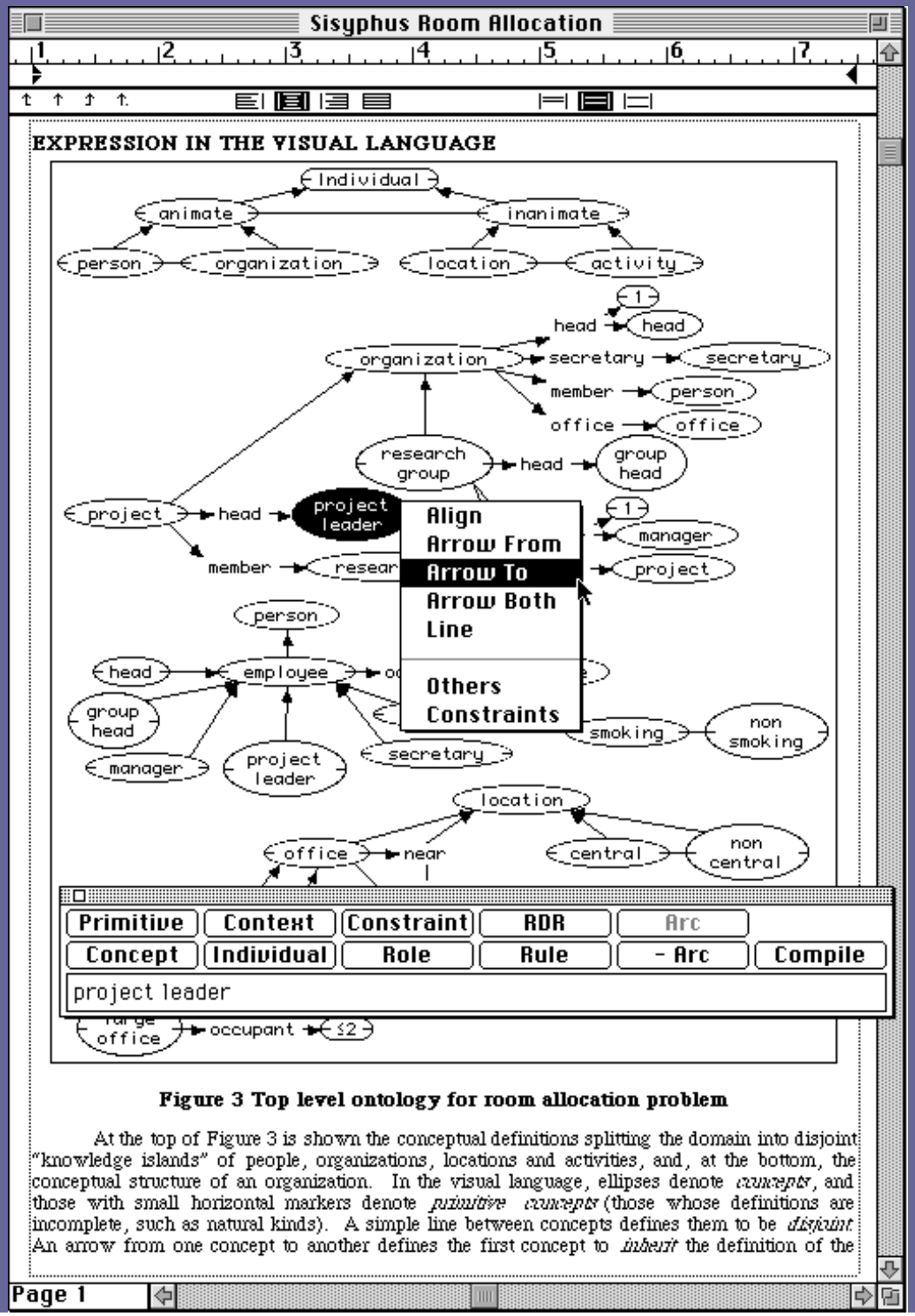


Figure 3 Top level ontology for room allocation problem

At the top of Figure 3 is shown the conceptual definitions splitting the domain into disjoint "knowledge islands" of people, organizations, locations and activities, and, at the bottom, the conceptual structure of an organization. In the visual language, ellipses denote *concepts*, and those with small horizontal markers denote *primitive concepts* (those whose definitions are incomplete, such as natural kinds). A simple line between concepts defines them to be *disjoint*. An arrow from one concept to another defines the first concept to *inherit* the definition of the

Active documents with embedded knowledge structures (Gaines & Shaw)

Paper published in BCSES'92 proceedings, and on CD-ROM and through ftp as active documents where embedded semantic nets could be edited and used to infer a situated-action solution of a room allocation problem.

Using Knowledge Acquisition and Representation Tools to Support Scientific Communities (1994) (Make Corrections) (15 citations)

Brian Gaines
National Conference on Artificial Intelligence

CiteSeer Scientific Literature Digital Library Home/Search Context Related

View or download:
repgrid.com/reports/KBS/AA...AAAI94.pdf
Cached: PS.gz PS PDF DjVu Image Update Help

From: repgrid.com/reports/abstracts (more)
(Enter author homepages)

(Enter summary)

Rate this article: 1 2 3 4 5 (best)
Comment on this article

Abstract: Widespread access to the Internet has led to the formation of geographically dispersed scientific communities collaborating through the network. The tools supporting such collaboration currently are based primarily on electronic mail through mailing list servers, and access to archives of research reports through ftp, gopher and world wide web. However, electronic communication can support the knowledge processes of scientific communities more directly through overtly represented knowledge... (Update)

Context of citations to this paper: More

...and numerical data types in addition to rating scales, and also allows symbolic values such as any , none , and not applicable . **Gaines (1993, 1994) has developed a complete semantics for CLASSIC and similar KL ONE derivatives in terms of a matrix in which the basic**...

.... **and Gaines, 1987) and to model the conceptual structure of documents specifying the objectives of collaborative communities (Gaines and Shaw, 1994).** Knowledge Elicitation System The knowledge elicitation sub system supports the development of conceptual structures through...

Cited by: More
Utility Revision in A Java-based Group Decision Support System - Ito, Shintani (Correct)
Mobile Digital Assistants for Community Support - Ishida (1998) (Correct)
Supporting Collaboration through Multimedia Digital Document.. - Gaines (Correct)

Similar documents (at the sentence level):
17.9%: Knowledge Management for Distributed Enterprises - Gaines, Norrie, Lapsley, Shaw (1996) (Correct)

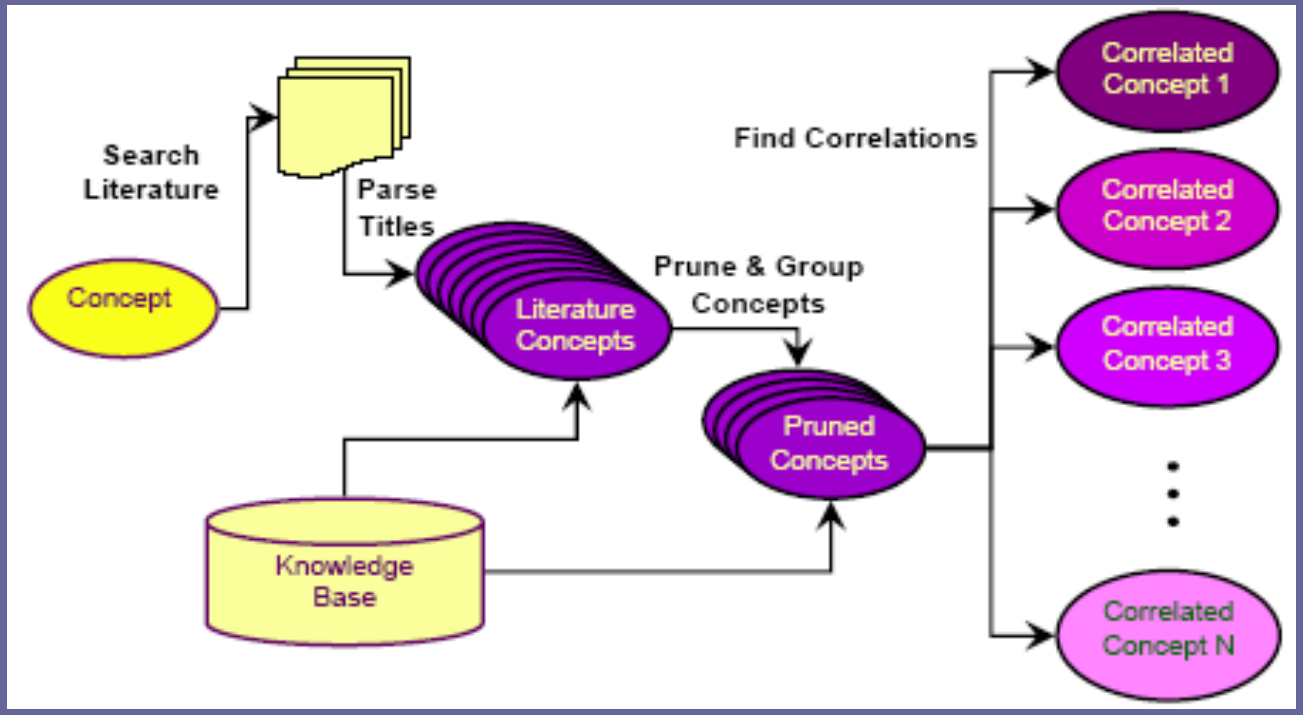
Active bibliography (related documents): More All
1.2: Knowledge Support Systems For Constructively Channeling.. - Mildred Shaw And (Correct)
1.1: Concept Maps as Hypermedia Components - Gaines, Shaw (1995) (Correct)
1.0: An Agenda For Digital Journals: The Socio-Technical.. - Gaines (1993) (Correct)

Similar documents based on text: More All
0.3: Eliciting Knowledge And Transferring It Effectively To A.. - Gaines (1993) (Correct)
0.3: Foundations for the Learning Web - Brian Gaines (Correct)
0.3: Integrated Knowledge Acquisition Architectures - Gaines, Shaw (1992) (Correct)

Related documents from co-citation: More All
4: Enabling agents to work together (context) - Guha, Lenat - 1994
4: Semantic Memory (context) - Quillian - 1968
4: Open architecture multimedia documents (context) - Gaines, Shaw

CiteSeer
(NEC, Steve Lawrence, *et al*)

Scientific innovation through literature linkages (Don Swanson, Sandra Pratt, *et al*)



Arrowsmith
LitLinker
et al

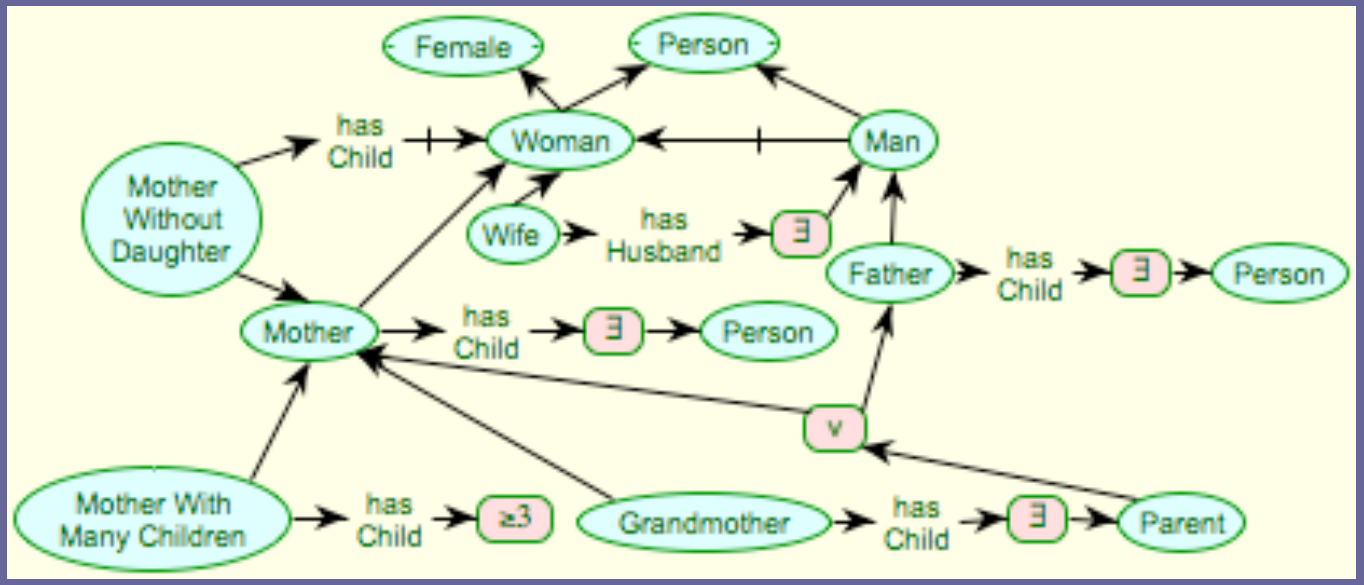
Suggest innovative conjectures through analysis of content & linkages in document space

Problems with scientific definitions

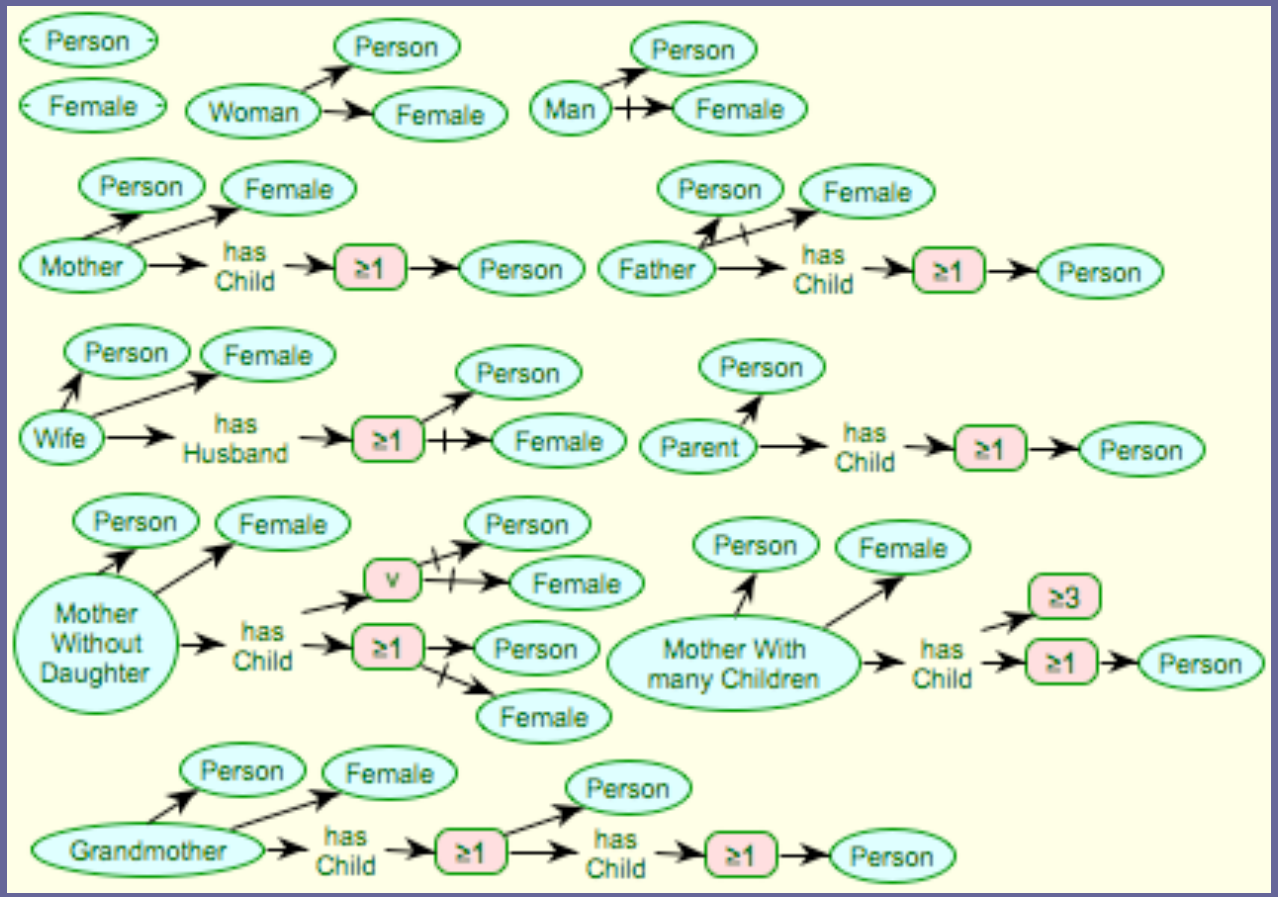
- Von Neumann's "proof" of no hidden variable models for quantum mechanics
- Popper's definition of an ordering over false theories such that one was better than another
- Arrow's proof that generally accepted 'fairness' conditions were mutually inconsistent

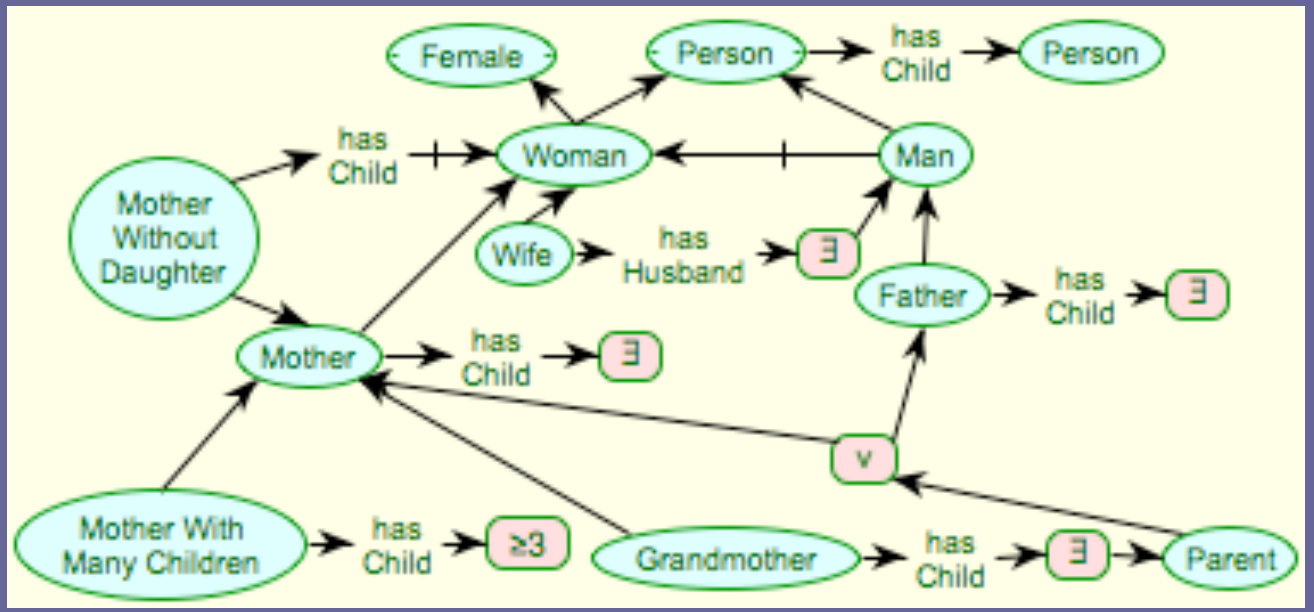
Use modern logical analysis tools to compute the availability of coherent models satisfying a given set of definitions

Definitions and their consequences



Woman	≡	Person ⊓ Female
Man	≡	Person ⊓ ¬Woman
Mother	≡	Woman ⊓ ∃ has_Child.Person
Father	≡	Man ⊓ ∃ has_Child.Person
Parent	≡	Father ⊔ Mother
Grandmother	≡	Mother ⊓ ∃ has_Child.Parent
Mother_With_Many_Children	≡	Mother ⊓ ≥3 has_Child
Mother Without Daughter	≡	Mother ⊓ ∀ has_Child.¬Woman
Wife	≡	Woman ⊓ ∃ has_Husband.Man





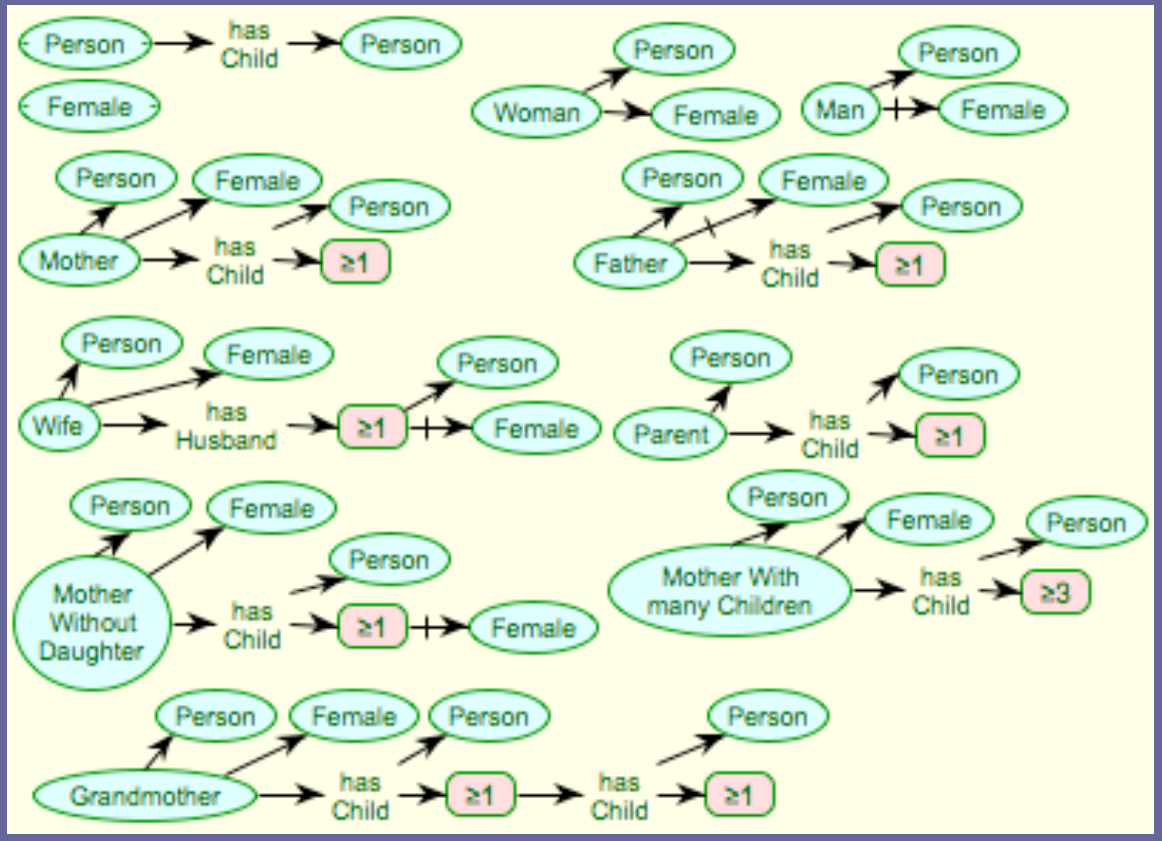
Improved definitions with intended consequences

Racer Server

Volker Haarslev
& Ralf Möller

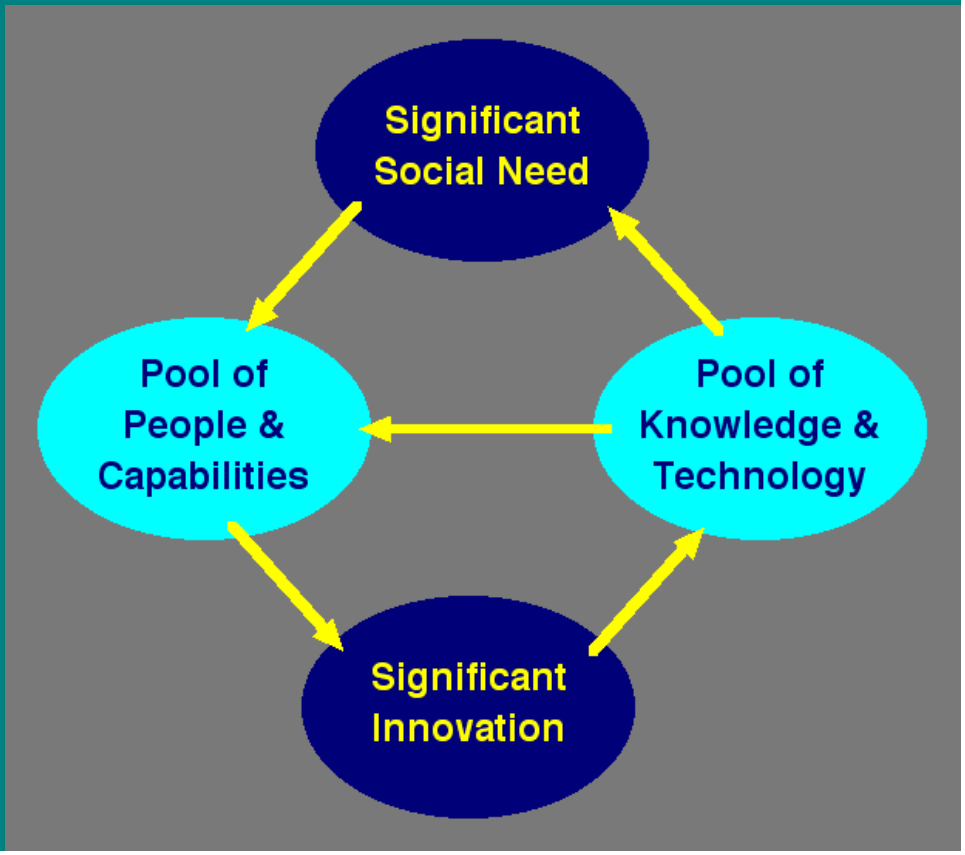
Adaptive logics

Diderik Batens
& Joke Meheus



Social Need and Critical Issues

- Access—the World Brain
- Overload—the historic constraint
- How much information?
- Addressing the social needs of the 21st century



World Brain, H G Wells 1938

Our contemporary encyclopaedias are still in
the coach-and-horses phase of development,
rather than in the phase of the automobile and the aeroplane.
Encyclopaedic enterprise
has not kept pace with material progress.
These observers realize that the modern facilities of
transport, radio, photographic reproduction and so forth
are rendering practicable
a much more fully succinct and accessible
assembly of facts and ideas than was ever possible before.

(H G Wells, World Brain, 1938)

Overload, Vannevar Bush 1945

There is a growing mountain of research.
But there is increased evidence that
we are being bogged down today
as our methods of transmitting and reviewing
the results of research are generations old
and by now are totally inadequate for their purpose
...publication has been extended
far beyond our present ability
to make real use of the record.

(Vannevar Bush, As we may think, 1945)

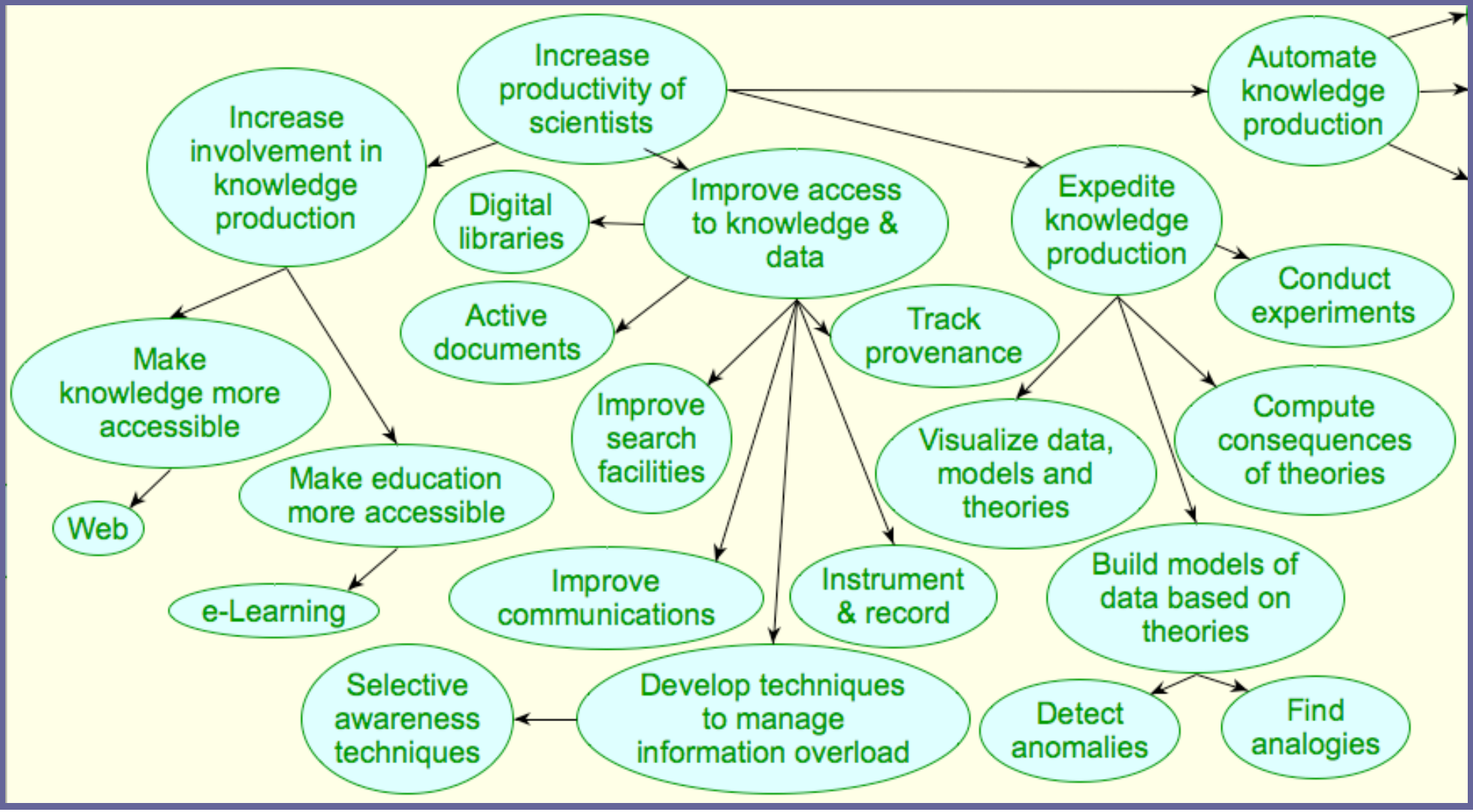
How much information? (Peter Lyman & Hal Varian, 2003)

- 5 exabytes of new stored information in 2002
- 92% stored on magnetic media
- Growth rate, 30% a year
- Web, 170 TB (surface), 92 PB (deep)
- Books, 8 TB/year
- Journals, 1.3 TB/year
- Library of Congress, 10 TB

Exabyte (EB) 10^{18}, Petabyte (PB) 10^{15}, Terabyte (TB) 10^{12}

Social needs of 21st century

- Sustain the exponential growth of knowledge despite the decline in growth of the scientific community
- Manage information overload



Challenges to the next generation

- Sustain the exponential growth of knowledge despite the decline in growth of the scientific community

— Increase involvement in knowledge production
— Improve access to knowledge and data
— Manage information overload (???)
— Expedite knowledge production
— Automate knowledge production (???)