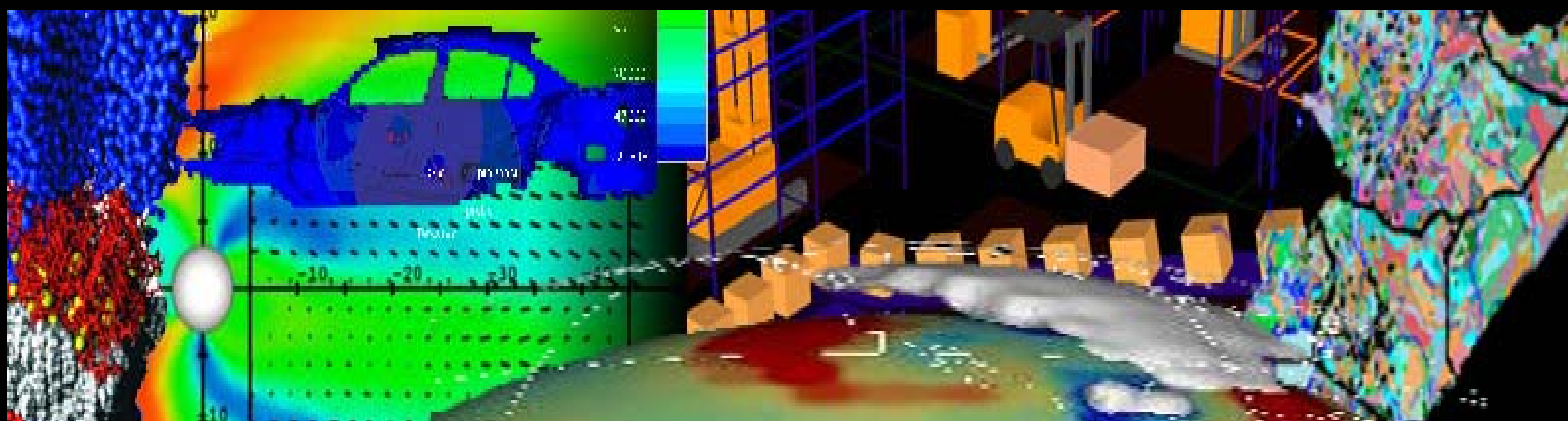
DATA CENTER

Making sense of your data™

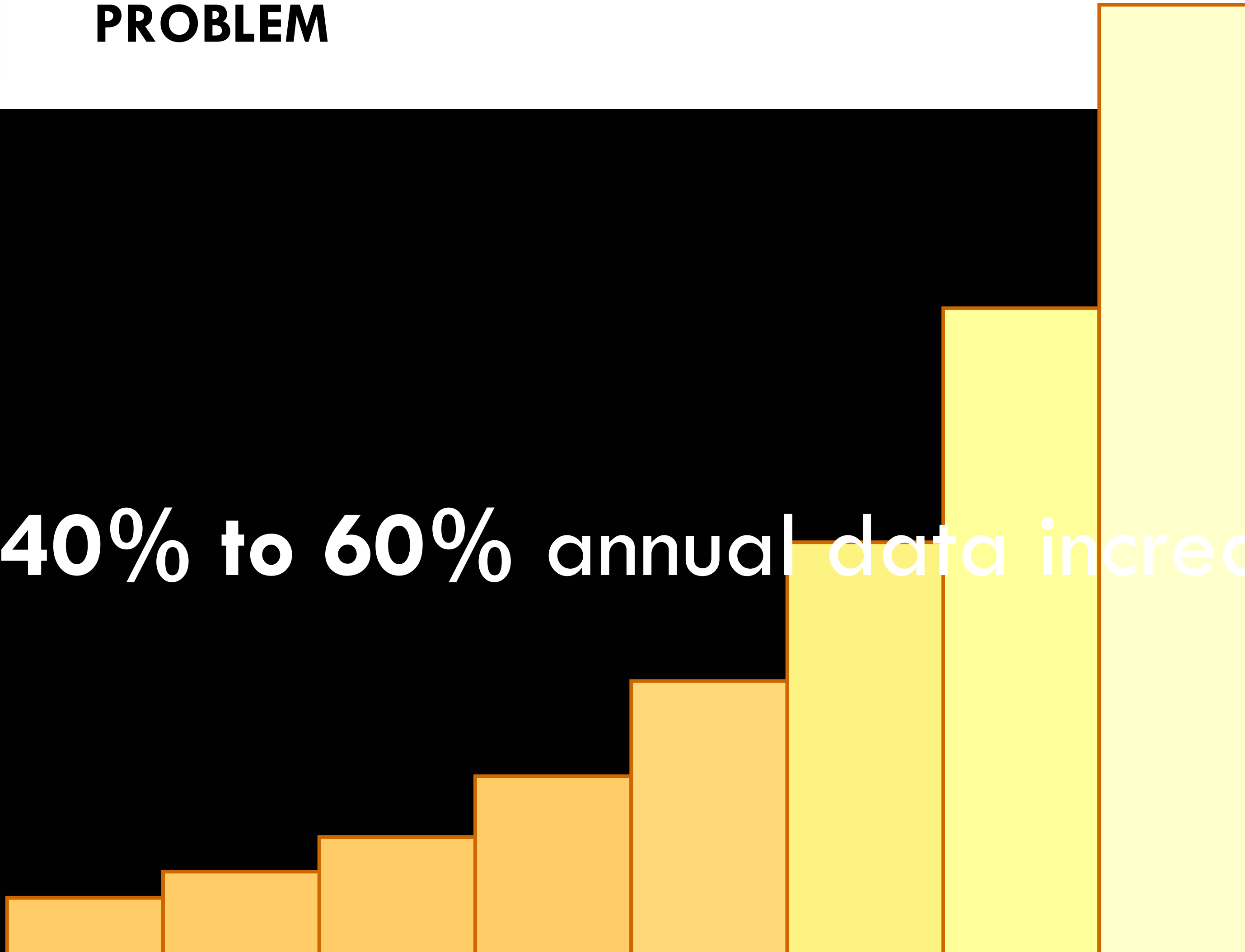


David L. Brock, Founder and Director

Edmund W. Schuster, Research Affiliate

Massachusetts Institute of Technology

PROBLEM



40% to 60% annual data increase

PROBLEM

"companies are struggling to figure out how to turn all those bits and bytes from a liability into a competitive advantage."

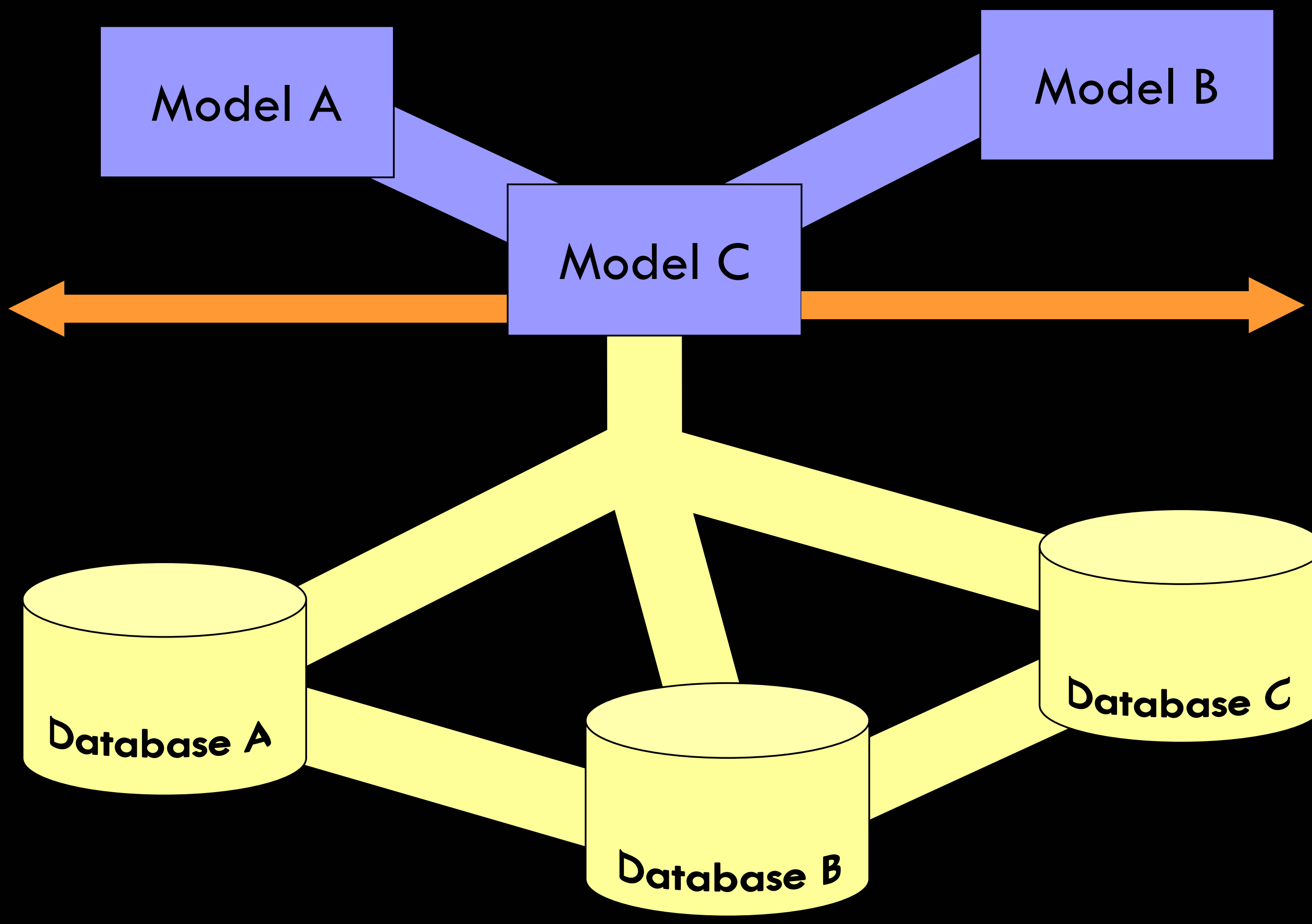
Park, Andrew (2004), "Can EMC Find Growth Beyond Hardware?," *BusinessWeek*, November 1, 2004.

DATA CENTER

PROBLEM

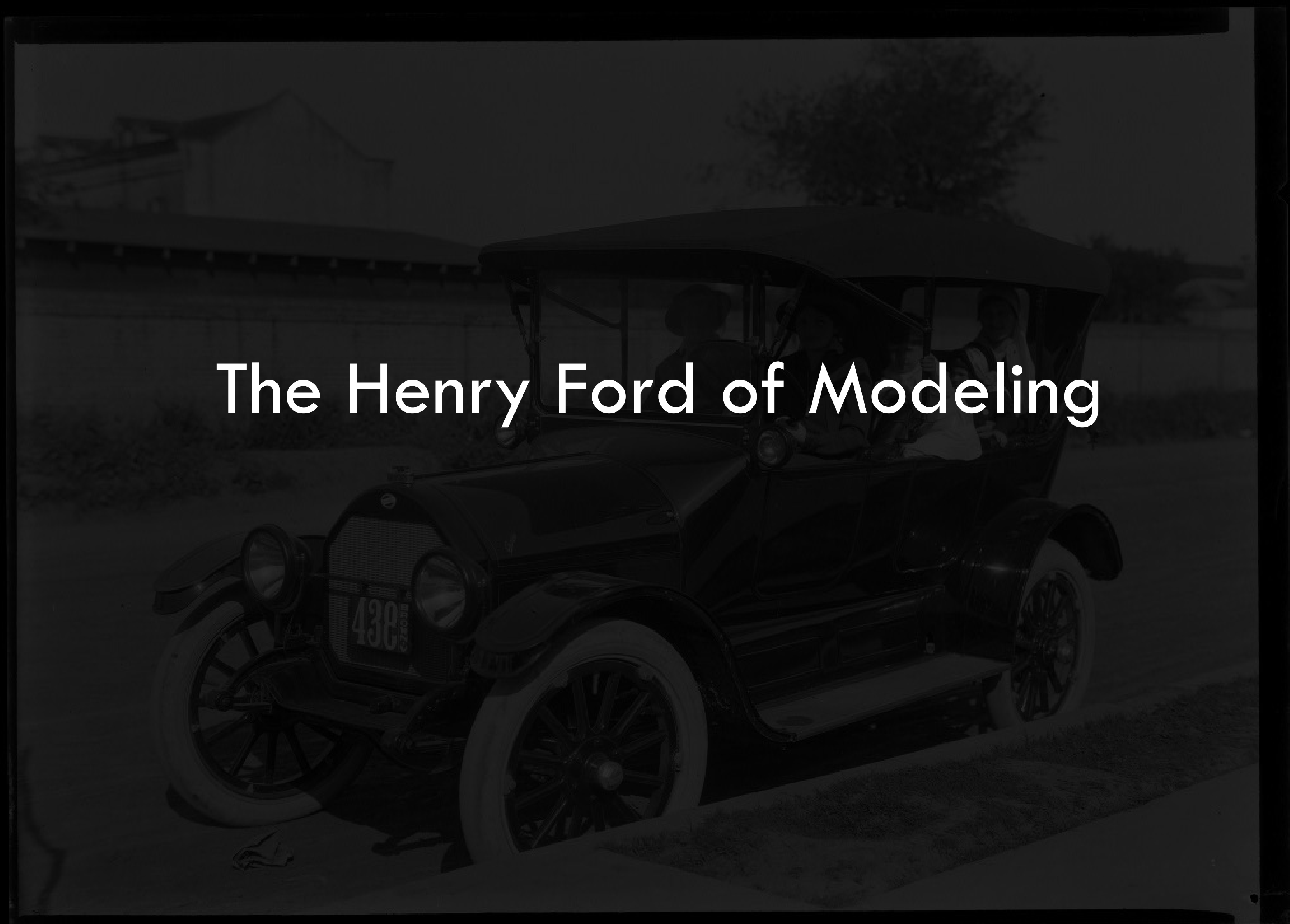
What are you going to do
with all your
Data?

VISION



DATA CENTER

VISION



The Henry Ford of Modeling

APPLICATIONS

- Manage
- Synchronize
- Distributed
- Store
- Analyze
- Predict
- Plan

DATA CENTER

INTEGRATION

Amazon has recently announced "A9" a new tool that can accomplish searches of information located on HTML web pages in addition to the text of thousands of books.

Hof, Robert, D. (2004). "Amazon Joins the Search Party." *BusinessWeek*, September 27.

INTEGRATION (CONTINUED)

In the US economy, there are billions of embedded microcontrollers in cars, traffic lights, and air conditioners that give specialized instructions for control based on sensing specific aspects of the environment.

All of these microcontrollers act in total isolation from one another.

Corcoran, Elizabeth (2004). "Giving Voice to a Billion Things." *Forbes,* September 5.

MOVE TO INTEGRATION (CONTINUED)

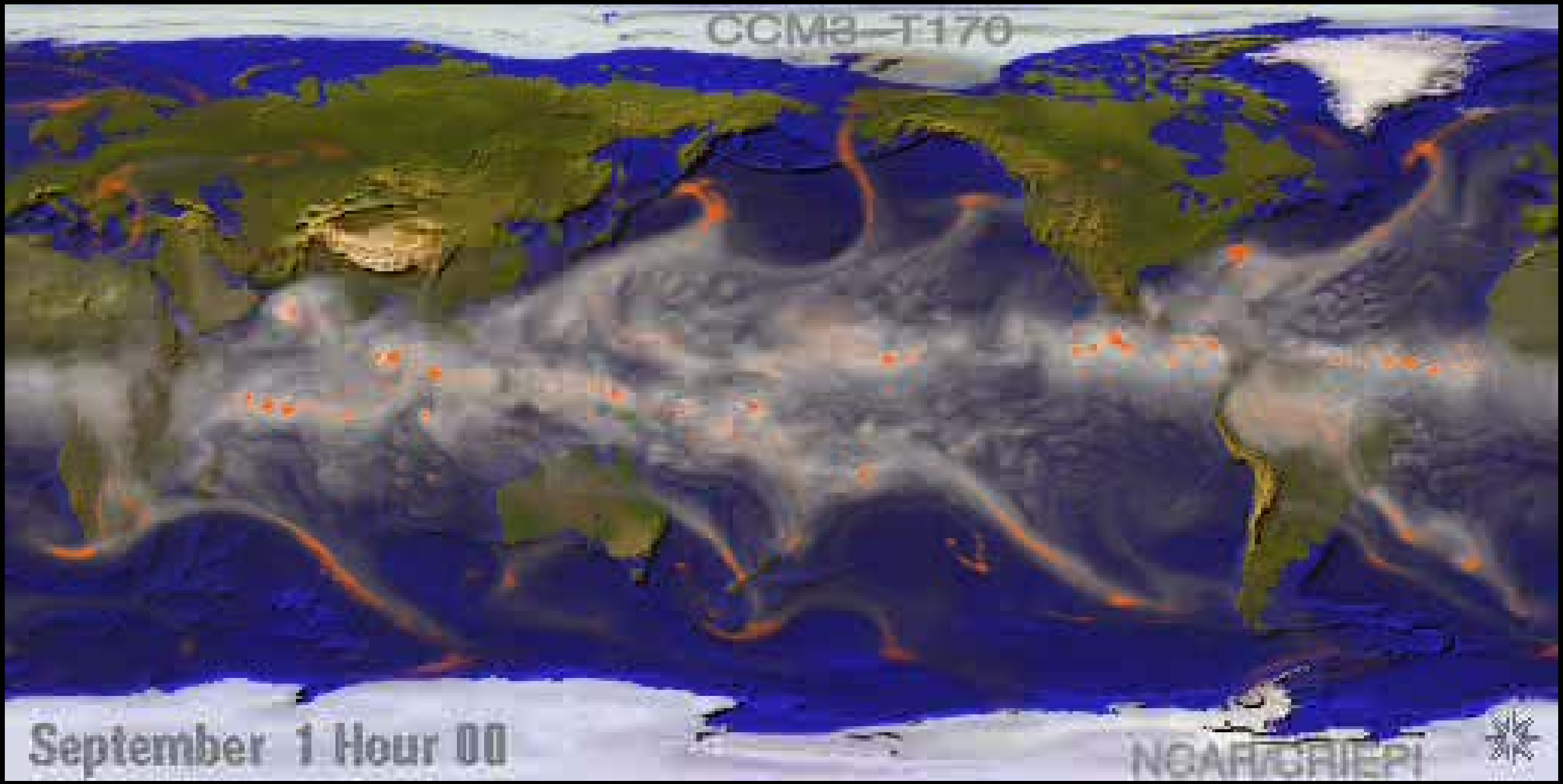
Ember, a company located in Cambridge, MA, has developed a "mesh network" that holds the potential of allowing all of these microcontrollers to communicate with each other.

Corcoran, Elizabeth (2004). "Giving Voice to a Billion Things." *Forbes*, September 5.

DATA CENTER

MODELS

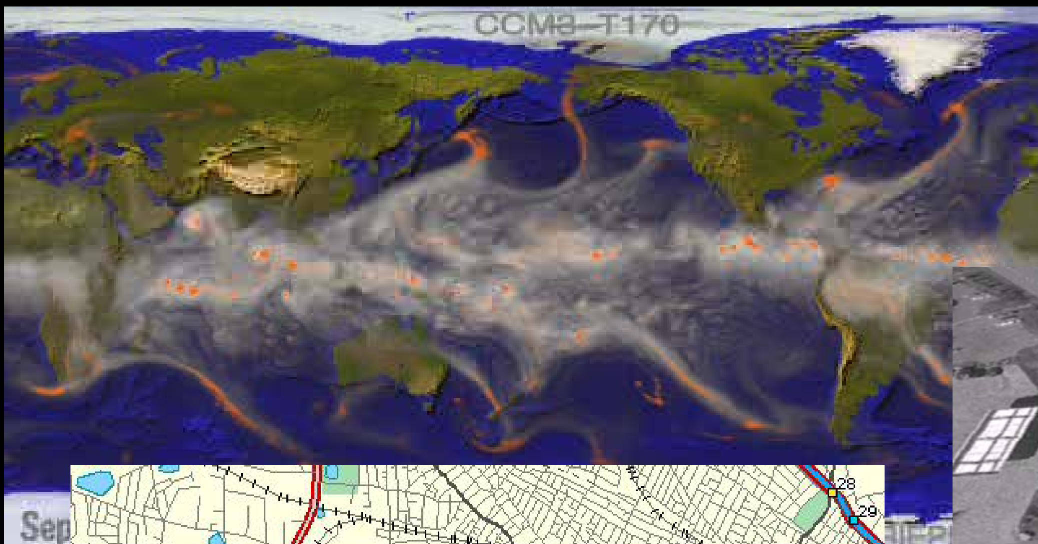
MODEL - WEATHER

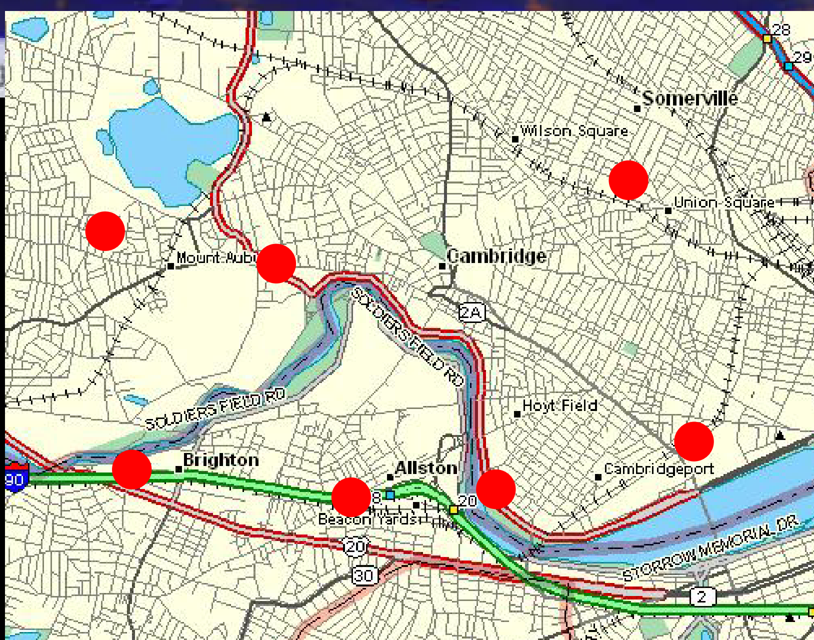


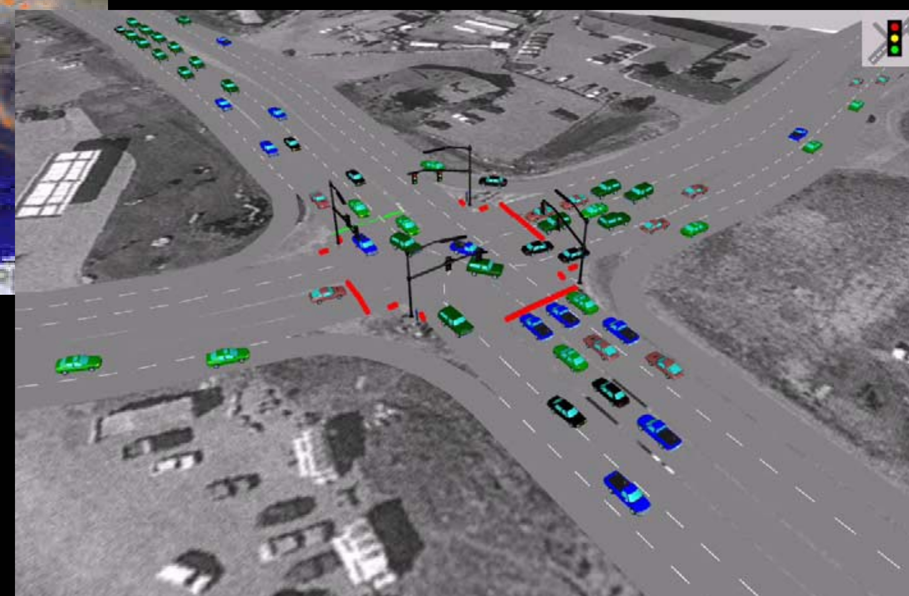
DATA CENTER

MODEL – TRAFFIC FLOW

EXAMPLE – INTEGRATE MODELS FOR LOGISTICS

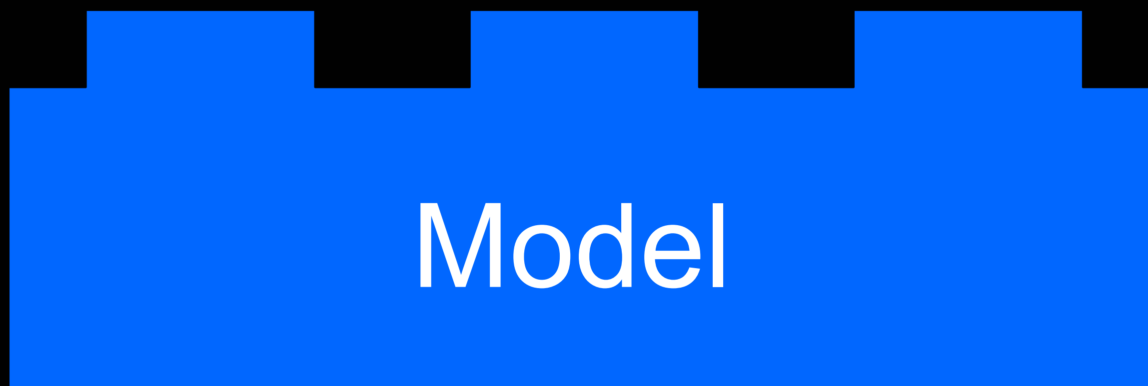




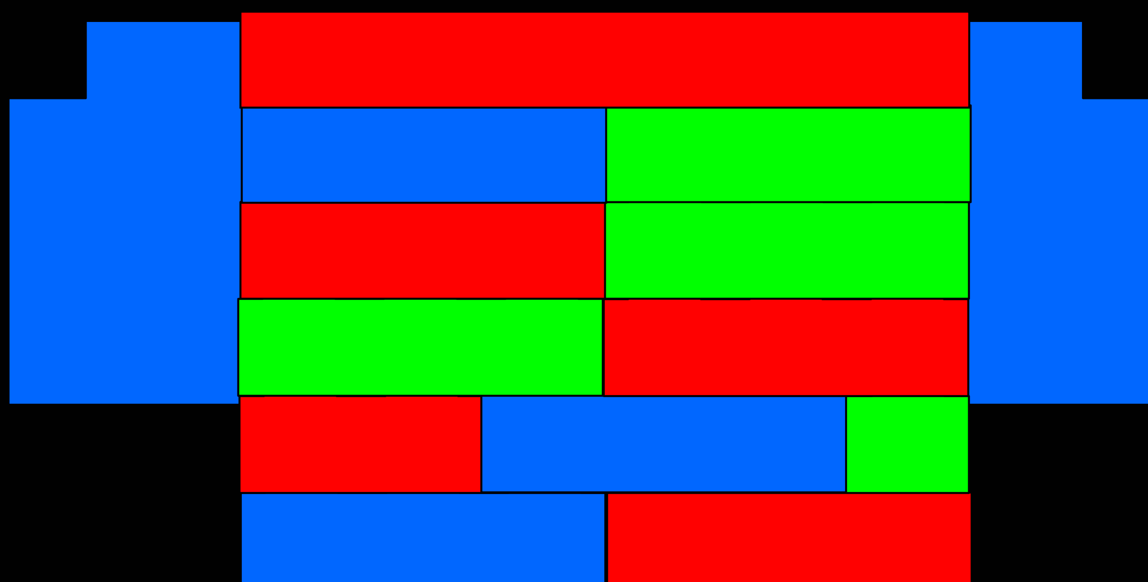


PROBLEM

MODELS

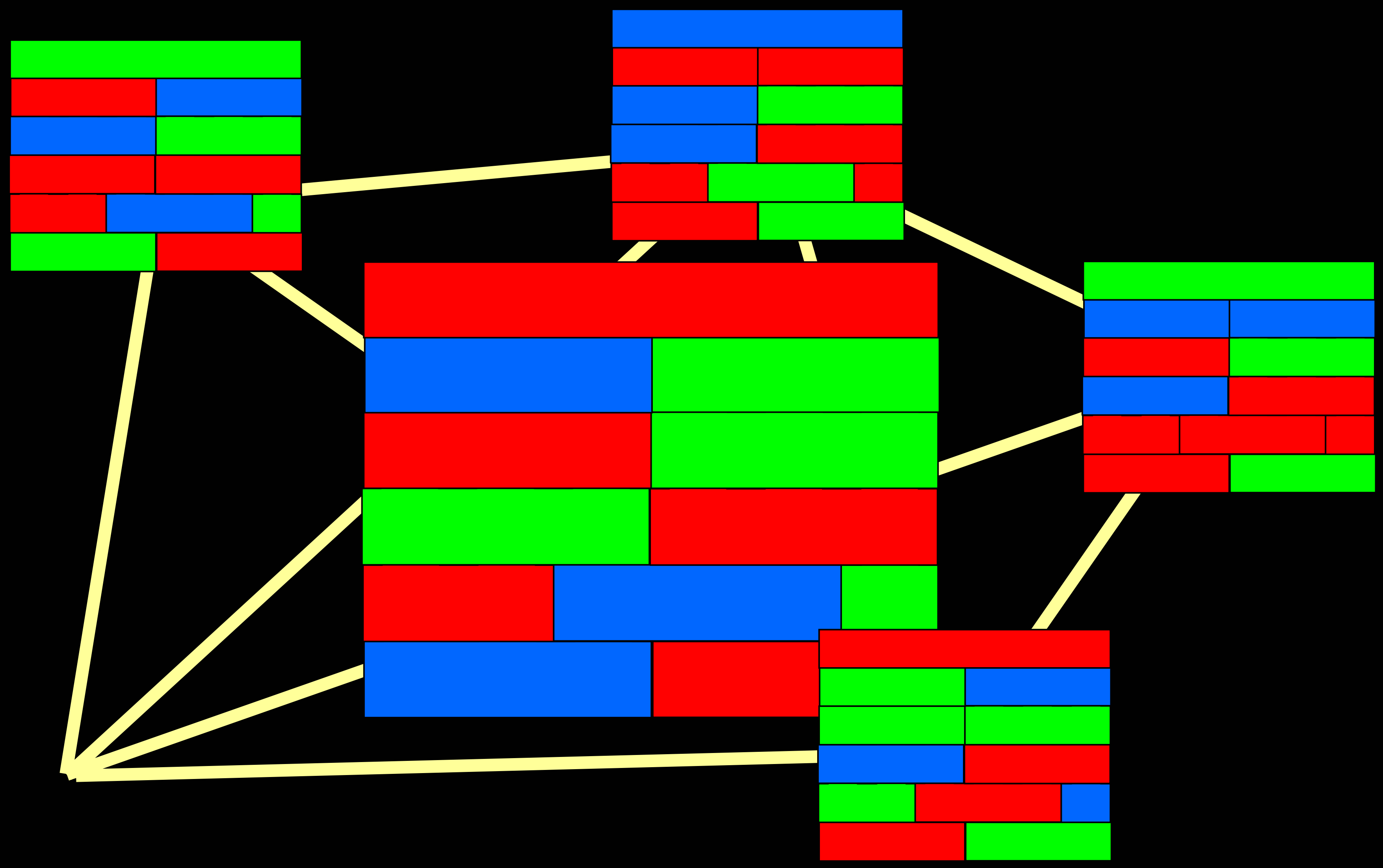


MODELS

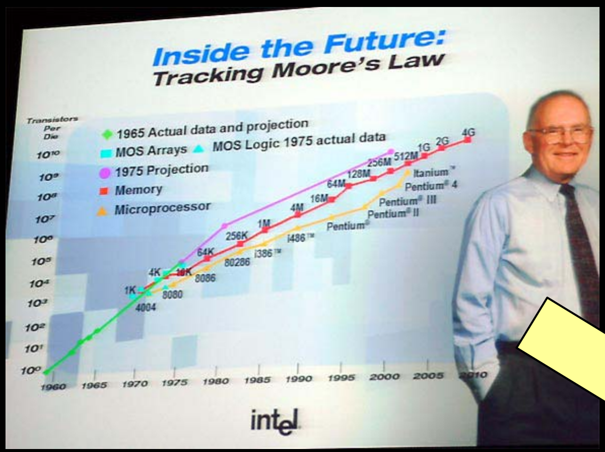


DATA CENTER

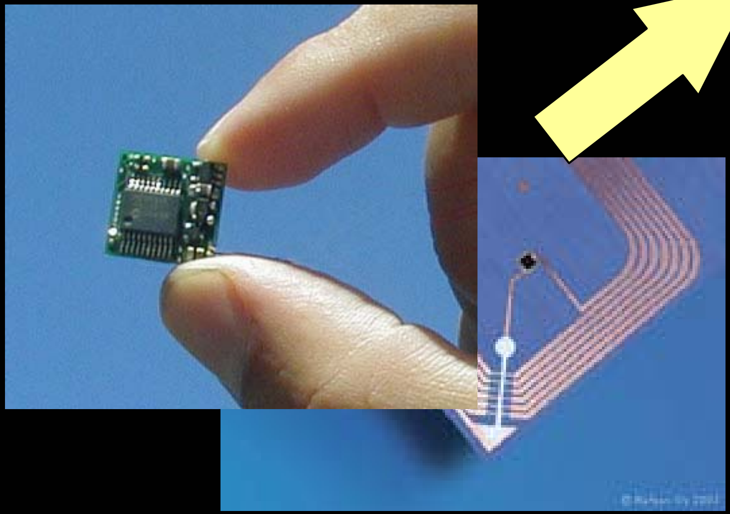
MODELS

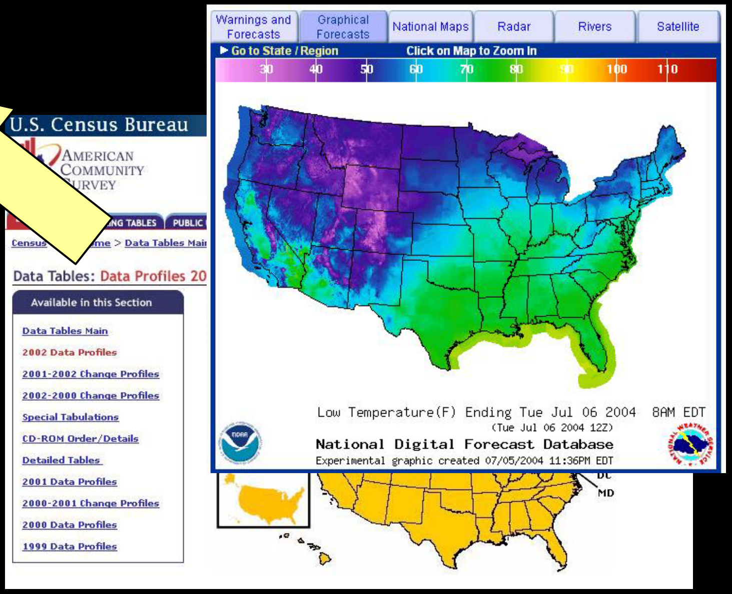


DRIVERS



HTML
XML
EPC
TCP/IP
HTTP
SOAP

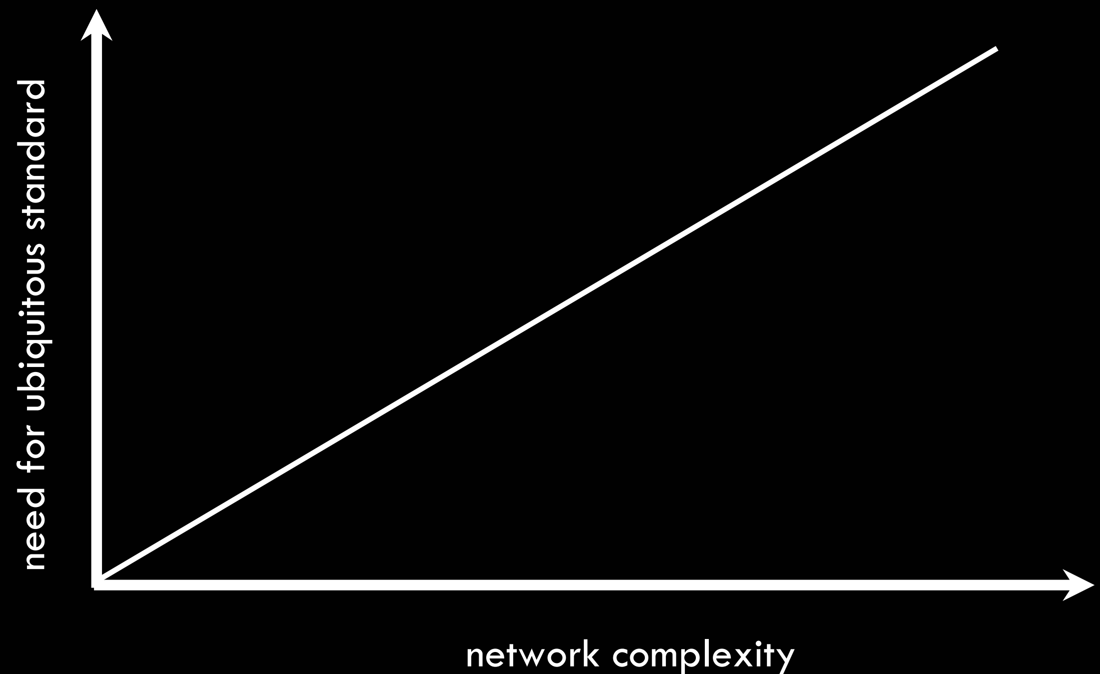




VISION

Create an infrastructure for
wide spread, general purpose
Interoperable data and modeling.

STANDARDS



The more complex the network, the more you need standards

DATA CENTER

VISION

Mission

- Make sense of your data

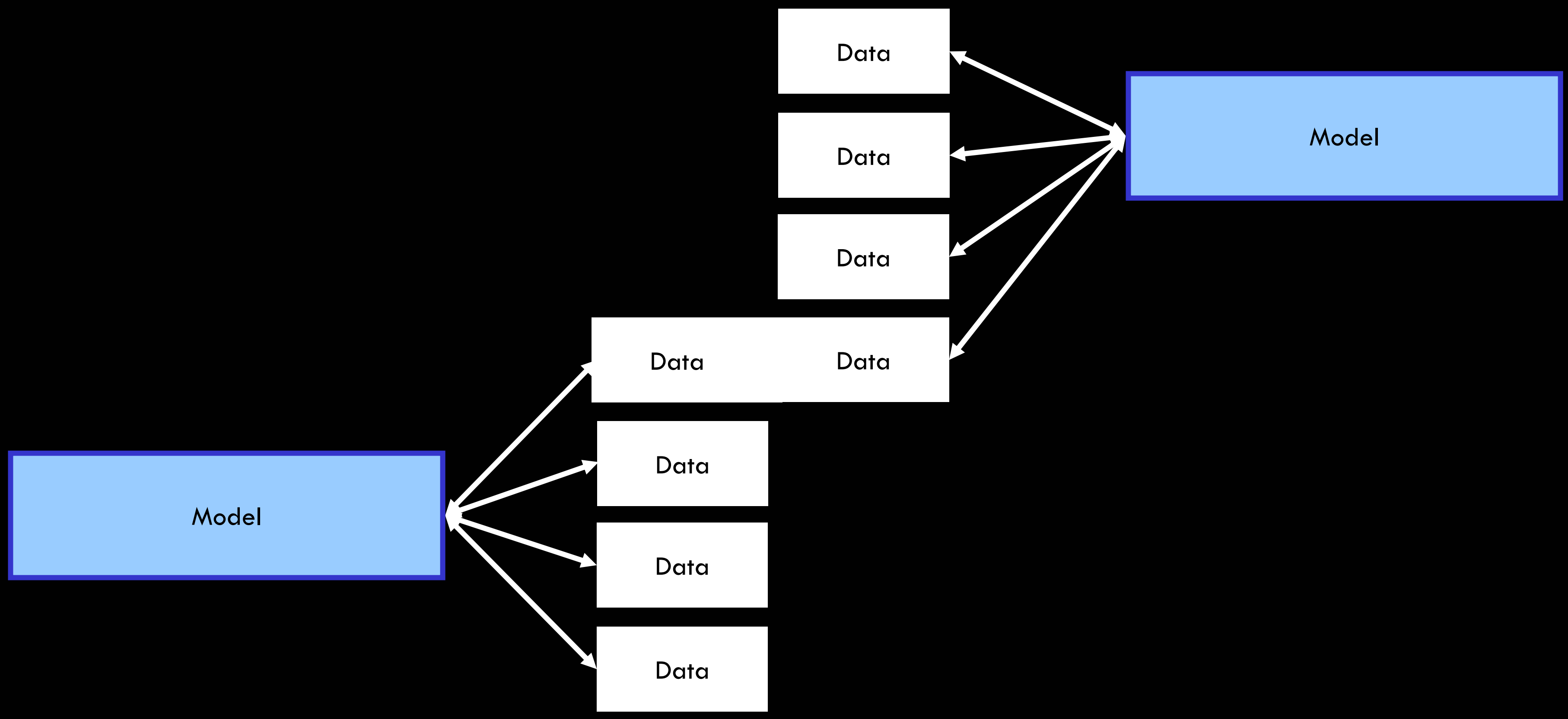
Task

- Create the standards and systems for interoperable data and modeling

M

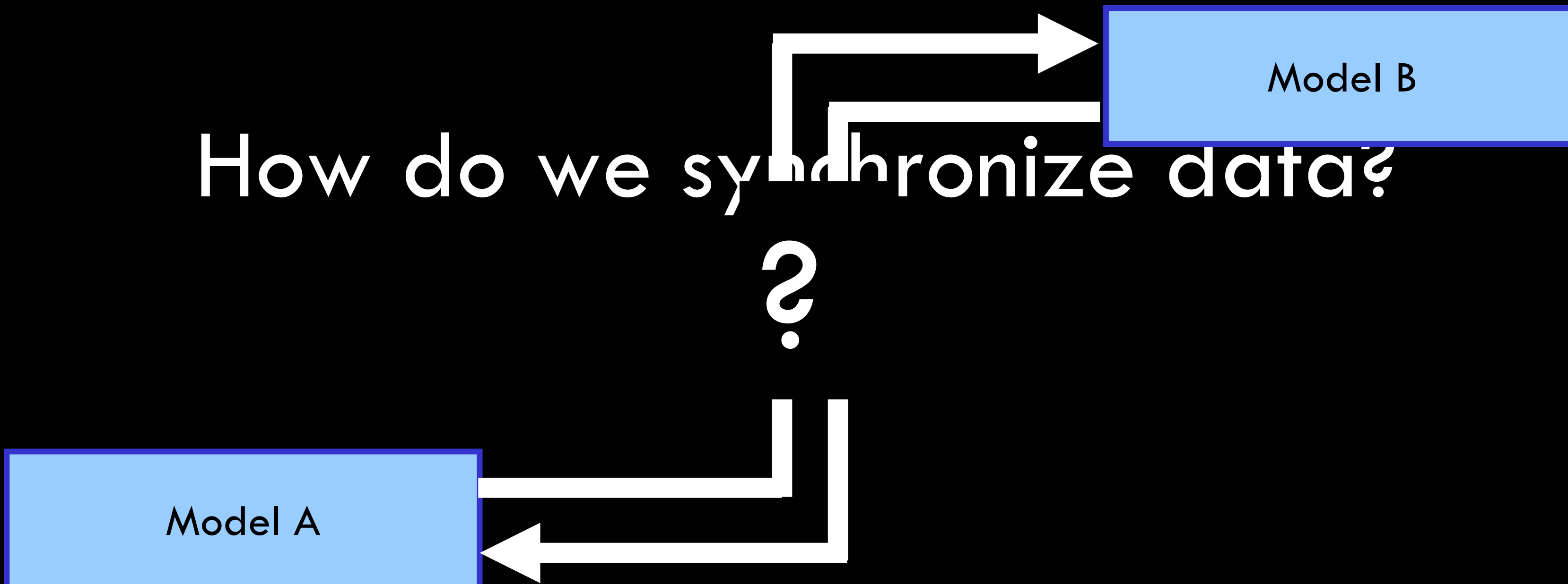
A Modeling Language

SYNCHRONIZE MODELS



SYNCHRONIZE DATA

How do we synchronize data?



SOLUTION

Standards!

STANDARDS ?

4ML	ARML	BiblioML	CIDX	eBIS-XML	HTTP-DRP	MatML	ODRL	PrintTalk	SHOE	UML	XML F
AML	ARML	BCXML	xCIL	ECML	HumanML	MathML	OeBPS	ProductionML	SIF	UBL	XML Key
AML	ASML	BEEP	CLT	eCo	HyTime	MBAM	OFX	PSL	SMML	UCLP	XMLife
AML	ASML	BGML	CNRP	EcoKnow	IML	MISML	OIL	PSI	SMBXML	UDDI	XML MP
AML	ASTM	BHTML	ComicsML	edaXML	ICML	MCF	OIM	QML	SMDL	UDEF	XML News
AML	ATML	BIBLIOML	Covad xLink	EMSA	IDE	MDDL	OLifE	QAML	SDML	UIML	XML RPC
AML	ATML	BIOML	CPL	eosML	IDML	MDSI-XML	OML	QuickData	SMIL	ULF	XML Schema
ABML	ATML	BIPS	CP eXchange	ESML	IDWG	Metarule	ONIX DTD	RBAC	SOAP	UMLS	XML Sign
ABML	ATML	BizCodes	CSS	ETD-ML	IEEE DTD	MFDX	OOPML	RDDI	SODL	UPnP	XML Query
ACML	AWML	BLM XML	CVML	FieldML	IFX	MIX	OPML	RDF	SOX	URI/URL	XML P7C
ACML	AXML	BPML	CWMI	FINML	IMPP	MMLL	OpenMath	RDL	SPML	UXF	XML TP
ACAP	AXML	BRML	CycML	FITS	IMS Global	MML	Office XML	RecipeML	SpeechML	VML	XMLVoc
ACS X12	AXML	BSML	DML	FIXML	InTML	MML	OPML	RELAX	SSML	vCalendar	XML XCI
ADML	AXML	CML	DAML	FLBC	IOTP	MML	OPX	RELAX NG	STML	vCard	XAML
AECM	BML	xCML	DaliML	FLOWML	IRML	MoDL	OSD	REXML	STEP	VCML	XACML
AFML	BML	CaXML	DaqXML	FPML	IXML	MOS	OTA	REPML	STEPML	VHG	XBL
AGML	BML	CaseXML	DAS	FSML	IXRetail	MPML	PML	ResumeXML	SVG	VIML	XSBEL
AHML	BML	xCBL	DASL	GML	JabberXML	MPXML	PML	RETML	SWAP	VISA XML	XBN
AIML	BML	CBML	DCMI	GML	JDF	MRML	PML	RFML	SWMS	VMML	XBRL
AIML	BML	CDA	DOI	GML	JDox	MSAML	PML	RightsLang	SyncML	VocML	XCFF
AIF	BannerML	CDF	DeltaV	GXML	JECMM	MTML	PML	RIXML	TML	VoiceXML	XCES
AL3	BCXML	CDISC	DIG35	GAME	JLife	MTML	PML	RoadmOPS	TML	VRML	Xchart
ANML	BEEP	CELLML	DLML	GBXML	JSML	MusicXML	PML	RosettaNet PIP	TML	WAP	Xdelta
ANNOTEA	BGML	ChessGML	DMML	GDML	JSML	NAML	PML	RSS	TalkML	WDDX	XDF
ANATML	BHTML	ChordML	DocBook	GEML	JScoreML	xNAL	P3P	RuleML	TaxML	WebML	XForms
APML	BIBLIOML	ChordQL	DocScope	GEDML	KBML	NAA Ads	PDML	SML	TDL	WebDAV	XGF
APPML	BIOML	CIM	DoD XML	GEN	LACITO	Navy DTD	PDX	SML	TDML	WellML	XGL
AQL	BIPS	CIML	DPRL	GeoLang	LandXML	NewsML	PEF XML	SML	TEI	WeldingXML	XGMML
APPEL	BizCodes	CIDS	DRI	GIML	LEDES	NML	PetroML	SML	ThML	Wf-XML	XHTML
ARML	BLM XML	CIDX	DSML	GXD	LegalXML	NISO DTB	PGML	SAML	TIM	WIDL	XIOP
ARML	BPML	xCIL	DSD	GXL	Life Data	NITF	PhysicsML	SABLE	TIM	WITSML	XLF
ASML	BRML	CLT	DXS	Hy XM	LitML	NLMXML	PICS	SAE J2008	TMML	WorldOS	XLIFF
ASML	BSML	CNRP	EML	HITIS	LMML	NVML	PMML	SBML	TMX	WSML	XLink
ASTM	BCXML	ComicsML	EML	HR-XML	LogML	OAGIS	PNML	Schemtron	TP	WSIA	XMI
ARML	BEEP	CIM	DLML	HRMML	LogML	OBI	PNML	SDML	TPAML	XML	XMSG
ARML	BGML	CIML	EAD	HTML	LTSC XML	OCF	PNG	SearchDM-XML	TREX	XML Court	XMTP
ASML	BHTML	CIDS	ebXML	HTTPL	MAML	ODF	PrintML	SGML	TxLife	XML EDI	XNS

DATA CENTER

a proposal …

- Dictionary
- Grammar

Dictionary

DATA CENTER

PROPOSAL

Use actual words

Calls

Days

Amount

Units

DATA CENTER

DICTIONARY ENTRY

a key *a word*

call *n.*

1. A loud cry, a shout.
2. The characteristic cry of an animal.
3. **A telephone communication or connection.**
4. Need or occasion.

call.3

ENTRY: call

call ***n.*** (*call*)

A telephone connection or conversation.

Syn. telephone call, phone call

Type of telephone.2, telephony

Attributes telephone number

DATA CENTER

ENTRY: model

model *n.* (***model***)

A simplified or idealized description or conception of a particular system, situation, or process, often in mathematical terms, that is put forward as a basis for theoretical or empirical understanding, or for calculations, predictions, etc.; a conceptual or mental representation of something.

Type of hypothesis, possibility.5, theory.2

Types simulation.4, computer simulation, stochastic process

Attributes name, identification.3, description.2, state, expression.4, model

ENTRY: AuthorizedPricingInformation

AuthorizedPricingInformation *p.* (*AuthorizedPricingInformation*)

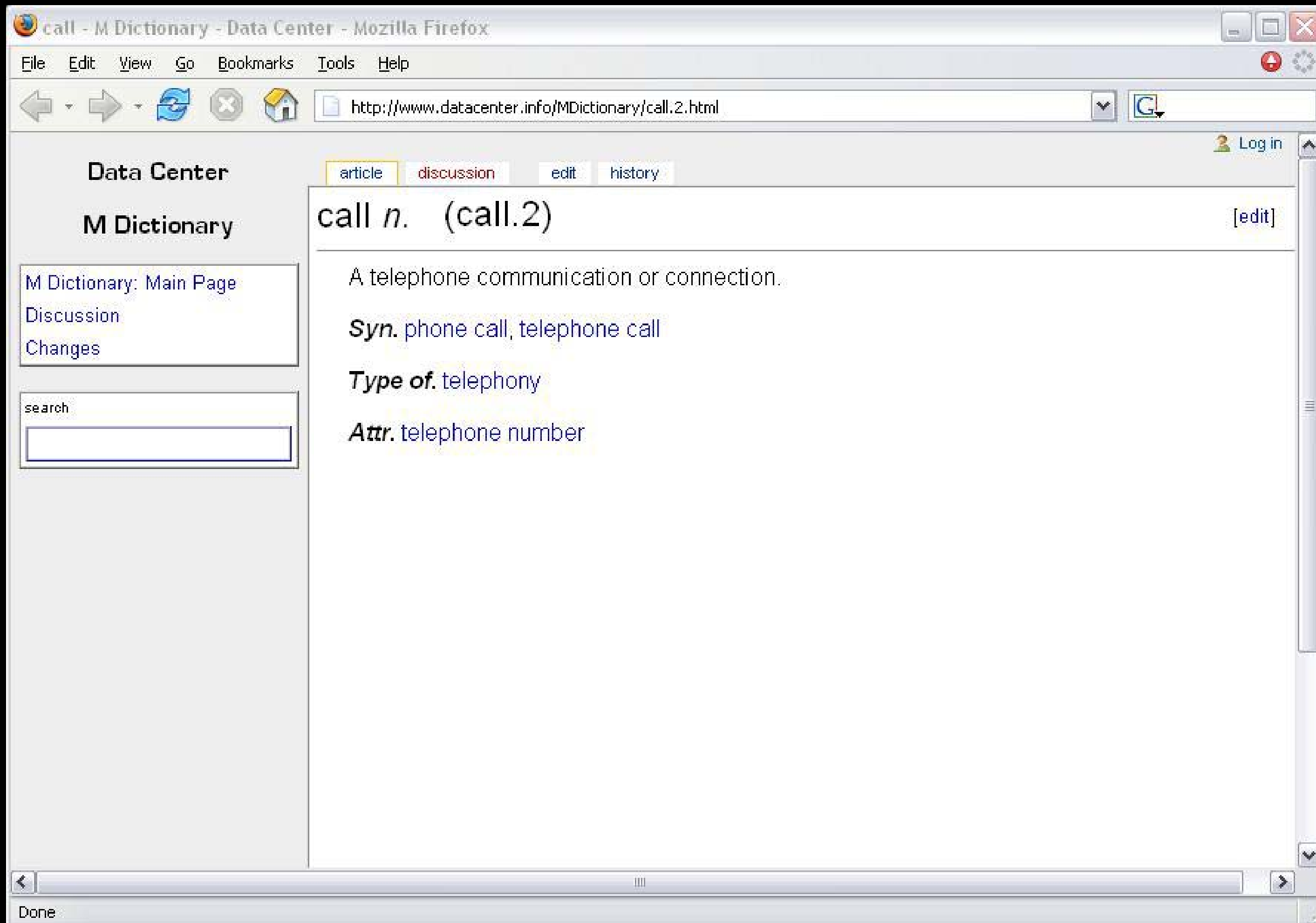
The collection of business properties that describe the supplier's product cost issued to a distributor that is below distributor's book cost.

Phrase. information, pricing; information, authorized

DICTIONARY DEVELOPMENT

- Web accessible
- Web editable!
- Web community
- Staged approval
  - Proposal – Universal accessible and editable
  - Draft – Universal accessible and limited editable
  - Pre-approval – Universal accessible and limited comments
  - Recommendation – Universal accessible

DICTIONARY


call - M Dictionary - Data Center - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
http://www.datacenter.info/MDictionary/call.2.html
Log in
Data Center
M Dictionary
article
discussion
edit
history
call n. (call.2)
[edit]
A telephone communication or connection.
Syn. phone call, telephone call
Type of. telephony
Attr. telephone number
M Dictionary: Main Page
Discussion
Changes
search
Done

DATA CENTER

DICTIONARY DEVELOPMENT

Oxford English Dictionary

National Library of Medicine
Unified Medical Language System

United States Department of Defense

Princeton University, WordNet

American Chemical Society
Chemical Abstracts Service

West Law Publishing
Black's Law Dictionary

Acronym Finder
Acronym Dictionary

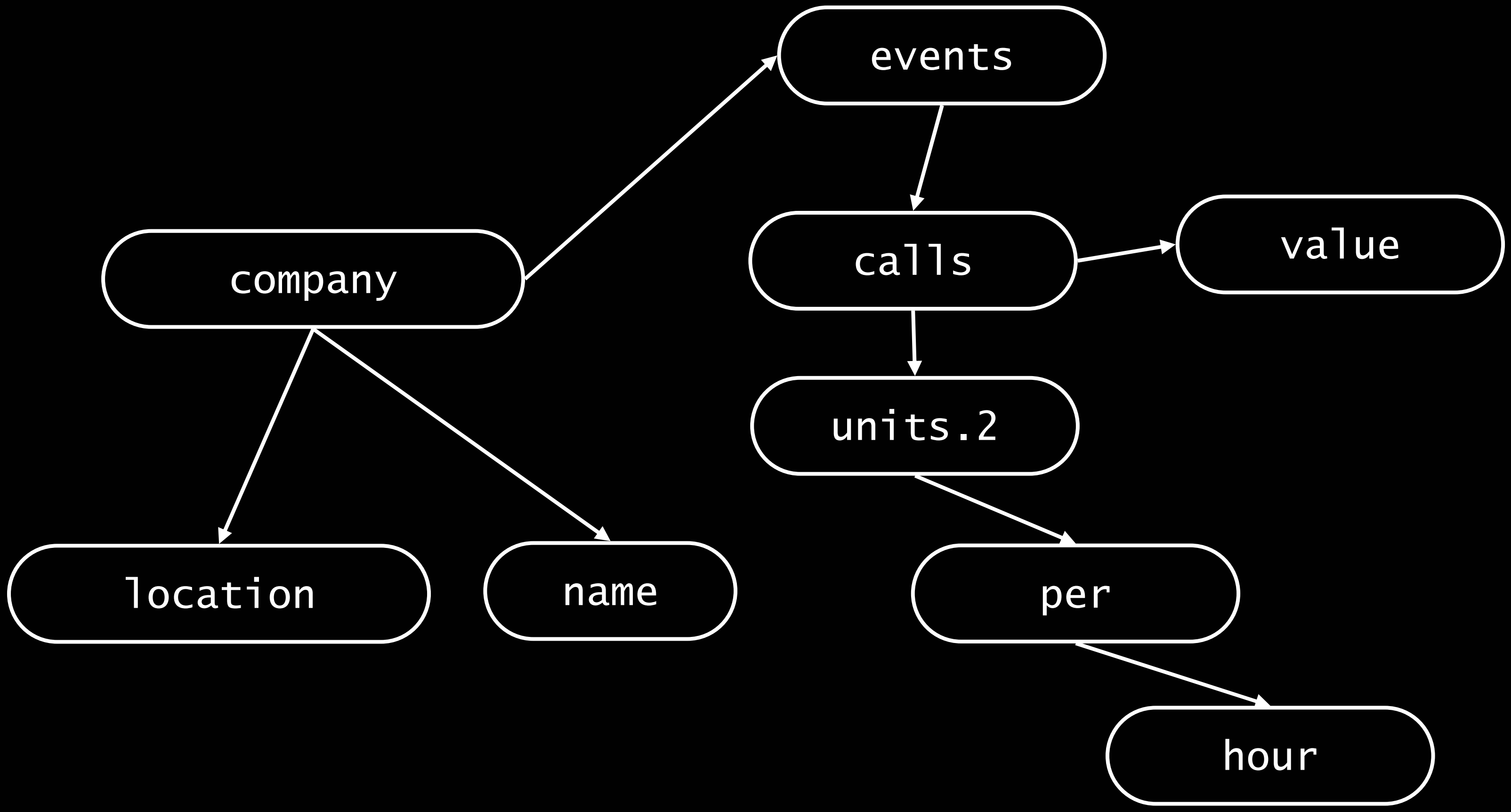
DATA CENTER

Grammar

XML "GRAMMAR"

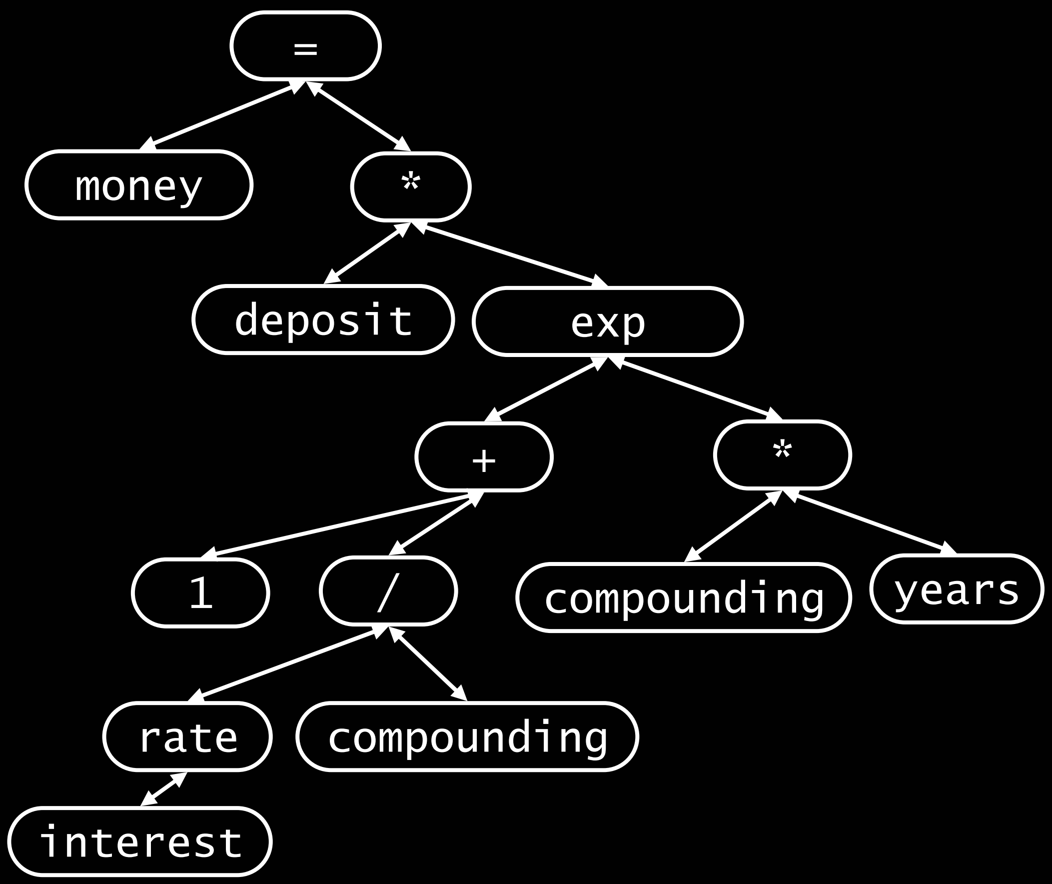
```
<CompanyData>
   <CompanyName>
      Fidelity Employer Services Company
   </CompanyName>
   <Location>
      Merrimack
   </location>
   <CallData>
     <RecordDate>
        Tue Aug 11, 2004
     </RecordDate>
     <CallsPerDay>
        2575
     </CallsPerDay>
   </CallData>
</CompanyData>
```

GRAMMAR



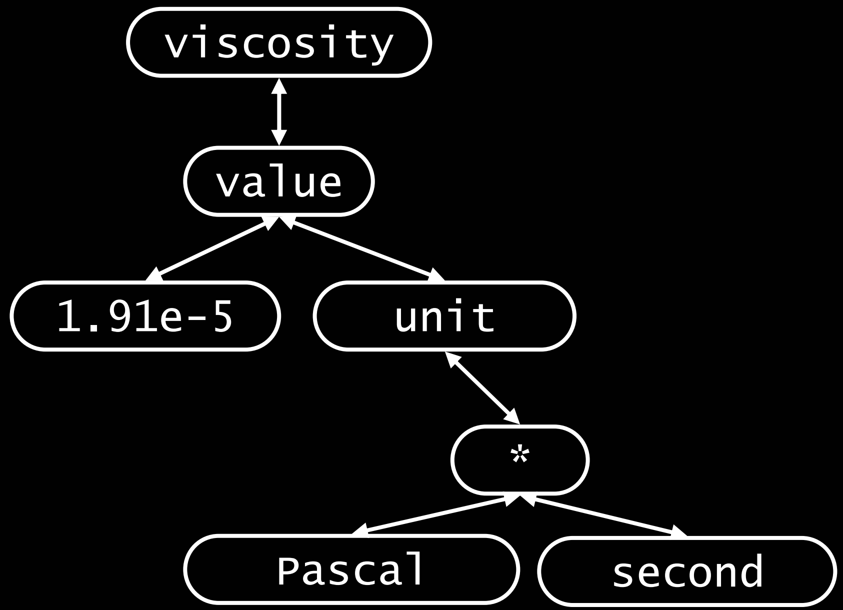
Equations ...

$$P = C\,(1 + r/n)^{nt}$$



Engineering units ...

viscosity 1.91×10^{-5} Pa s

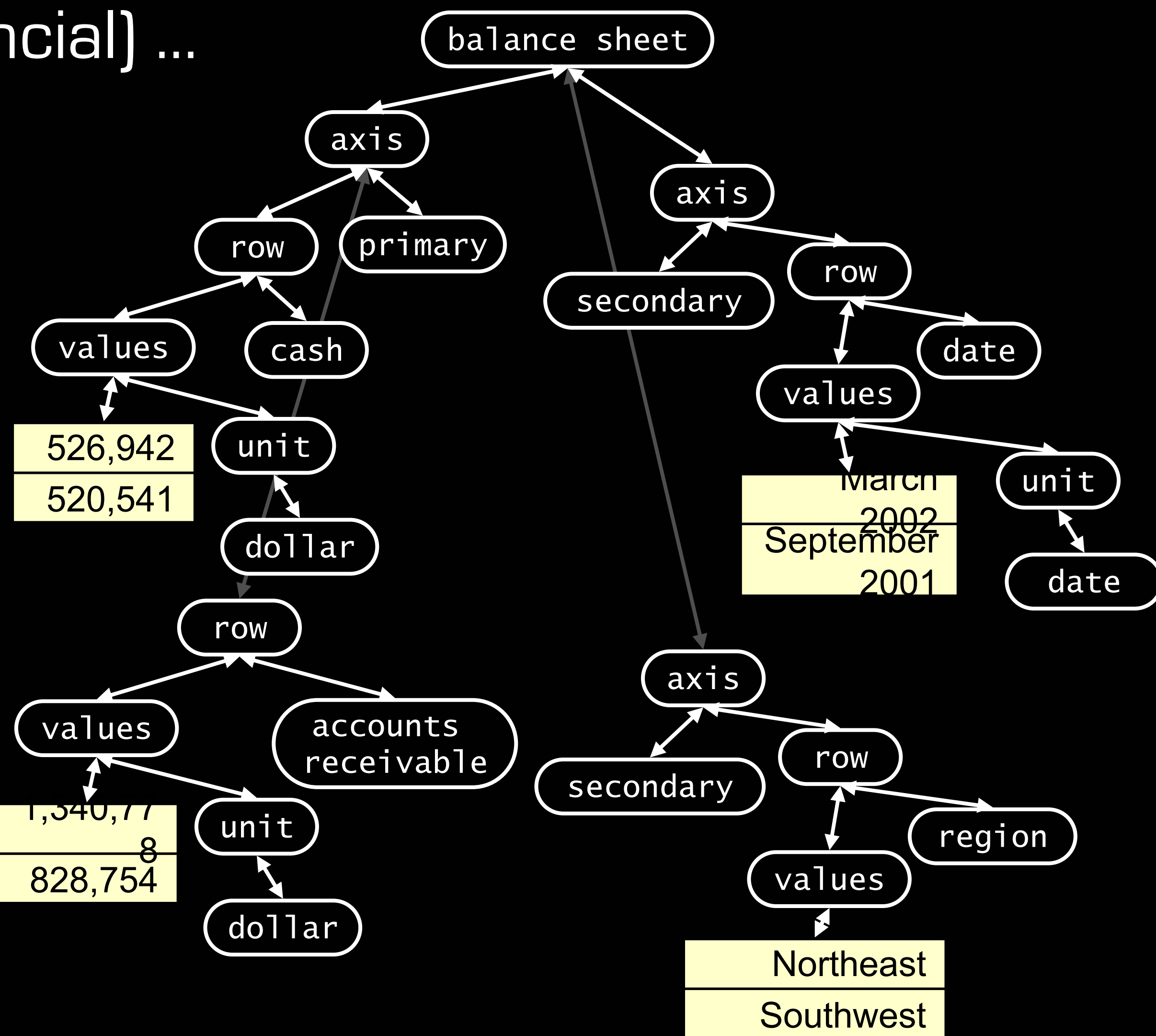


DATA CENTER

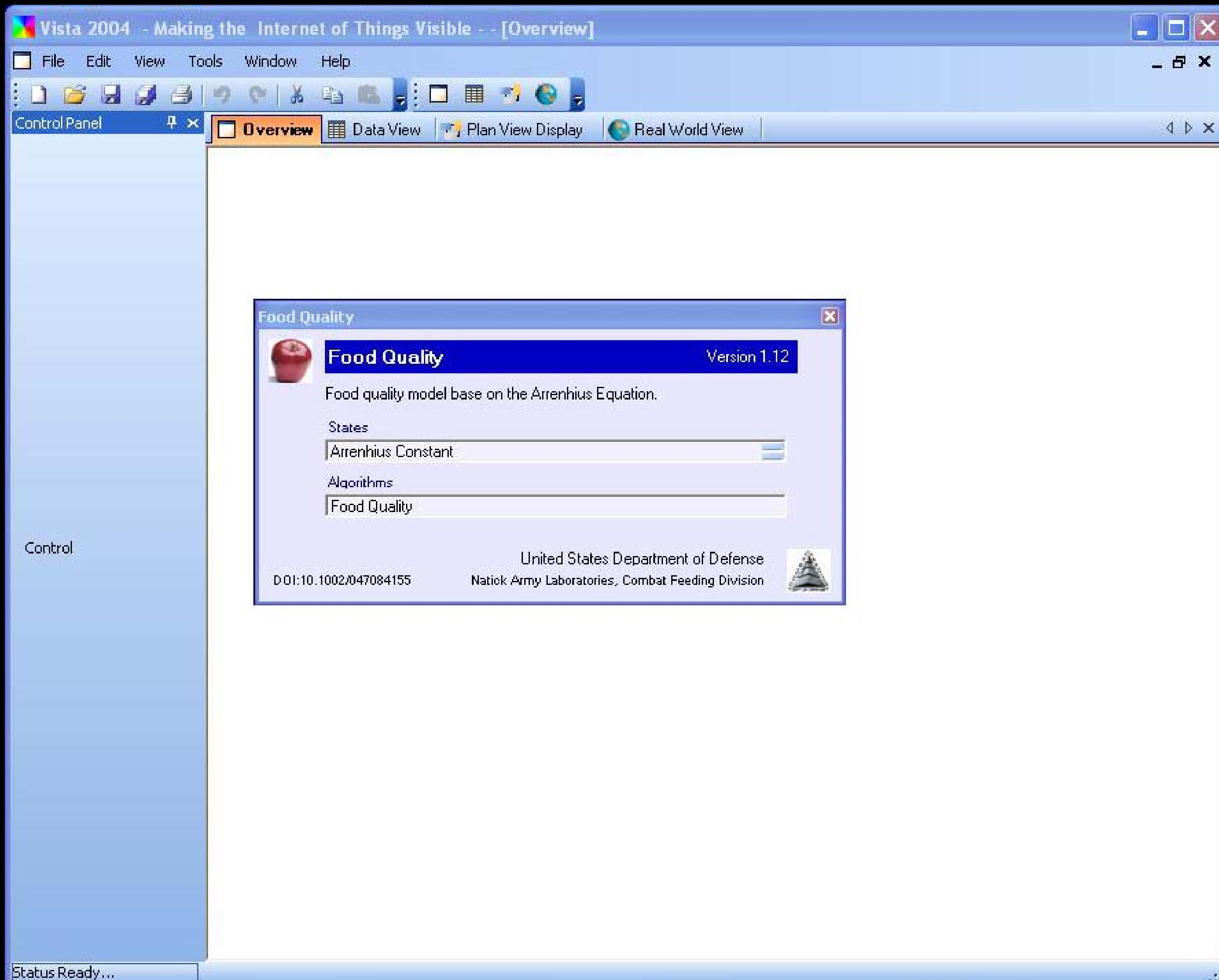
FINANCE

Spreadsheets (financial) ...

Assets	March 2002	September 2001
Cash	526,942	520,541
Accounts Receivable	1,340,778	828,754
Prepaid Expense	75,429	78,639
Equipment	84,102	93.393
Total Assets	**2,027,251**	**1,521,327**
Liabilities		
Accounts Payable	16,354	10,446
Deferred License	869,119	881,014
Unearned Support Revenue	295,957	312,110
Accrued Payroll/Expenses	87,861	80,372
Total Liabilities	**1,269,291**	**1,283,942**
Retained Earnings	**757,960**	**237,385**
Total Liabilities and Retained Earnings	**2,027,251**	**1,521,327**

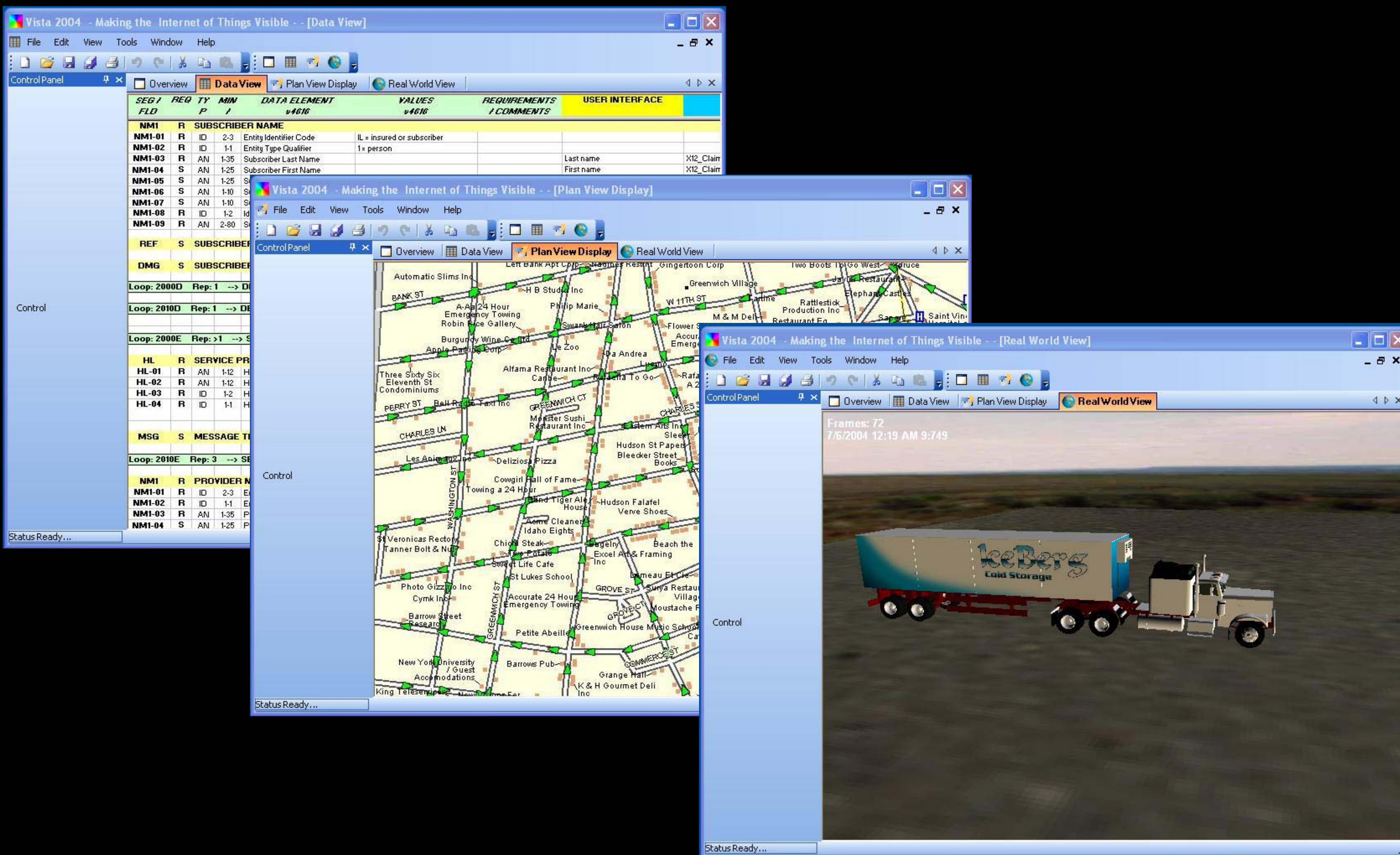


M-ENGINE

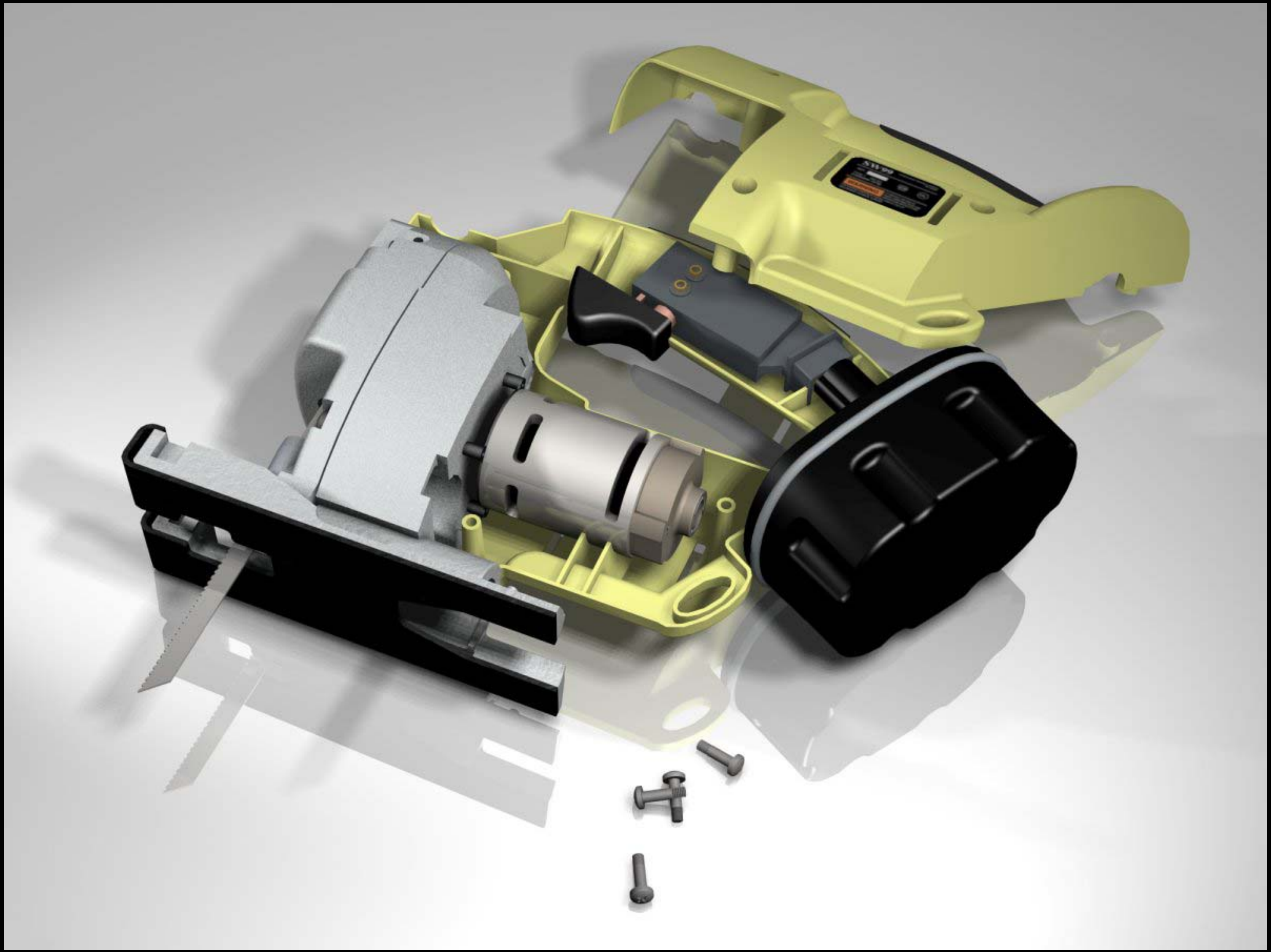

Vista 2004 - Making the Internet of Things Visible - - [Overview]
File Edit View Tools Window Help
Control Panel
Overview
Data View
Plan View Display
Real World View
Food Quality
Food Quality
Version 1.12
Food quality model base on the Arrenhius Equation.
States
Arrenhius Constant
Algorithms
Food Quality
United States Department of Defense
DOI:10.1002/047084155
Natick Army Laboratories, Combat Feeding Division
Control
Status Ready...

DATA CENTER

M-ENGINE



APPLICATIONS – PRODUCT DESIGN

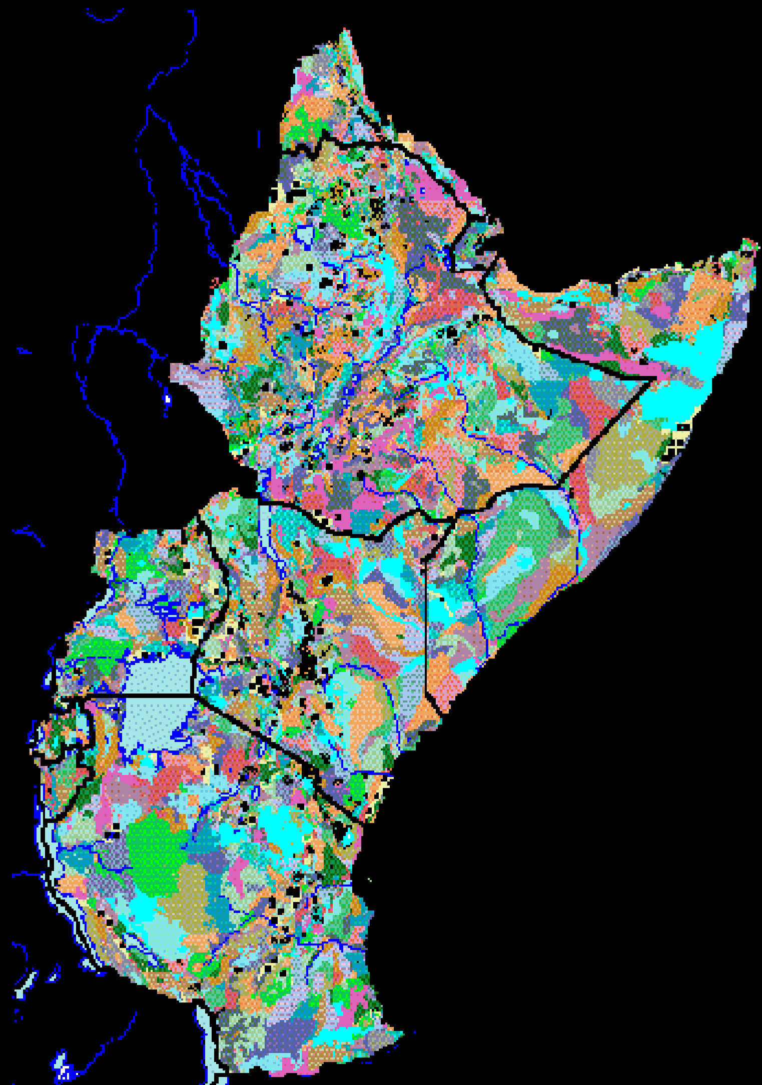


APPLICATIONS – LOGISTICS

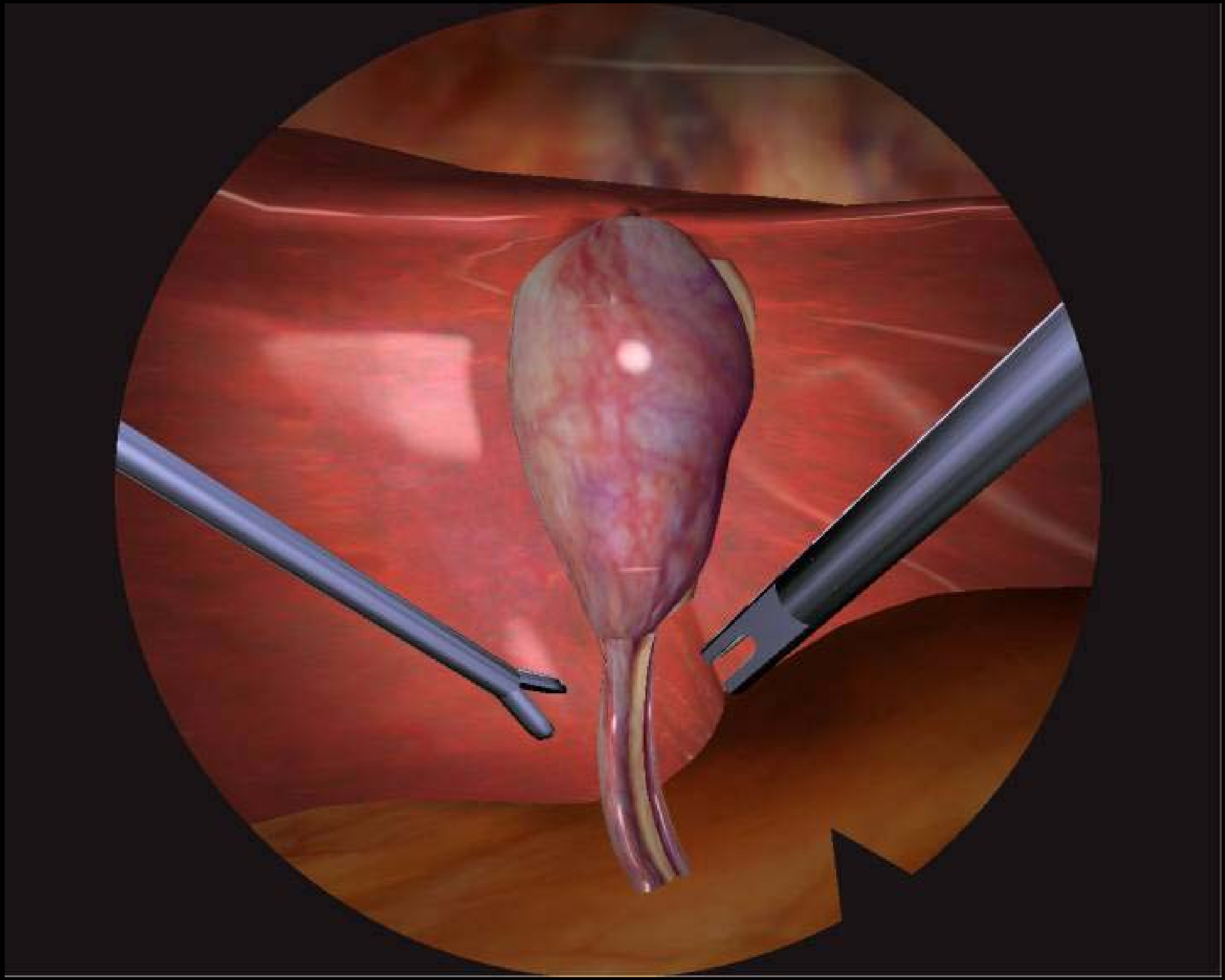
APPLICATIONS –TRAFFIC AND ROUTING

APPLICATIONS – MANUFACTURING

APPLICATIONS – ENVIRONMENTAL IMPACT



APPLICATIONS – HEALTHCARE



SUMMARY

Summary

DATA CENTER

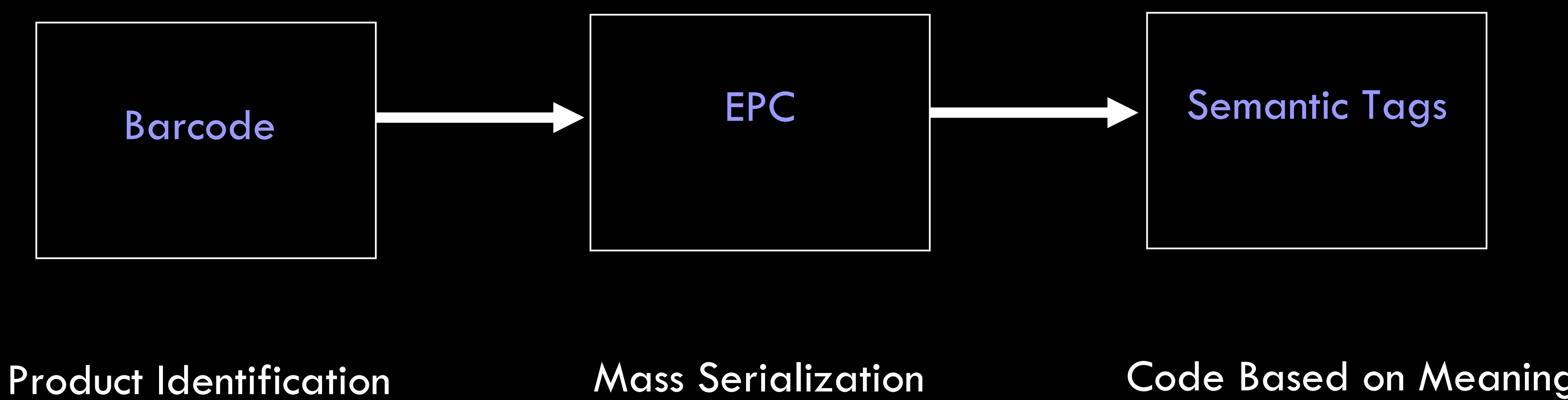
TASKS

- Establish a common dictionary
  - One definition per key
- Use graphs to express relationship
  - Graphs represent semantic relationship
- Establish an infrastructure that handles model interoperability

TYPES OF "WEBS"

- The Web of Information
  - –HTML and the World Wide Web
- The Web of Things
  - –Linking physical objects together using Auto-ID
- The Web of Abstractions
  - –Building a network of mathematical models
  - –Link models together
  - –Link data to models
  - –Computer languages & protocols to create a free flow of models in a network (Internet or Intranet)

THE PATH OF IDENTIFICATION



THE FUTURE...

Supply chains that sense and respond to the physical world.

This requires an **Intelligent Infrastructure** for management, control, and automation.

The initial base of the infrastructure is the Electronic Product Code (EPC).

A serial number does not adequately describe an abstraction like a model.

DATA CENTER

SEMANTIC MODELING – THE GOAL

- Communication of models between computers to create **interoperability**
- Run **distributed** models across the Internet
- Increased model **sharing** and **re-use** of model elements

SEMANTIC MODELING – THE GOAL (CONTINUED)

- Increase the **productivity** of modeling
  - Reduce trial & error
  - Improve mathematical intuition
  - Reduce dependence on literature search
- Redefine the **link** between models and data…and data to data
- Share models across **domains**

IMPLICATIONS FOR LOGISTICS AND OPERATIONS

- Logistics depends on the flow of data for effective management.
- Auto-ID and other sensing technologies will increase the flow of data.
- Practitioners will need models to interpret data streams
  - Inventory, transportation, warehousing, customer service, purchasing…

BASIC QUESTIONS

What are the relationships between models?

How are models connected?

In the future, the definition of a model and the sharing of models though a network will become as important as the model itself.

Meaning arises by the way one model is connected or related to other models

EARLY WORK IN THE FIELD

- GEOFFRION, A.M., **1987**. “An Introduction to Structured Modeling.” *Management Science* 33:5.

- GEOFFRION, A.M., **1989**. “The Formal Aspects of Structural Modeling.” *Operations Research* 37:1.

- MUHANNA, W.A. and R.A. PICK, **1994**. “Meta-modeling Concepts and Tools for Model Management: A Systems Approach.” *Management Science* 40:9.

RECENT CONCEPTUAL WORK

- BROCK, D.L. **2000.** "Intelligent Infrastructure – A Method for Networking Physical Objects." *MIT Smart World Conference.*
- BROCK, D.L. **2003.** "The Data Project – Technologies, Infrastructure and Standards for Distributed Interoperable Modeling and Simulation." *MIT Data Project Workshop,* September.

2004 E. GROSVENOR PLOWMAN AWARD

- BROCK, D.L., E.W. SCHUSTER, S.J. ALLEN and P. KAR, **2005**. "An Introduction to Semantic Modeling for Logistical Systems." *Journal of Business Logistics*, forthcoming.

- SCHUSTER, E.W., D.L. BROCK, S.J. ALLEN and P. KAR, 2004. "Prototype Applications for Semantic Modeling." *Smart World 2004*, MIT.

RECENT APPLIED WORK

- GAZMURI, P and MATURANA, S. **2001**. “Developing and Implementing a Production Planning DSS for CTI Using Structured Modeling.” *Interfaces* 31:4.

PROPOSED SYSTEM - M

- David Brock, Inventor and Chief Architect
- Initial Design – a System of Languages and Protocols – merged into "M"
  - **Data Modeling Language** (DML) is a semantic for describing modular, interoperable model components in terms of individual outputs, inputs and data elements.
  - **Data Modeling Protocol** (DMP), once a connection between models and data is established, the DMP coordinates the communication sequence between the computing machines that host models in terms of outputs and inputs.

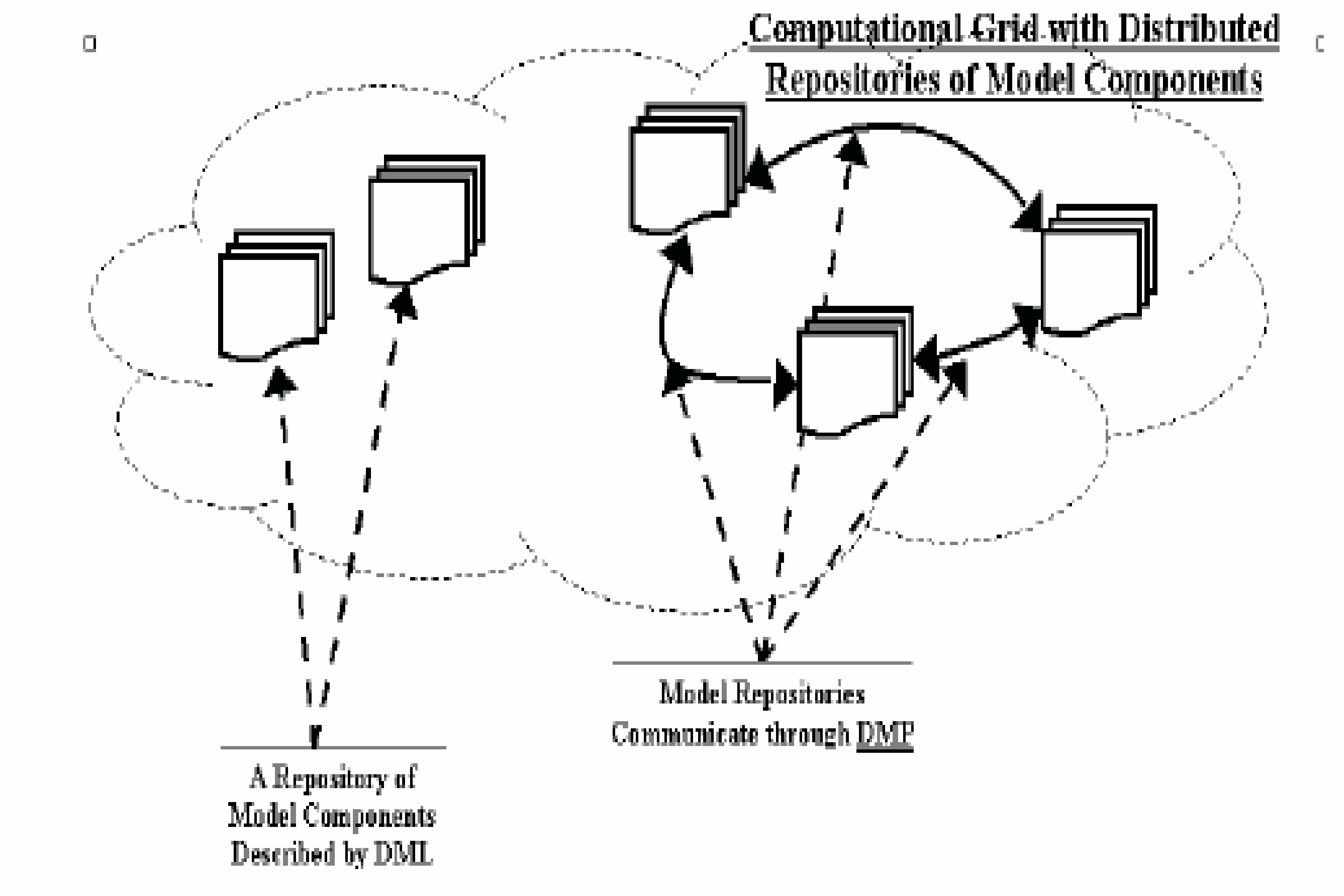
PROPOSED SYSTEM – M (CONTINUED)

- Initial Design – a System of Languages and Protocols – merged into "M"
  - **Automated Control Language** (ACL) establishes the connection between models and data based on DML (descriptor of inputs, outputs and data) and the ACP, which locates the appropriate connections.
  - **Automated Control Protocol** (ACP) helps model outputs and inputs locate one another within a network, even though the individual models may exist in different host systems and organizations. The ACP identifies potential connections and takes priority over the DMP, which is a coordinating activity after achieving connections through the ACL.

PROPOSED SYSTEM – M (CONTINUED)

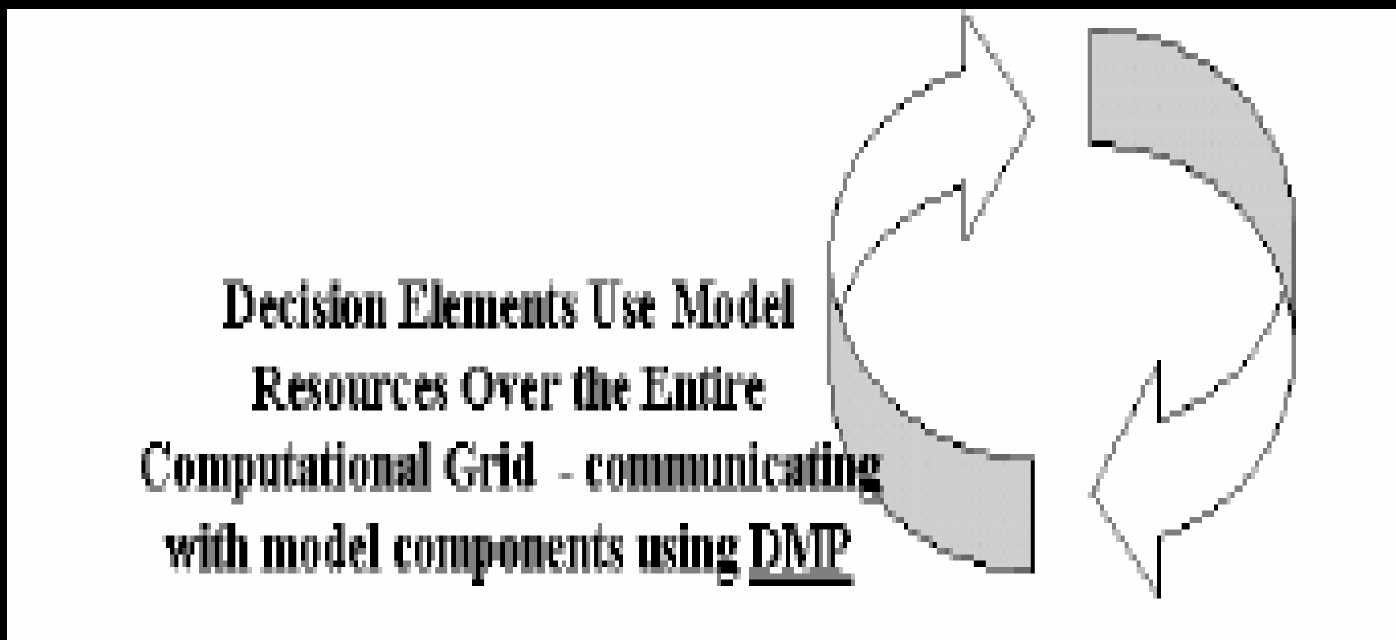
- The Dictionary
  - **Dictionary** a common resource containing words with multiple meanings. The dictionary will utilize established sources such as the Oxford English Dictionary, WordNet, and various specialty dictionaries from the medical field, operations, logistics and other disciplines.

A VISUALIZATION OF M

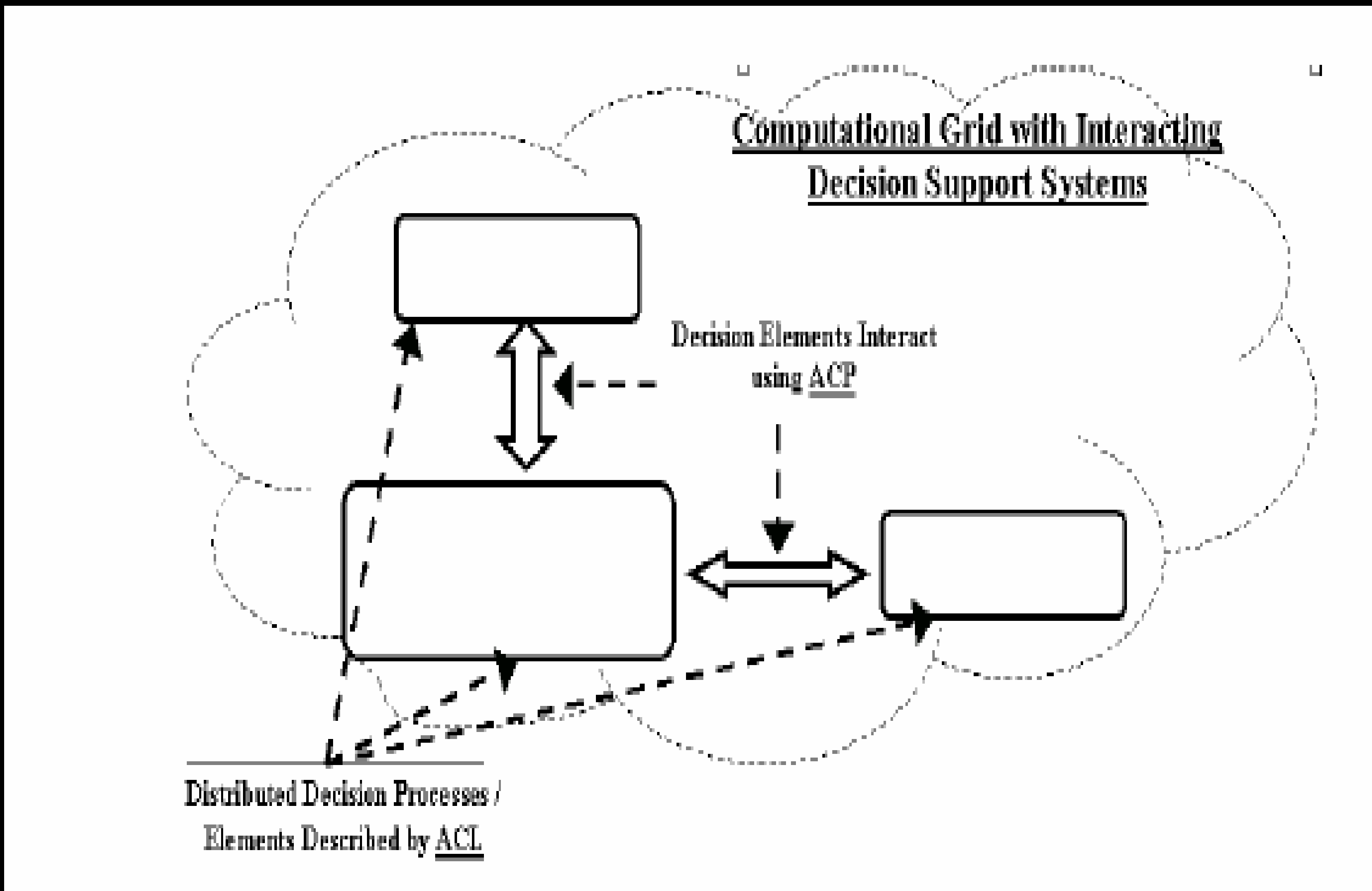


DATA CENTER

GRID COMPUTING COORDINATION



A VISUALIZATION OF M



DATA CENTER

APPROACHES TO PLANNING AND SCHEDULING

Attribute	Math Programming	Simulation	Heuristic
Hold Time		x	x
Queue Time		x	x
Customer Service		x	
Forecast Bias		x	
Set-up Cost	x		x
Holding Cost	x		x
Overtime Cost	x		x
Capacity	x		x
Production Lot Size	x		x
Production Sequence	x		x
Customer Due Date	x	x	x
Family Structure	x		

MODEL A - Deterministic Simulation (Schuster and Finch 1990) – With bias adjusted safety stocks that use customer service levels as an input, production planning occurs for each item independently. All items run on a production line are summed to give a total capacity load. This model initially assumes infinite capacity is available for production and does not consider set-up or inventory carrying cost. However, the model does provide a method for safety stock planning that considers dynamic forecasts and the impact of forecast bias in planning safety stock levels.

Schuster, Edmund W. and Byron J. Finch (1990), "A Deterministic Spreadsheet Simulation Model for Production Scheduling in a Lumpy Demand Environment," *Production and Inventory Management Journal*, Vol. 31, No. 1, pp. 39 – 43.

MODEL B (1994) - Mathematical Programming (Allen and Schuster 1994) – Exploiting the fact that consumer goods have a family structure defined by package size, production can be planned using a two-tier hierarichal structure where product families are sequenced with disaggregation taking place to form end item schedules. This approach provides optimal solutions based on cost and utilizes an innovative mathematical formulation that yields near instantaneous solutions to mixed integer math programming problems.

Allen, Stuart J. and Edmund W. Schuster (1994), "Practical Production Scheduling with Capacity Constraints and Dynamic Demand: Family Planning and Disaggregation," *Production and Inventory Management Journal*, Vol. 35, No. 4, pp. 15-21.

MODEL C (1997) - The MODS Heuristic, Sequence Independent (Allen et al. 1997) – An approach to scheduling using the Modified Dixon Silver (MODS) method to calculate near optimum production schedules based on inventory and set-up costs, and inventory set-up time.

Allen, Stuart J., Jack L. Martin, and Edmund W. Schuster (1997), "A Simple Method for the Multi-item, Single-level, Capacitated Scheduling Problem with Set-up Times and Costs," *Production and Inventory Management Journal*, Vol. 38, No. 4, pp. 39-47.

MODEL D (1998) - The MODS Heuristic, Sequence Dependent (D'Itri et al. 1998) – Building on the Modified Dixon Silver method, this approach utilizes the nearest neighbor variable origin (NNVO) heuristic as a second step to sequence production based on a "from-to" table of changeover costs between items.

D'Itri, Michael P., Stuart J. Allen, and Edmund W. Schuster (1998), "Capacitated Scheduling of Multiple Products on a Single Processor with Sequence Dependencies," *Production and Inventory Management Journal*, Vol. 40, No. 5, pp.

DATA CENTER

DATA INPUTS AS A SEMANTIC

Data Input	Model A	Model B	Model C	Model D
D1. Beginning Inventory	X	X	X	X
D2. Forecast Demand (by week)	X	X	X	X
D3. Historical Shipments (by week)	X	X	X	X
D4. Historical Forecast (by week)	X	X	X	X
D5. Hold Time (days)	X			
D6. Queue Time (days)	X			
D7. Service Level (% in stock)	X	X	X	X
D8. Set-up Cost ($/changeover)		X	X	X
D9. Set-up Time (hrs/set-up)			X	X
D10. Holding Cost ($/week)		X	X	X
D11. Capacity Limit (hrs/day)		X	X	X
D12. Family Structure (end items per group)		X		
D13. Overtime Cost ($/hr)			X	X
D14. Sequence Dependent Set-up Cost (From–To table of change-over costs)				X

FIRST PROTOTYPES

- Logistical Systems Including ERP
  - Forecasting, planning, scheduling, and inventory models
- Agricultural Models
  - Harvest risk and planning
- Retail
  - Lot sizing for short life-cycle products
  - Lillian Vernon, Inc.

ERP

- (ERP) system identifies and plans "the…resources needed to take, make, ship and account for customer orders."

  from *APICS Dictionary* (2004), Published by the American Production and Inventory Control Society, Alexandria, VA.

- To achieve these important tasks, ERP uses a variety of models and data to plan and control all the resources in a manufacturing or service-oriented company.

ERP (CONTINUED)

- Most organizations implement packaged ERP software that contains a single model for a specific business process.
- If the model does not exactly fit, substantial modifications are required.
- Managers often complain that this process of adaptation reduces overall organizational productivity.

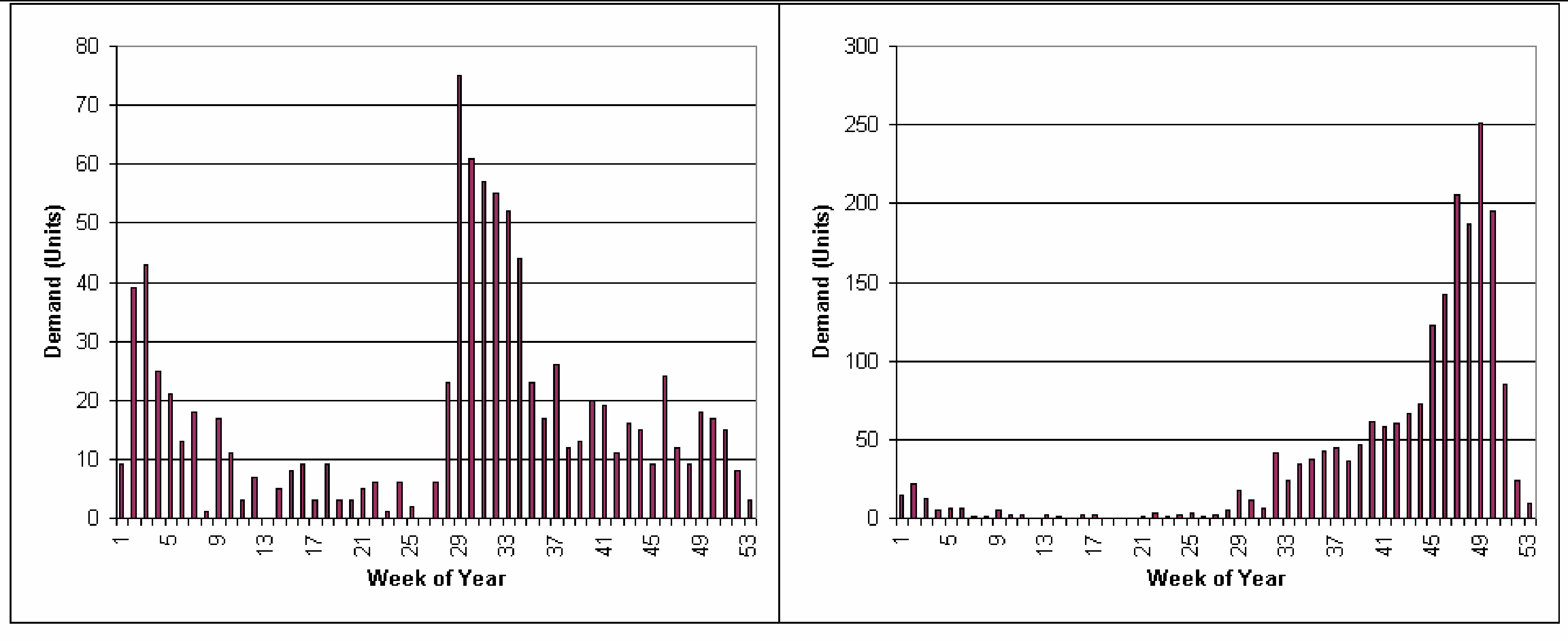
ERP (CONTINUED)

- Build a network of ERP models that could automatically match to data within organizations.
- These models include forecasting, production planning and scheduling, lot sizing, logistical, and financials.
- Create new models by combining the outputs of one model with the inputs of another model.

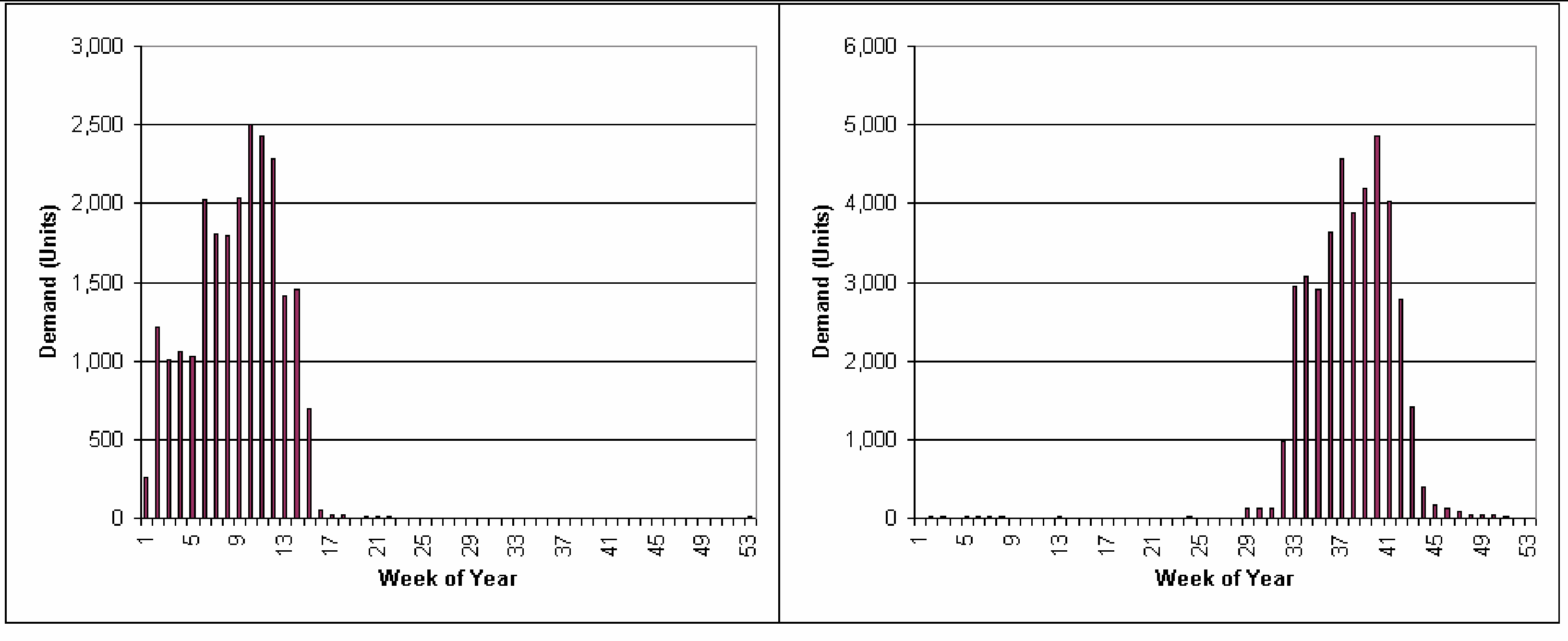
RETAILING – LILLIAN VERNON

- Direct Marketing – Data Rich Environment
- Entrepreneurial Company Established in 1951
- Product Line Equals 6,000 Items
- Short Life Cycles
- The Need to Quickly Apply Models to Data for Risk Analysis

LILLIAN VERNON – ONGOING ITEMS



LILLIAN VERNON – SEASONAL ITEMS



AGRICULTURE

- Lack of Model Use in Agriculture
- Move Toward Precision Agriculture
- The Need to do Rapid Modeling
- Combine Data Sets (GIS)
- A More Detailed Presentation

FIRST PROTOTYPES (CONTINUED)

- More General View of Semantic Modeling
  - Method to search and re-use elements of mechanical designs (**automobile industry**)
  - Communication between different divisions within a conglomerate (**medical industry**)
  - Analyzing news releases (**financial services**)

DATA CENTER

NEXT STEPS…

- Smart World 2004 – Semantic Modeling, Dec, 8
  - Over 300 attended, mostly from industry
  - Attendees from Asia and Europe, as well as the US
- Support from the MIT Industrial Liaison Program
- Speakers representing Intel, IBM, Microsoft, SAP, P&G, J&J, and MIT
- Establish **The Data Center**, January 1, 2005
- This is large project that will take participation from industry and academia
- We are recruiting sponsors to move M and Semantic Modeling forward.

DATA CENTER

NEXT STEPS (CONTINUED)

“Interoperable Modeling and the M Language”

Workshop leader – David L. Brock

January 12, 2005

NEXT STEPS (CONTINUED)

“Defining the Dictionary”

Workshop Leader – David L. Brock

January 20, 2005

NEXT STEPS (CONTINUED)

"Business Applications"

Workshop Leaders – David L. Brock and Edmund W. Schuster

January 27, 2005

CONCLUSION