prepared for WETICE96

Value-added Services on the WEB

June 1996

Gio Wiederhold
Stanford University

Abstract

To exploit existing and potential resources on the Web for effective engineering a number of services are needed.

We will indicate some of the opportunities and prerequisites for such services.

Collaboration, security, and payment schemes are some of the issues.

Many traditional relationships among consumers and vendors will change.

Reliable predictions are not possible, but exploring the range of choices is a challenge in itself.

Industry Needs Information

- Engineering and Manufacturing
  ✔own capability ↑ suppliers' capabilities
  ✔demand ↑ global demand
- Distribution and Transportation
  ✔costs for alternate means of shipping
- Finance
  ✔project demand ✓ project cost of funds
- Marketing and Service
  ↑taste and style ↑ demographics

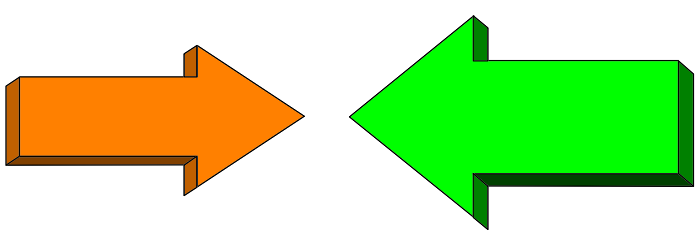
more from remote sources

Gio Wiederhold 1995 3

Information Leverage

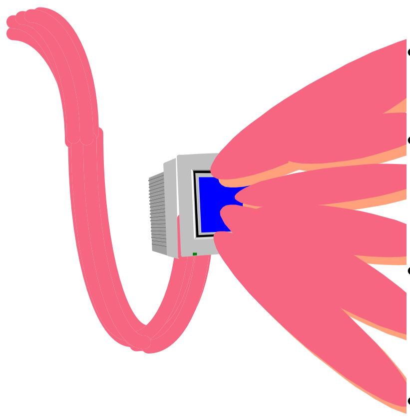
Tactical
- Customers
- Inventory
- Suppliers

Strategic
- Planning
- Capabilities
- Opportunities



Gio Wiederhold 1995 4

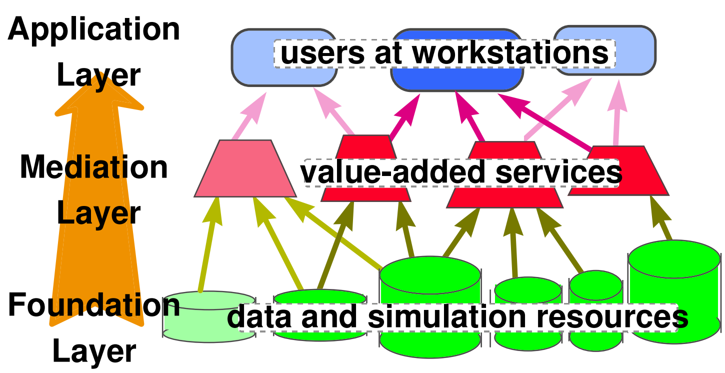
Information Data overload starvation



- More databases
  - public & corporate
- Faster communication
  - digital
  - packeting: TCP-IP, ATM
- World-wide connectivity
  - internet
  - world-wide web
- Disintermediation
  - ubiquitous publishing

Gio Wiederhold 1995 5

Transforming Data to Information



Gio Wiederhold 1995 6

Definition*

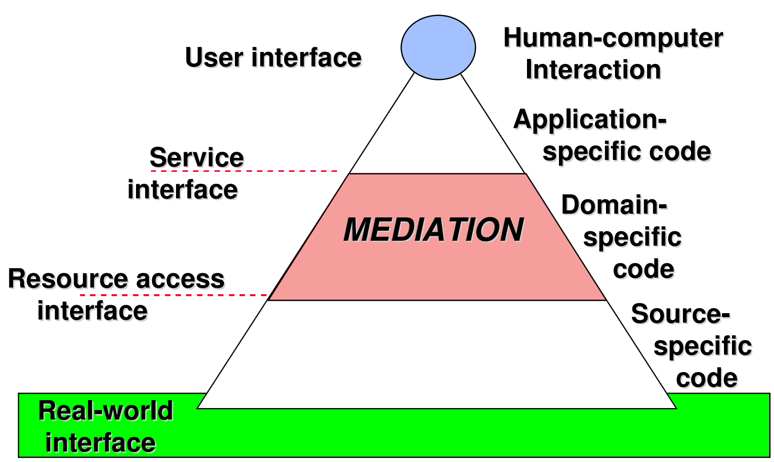
A *mediator* is a software module that exploits encoded knowledge about certain sets or subsets of data to create information for a higher layer of applications.

It should be small and simple, so that it can be maintained by one expert or, at most, a small and coherent group of experts.

* Wiederhold: IEEE Computer March 1992

Gio Wiederhold 1995 7

Functional Layer



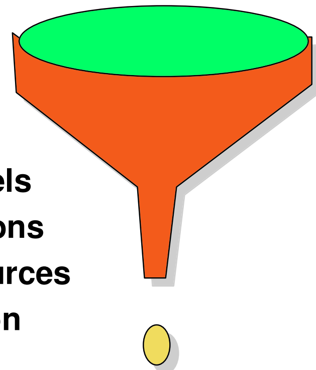
Gio Wiederhold 1995 8

Function of Mediation

Apply *Domain-specific* Specialist Knowledge to add value

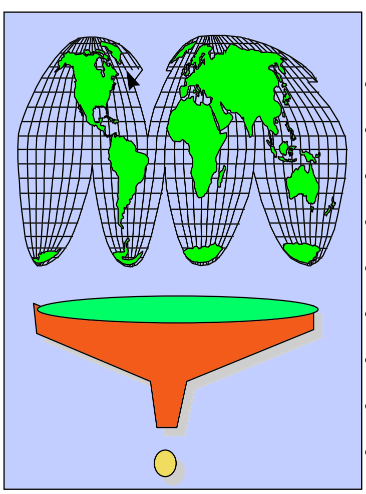
- to locate data sources
- to describe data for use
- to convert for consistency
- to abstract for insight / models
- to extrapolate to new situations
- to integrate from diverse sources
- to re-abstract for presentation

➜ ***INFORMATION***



Gio Wiederhold 1995 9

Mediation exploits *Knowledge*

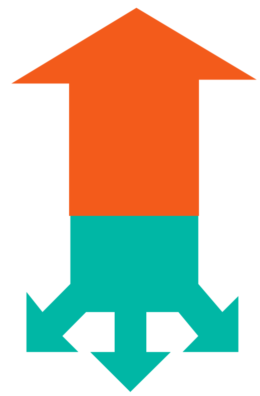


- *Discover* sources
- Search *likely files*
- Obtain *descriptions*
- Scan texts
- Select *relevant data*
- Abstract *to right level*
- Integrate from all
- Validate *consistency*
- Apply *to problem model*
- Test *stability* of extrapolations

Gio Wiederhold 1995 10

Making data relevant

- Data reduction
- Data abstraction
  – Summarization
  – Exception search
  – Level change to integrate with other data sources
- Follow Customer Model: hierarchical, divide-and-conquer, *a common paradigm*



Mediator Design Principle

Transform Data into Information

Match

Costumer Model

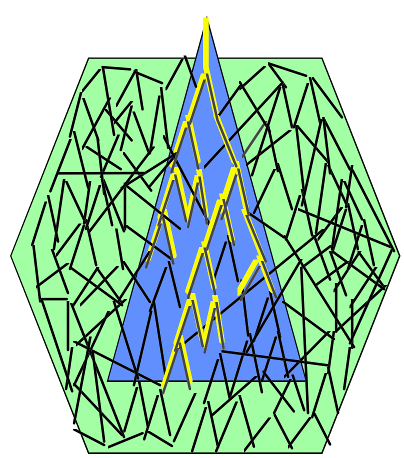
Hierarchical

to

Resource Model

General network

(and maintain models)



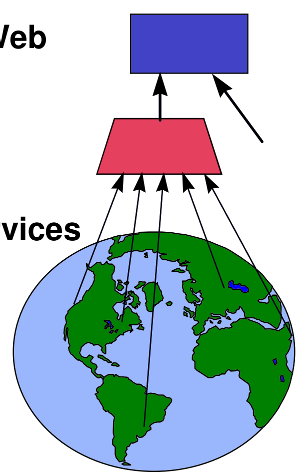
Access Maintenance Tasks

- Selection of relevant source material
  - using Yahoo, Knowbots, Harvest, federated schemas, GLOSS
  - evaluate descriptions. meta-data
- Focused access to the variety of resources
  - using SQL, wrappers, CORBA, . . .
- Caching
  - to resolve asynchrony in sources
  - create consistent histories
- Tracking Resources, their cost, and response

Gio Wiederhold 1995 13

Mediation on the WWW

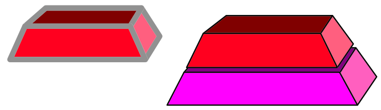
- Resources on the World-Wide-Web
  - are plentiful
  - autonomous
  - incoherent
- Opportunity for value-added services
  - select best source
  - improve coverage
  - minimize overlap
  - resolve inconsistencies
  - summarize results

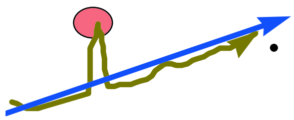


Gio Wiederhold 1995 14

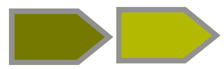
Abstraction / Summarization

- Abstraction to match levels of granularity





- Seeking exceptions from expected values or trends
- Assessment of quality of diverse sources

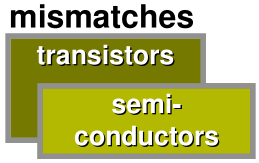


- Omission of replicated or known information

Gio Wiederhold 1995 15

Integration

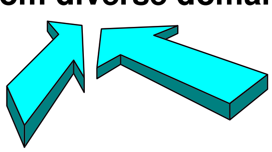
- Resolution of scope mismatches



- Ranking of material from diverse sources

1. 45
2. 43
3. 33
4. 28
.. . .

- Integration of material from diverse domains



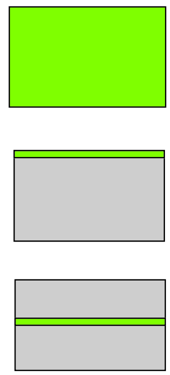
Gio Wiederhold 1995 16

Result modes for ranking

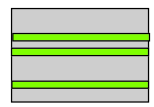
Databases:
- Completeness
- All the answers

Prolog
- Correctness
- The first answer

Optimization
- The best one
- Assumes all factors are known, no human decision



Customer:
- wants choices



- explanation
- background

Gio Wiederhold 1995 17

Ranking

Qualitative Significant Differences:
in terms of the customer model

Plan 1. UA59 dep.Wash.Dulles 17:10, arr. LAX 19:49
Plan 2. AA75 dep.Wash.Dulles 18:00, arr. LAX 20:10
Plan 3. UA119 dep.Wash.Dulles 9:25, arr. LAX 12:00

Busy Joe:
P1= P2, P3

Speedy Mike:
P2, P1=P3

Greedy Pete:
P1=P3, P2

Gio Wiederhold 1995 18

Opportunities in Engineering

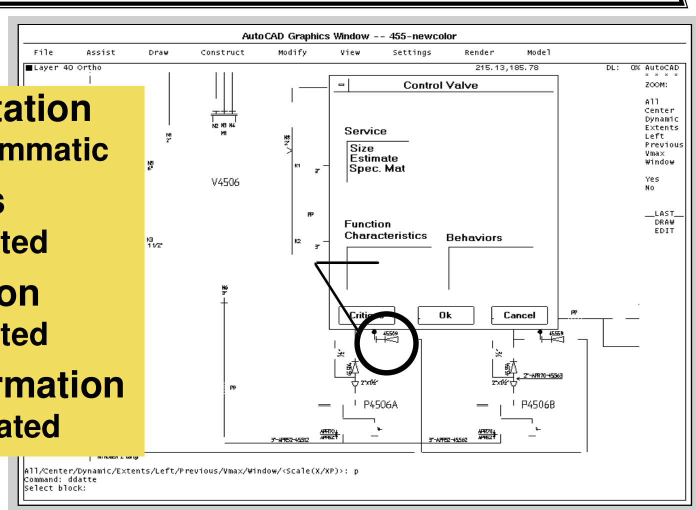
- **On-line presentation services,** appropriate for audience
- **Integration of documents and figures** for WWW access
- **Abstraction services:** summaries of papers, reports (with references to base mat.)
- **Review services** over suppliers, technologies, services
- **Alternative ranking** of suppliers, parts, materials, . . .
- **Active documents** with function evaluation, plotting
- **Test generators and checkers** (people, equipment)

Gio Wiederhold 1995 19

From Andrew Arnold: Civ. Eng. Qualification Exam

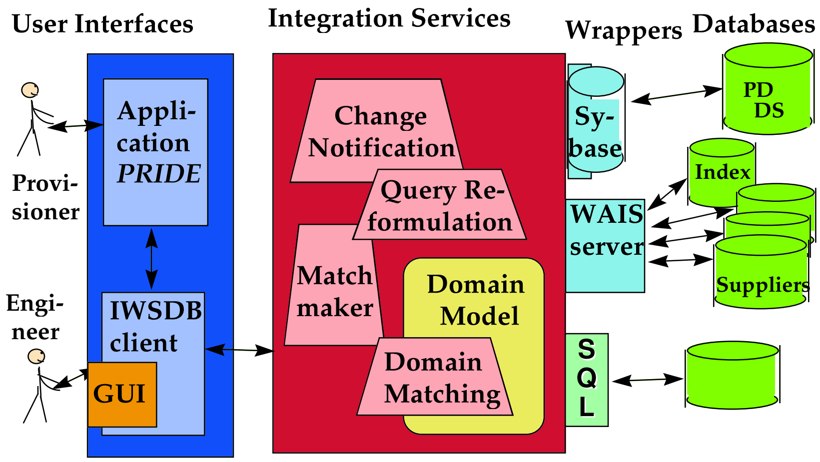
Control Valve Sizing, *Future*

- **Interpretation**
  - Programmatic
- **Analysis**
  - Integrated
- **Evaluation**
  - Integrated
- **Transformation**
  - Automated



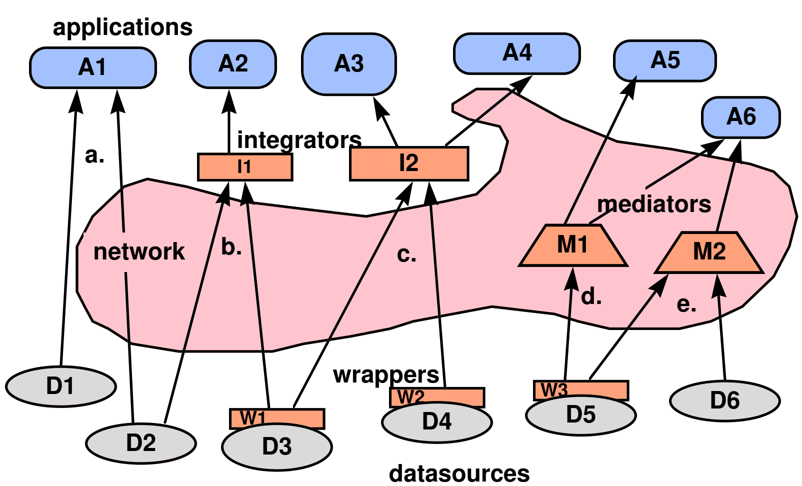
Gio Wiederhold 1995 20

F-22 IWSDB Phase 6



Gio Wiederhold 1995 21

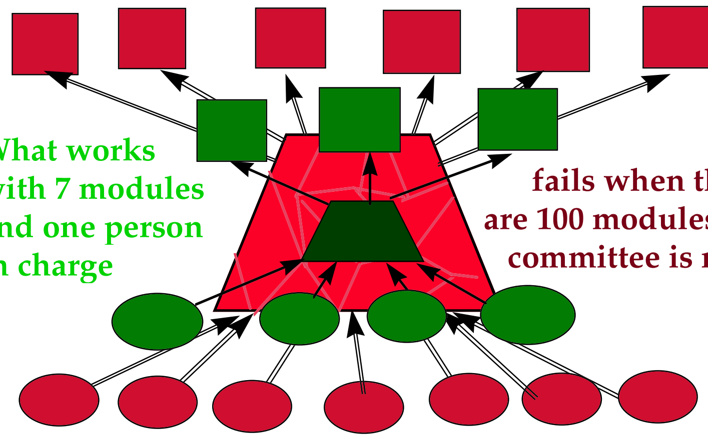
Evolution of mediation



Gio Wiederhold 1995 22

Central Solutions do not Scale

What works
with 7 modules
and one person
in charge

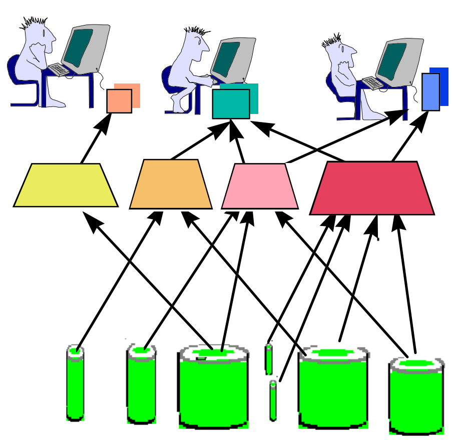


fails when there
are 100 modules and a
committee is needed

Any changes in resources affect the central module

Gio Wiederhold 1995 23

Domain-specific Mediation



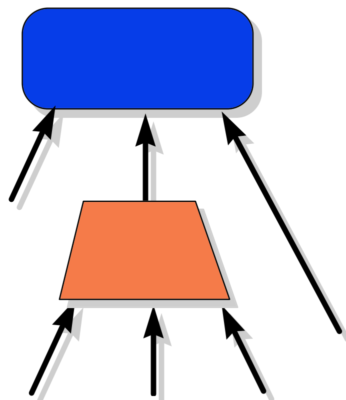
- User application
  - Workstations
- Mediator
  - Expert-owned nodes
- Data sources
  - Remote primary and byproduct services

Gio Wiederhold 1995 24

Integration at two levels

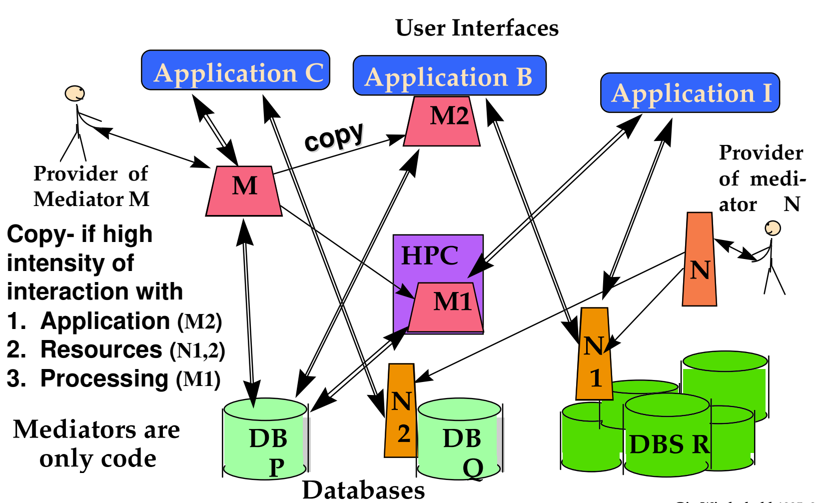
Application
- Informal, pragmatic
- User-control

Mediation
- Formal service
- Domain-Expert control



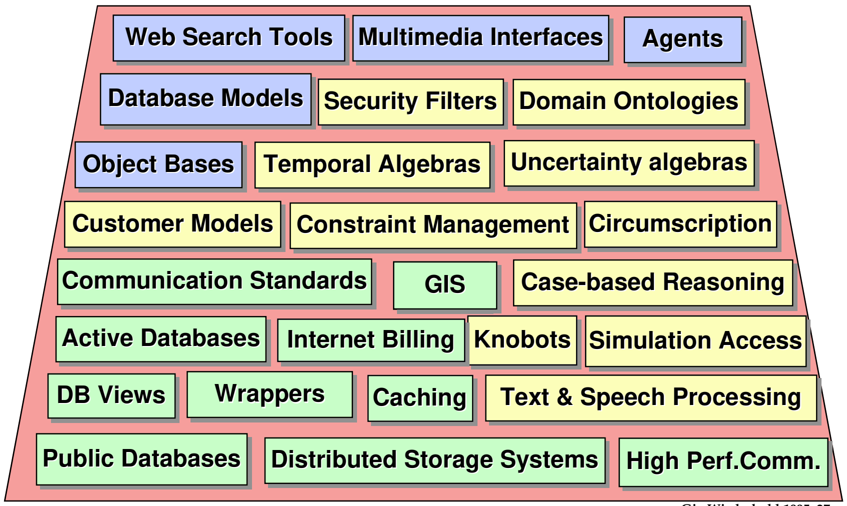
Gio Wiederhold 1995 25

Allocation Flexibility



Gio Wiederhold 1995 26

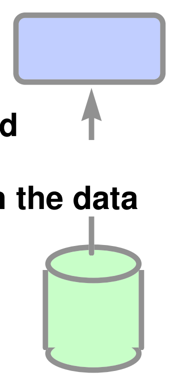
Getting there: Available Technology/Science



Gio Wiederhold 1995 27

Current Technologies

- SQL
  - One Verb - SELECT with primitive aggregation
  - One Database at a time
  - One Datatype: Tables
- Object-orientation
  - Group data into objects = predefined aggregation
  - Program snippets -- methods -- with the data
- Midleware (ex.: CORBA)
  - Fetch objects from server
  - Assume coherent domains

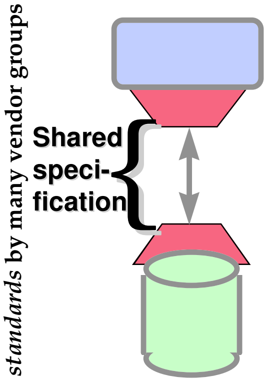


Gio Wiederhold 1995 28

Middleware

CORBA (Common Object Request Broker)
  – *IBM* SOM, DSOM
- DOE (Distributed Objects Everywhere)
  – SunSoft
- DOME
- EZ-bridge
  – System Strategies inc.
- ILU (InterLanguage Unification) *Xerox*
- ISIS
- KQML (Knowledge Query & Manipulation Lang.)
- MQM (Message Queing Middleware)
  – IBM (for mainframe connections)
- OLE (Microsoft: Object embedding and Linking)
- OpenDOC (*Apple*)
- PDES (Product Data Interchange using STEP)
- TIB (Teknekron Information Bus)

Many *standards* by many vendor groups



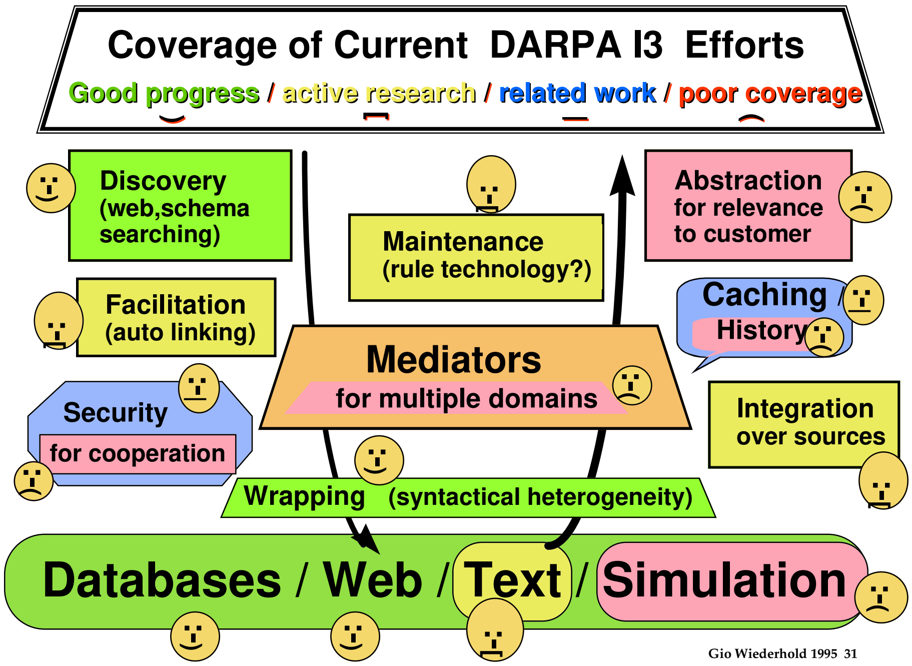
Gio Wiederhold 1995 29

Status of Mediation Technology

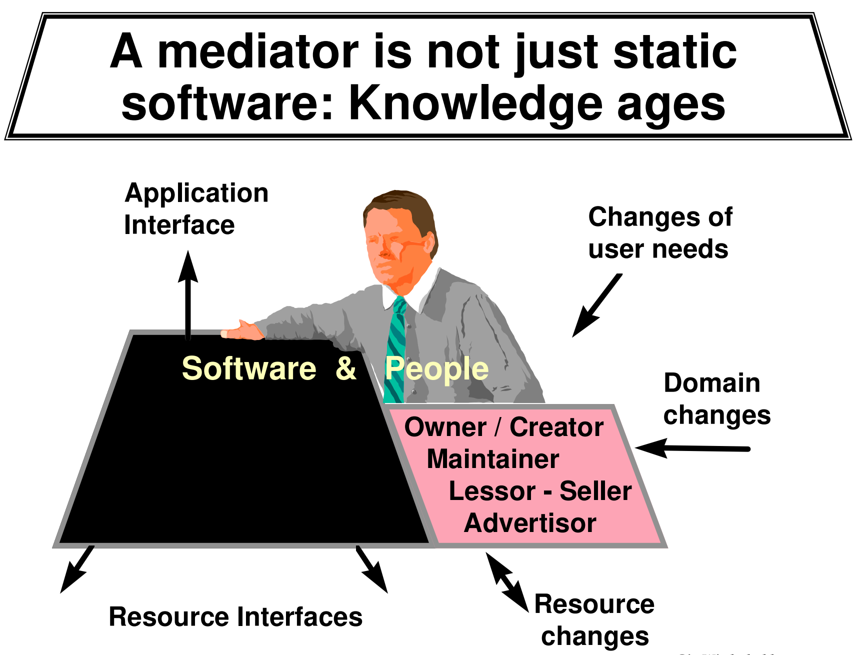
Today
- Handcrafted
- Expert consults with programmer
- Programmer codes the knowledge needed
- Resource changes require advise, program update

Future
- Generated from models
- Domain Expert maintains models
- Specification determines functions
- Resource changes trigger regeneration

Gio Wiederhold 1995 30

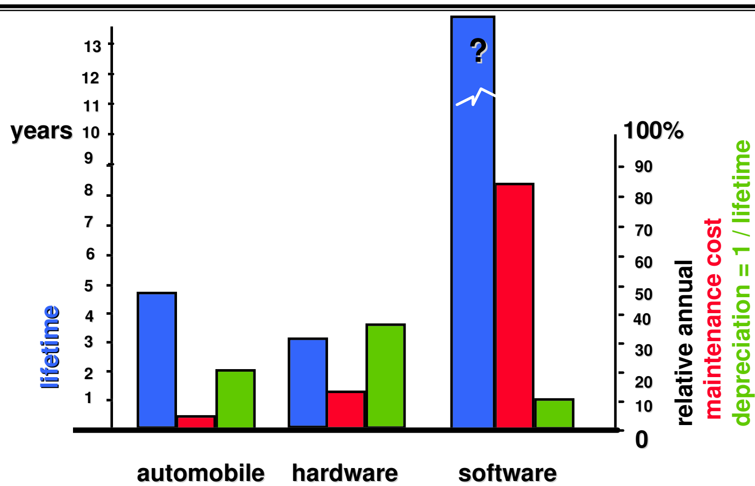


Coverage of Current DARPA I3 Efforts
Good progress / active research / related work / poor coverage
Discovery (web,schema searching)
Maintenance (rule technology?)
Abstraction for relevance to customer
Facilitation (auto linking)
Caching / History
Mediators
for multiple domains
Security
for cooperation
Integration over sources
Wrapping (syntactical heterogeneity)
Databases / Web / Text / Simulation
Gio Wiederhold 1995 31



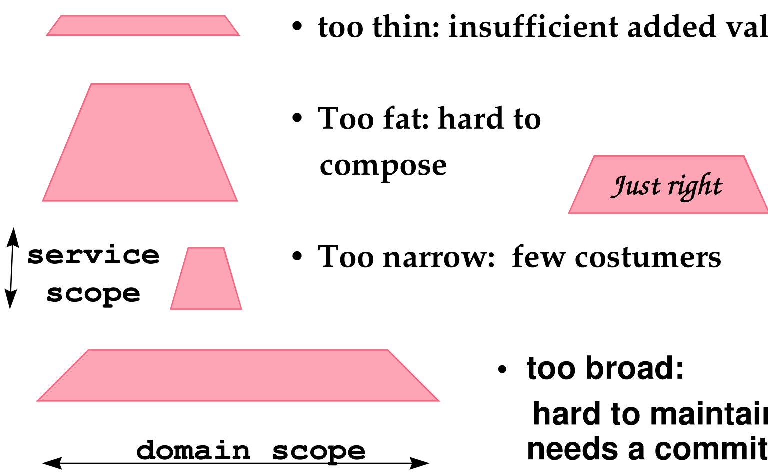
A mediator is not just static software: Knowledge ages
Application Interface
Changes of user needs
Software & People
Domain changes
Owner / Creator
Maintainer
Lessor - Seller
Advertisor
Resource Interfaces
Resource changes
Gio Wiederhold 1995 32

Maintenance is good for you



Gio Wiederhold 1995 33

Fat versus thin mediators



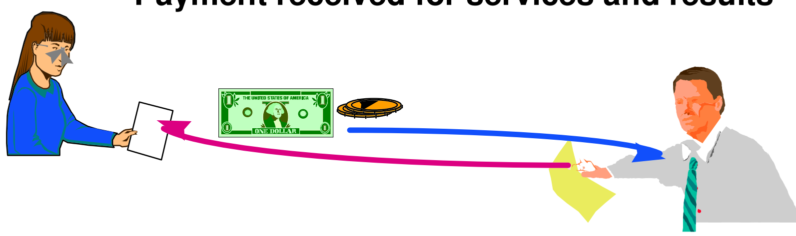
- too thin: insufficient added value
- Too fat: hard to compose
- Too narrow: few costumers
- too broad:
  hard to maintain,
  needs a committee

Gio Wiederhold 1995 34

Mediation as a Service

Service Paradigm

- **Customer views understood within server domain**
- **Processes use stored and maintained knowledge**
- **Processing adds value to data objects accessed**
- **Payment received for services and results**



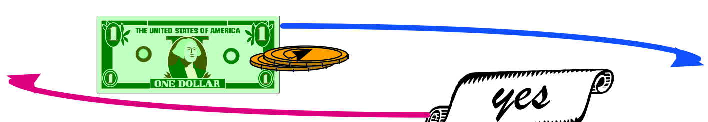
Gio Wiederhold 1995 35

E-money

Services must be paid for

- **Incentive for creation and improvement**
- **price proportional to value added**
- **profit *f* (cost, market, price, overhead)**
- **Price low per item, so overhead must be low**

Simple payment (no credit accounts, checks)



Enabled through secure signatures

Gio Wiederhold 1995 36

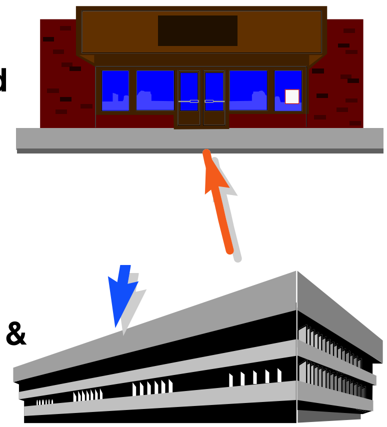
CommerceNet Vision

Purchase components
- initially electronics and electro-mecanical

Rapid, economical building of equipment

Get good choice
- competent supplier
- resolve variety of term & classifications

Use e-mail, e-money



Gio Wiederhold 1995 37

Simulation services

1. **Continously executing**: weather prediction
   – SimQL result reports best match samples
2. **Execution specific to query**: what-if assessment, spreadsheets
   – may require HPC power for adequate response
3. **Complement base data**: materials data, assembly
   – performs inter- or extra-polations to match query parameters
4. **Combinations of 2. and 3.**: top layer simulation using stored partial lower level results: weapon performance in setting
5. **Human-in-the-loop** (mediated by an agent program): SAFs

Note
- A simulation service program can be written in any language
- A simulation service must be compliant to the interface

Gio Wiederhold 1995 38

Domain Specialization

- Knowledge Acquisition &
- Knowledge Maintenance

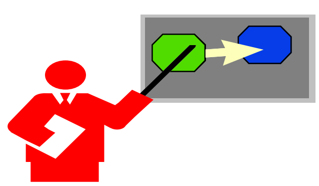
require

- Domain specialists
- Professional organizations

Empowerment

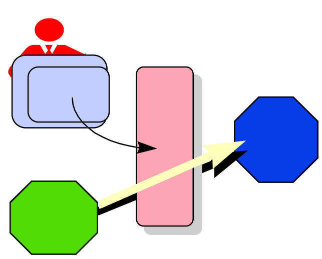
Gio Wiederhold 1995 39

New Role for Consultants



Old

- Used at Design Time
  and
- To Explain Failures



Future

- Available as a Service
- Responsible for Knowledge Maintenance

Gio Wiederhold 1995 40

Industrial Needs Served

NEEDS

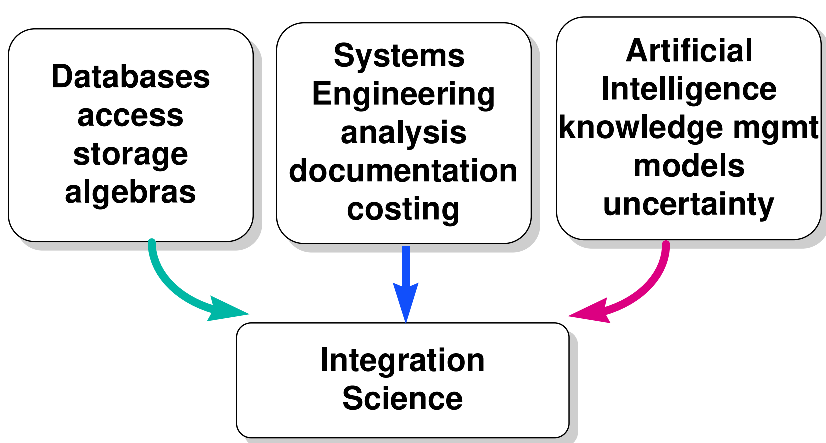
- Access to relevant Information
- Rapid response to changing situations
- Remain current with global conditions
- External services can be shared effectively

FEATURES

- Linkages to networks and resources
- Incremental update of information systems avoids legacy problem
- Equal access to local and remote sources
- Value-added services live *in* the network

Gio Wiederhold 1995 41

Integration Science



Gio Wiederhold 1995 42