Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

# MACHINE-TO-MACHINE COMMUNICATION

Edmund W. Schuster
Laboratory for Manufacturing and Productivity
Massachusetts Institute of Technology

Jun. 23, 2011

*The MIT Motto*

Mens et Manus

“Mind and Hand”

Prof. Ian Waitz on MIT History

-- First 75 years: **Build Infrastructure**

-- Second 75 years: **Build Industries**

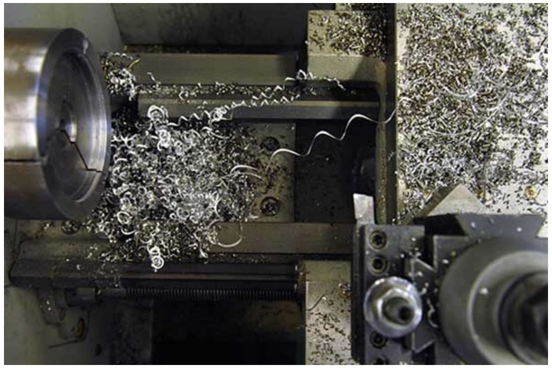
-- Next 75 years: **Build a better Planet**

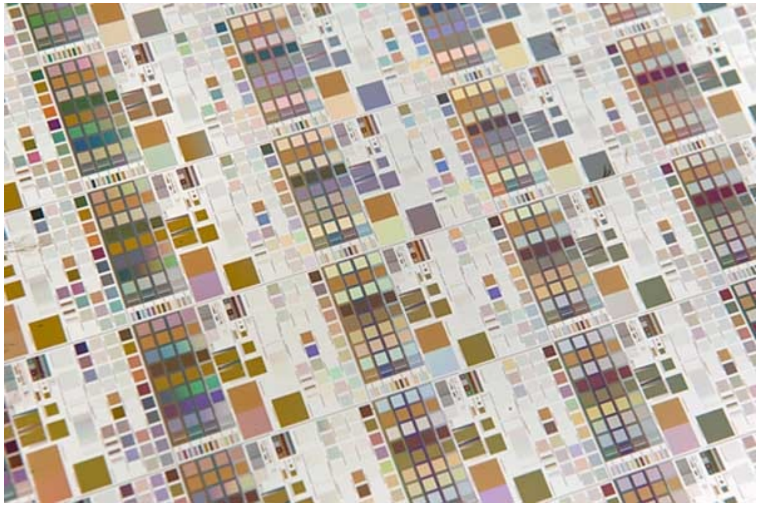
# INTRODUCTION

# MIT LABORATORY FOR MANUFACTURING AND PRODUCTIVITY

Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

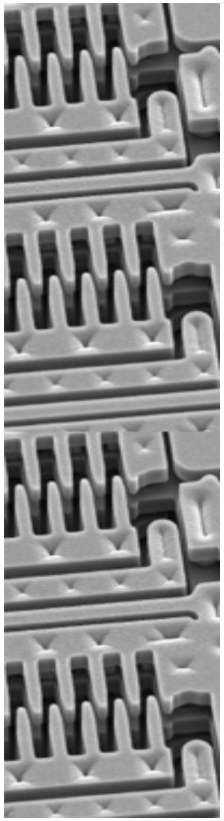
# CHANGE





Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

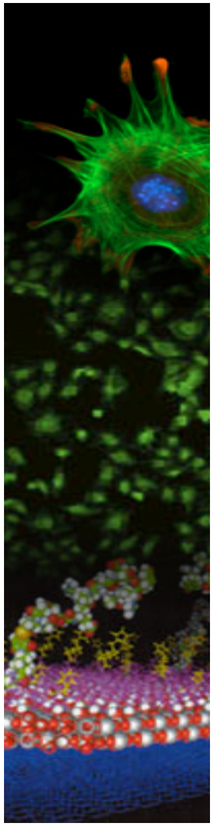
# INTERDISCIPLINARY



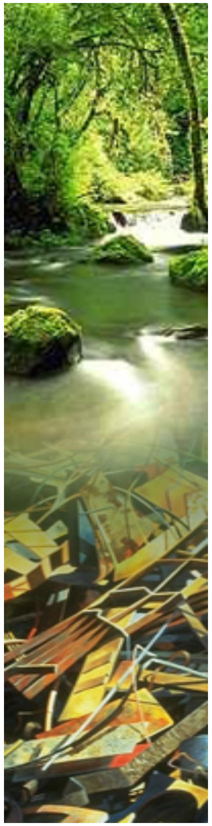
**nano**

**macro**

**info**



**bio**

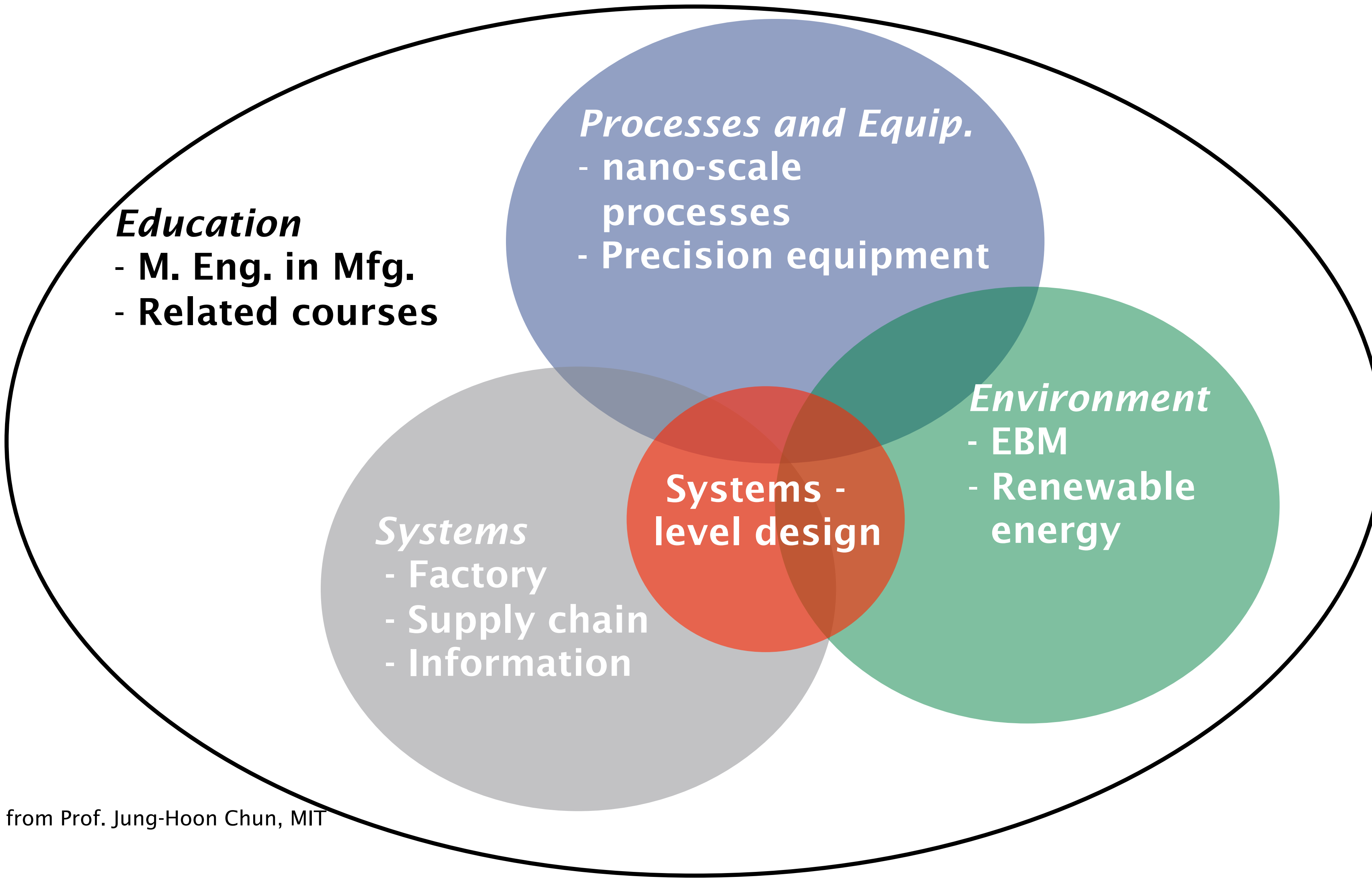


**eco**

Adapted from Prof. Jung-Hoon Chun, MIT

# THREE FOCUS AREAS

- **Renewable Energy and Environmentally Benign Manufacturing**
  - Sustainable technologies and manufacturing processes
  - Photovoltaics
- **Micro and Nano Scale Manufacturing**
  - Processes for micro fluidic devices
  - Nano positioning devices
  - Fuel cells
  - Surface science and engineering
- **Manufacturing Systems and Information Technology**
  - Ubiquitous computing, supply chain analysis, and Internet computing




Adapted from Prof. Jung-Hoon Chun, MIT

Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

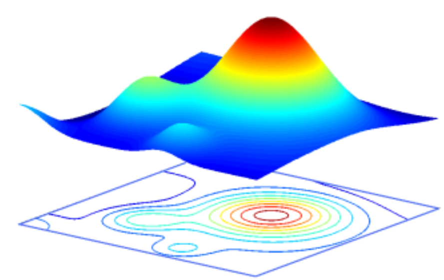
# ORGANIZATION

LMP – School of Engineering Interdisciplinary Group

About 150 Faculty, Researchers, Staff, and Students

Degrees: Ph.D., SM, Meng, SB

AUTO-ID LABS



Field Intelligence Lab

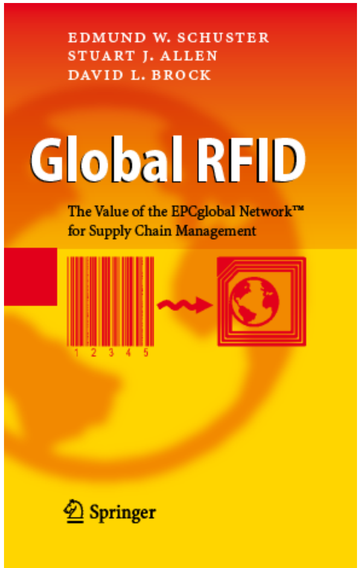
# Ubiquitous computing

Post-desktop model of human-computer interaction where information processing has been integrated into everyday objects and activities.

Ref. http://en.wikipedia.org/wiki/Ubiquitous_computing

Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

# AUTO-ID LABS




- **Historical**
  - Connect physical objects to the internet
  - Leader in RFID research for supply chain
  - Prototype industry academic consortium
  - Creation of EPCglobal in 2003, part of GS1
  - Network of Labs
    - Fudan University, Keio University, KAIST, U. of Adelaide, ETH, Cambridge

- **Future**
  - Wide concept of application
  - Mix with other technologies
    - Active tags, 2-d bar codes, linear bar codes, GPS, WiMax

# PROJECTS

- **RFID in Challenging Environments**
  - Improved performance with meta-material tags
- **Beyond Identification**
  - Tag Antenna based Sensors
- **Software Tools for RFID**
  - RFIDSim - a physical and logical layer simulation engine
  - Open Source EPC Network Toolkit
- **“Non Real-Time Location Systems”**
  - Object Localization using RFID readers mounted on autonomous vehicles
- **RFID Security and Privacy Issues**

Laboratory for Manufacturing and Productivity

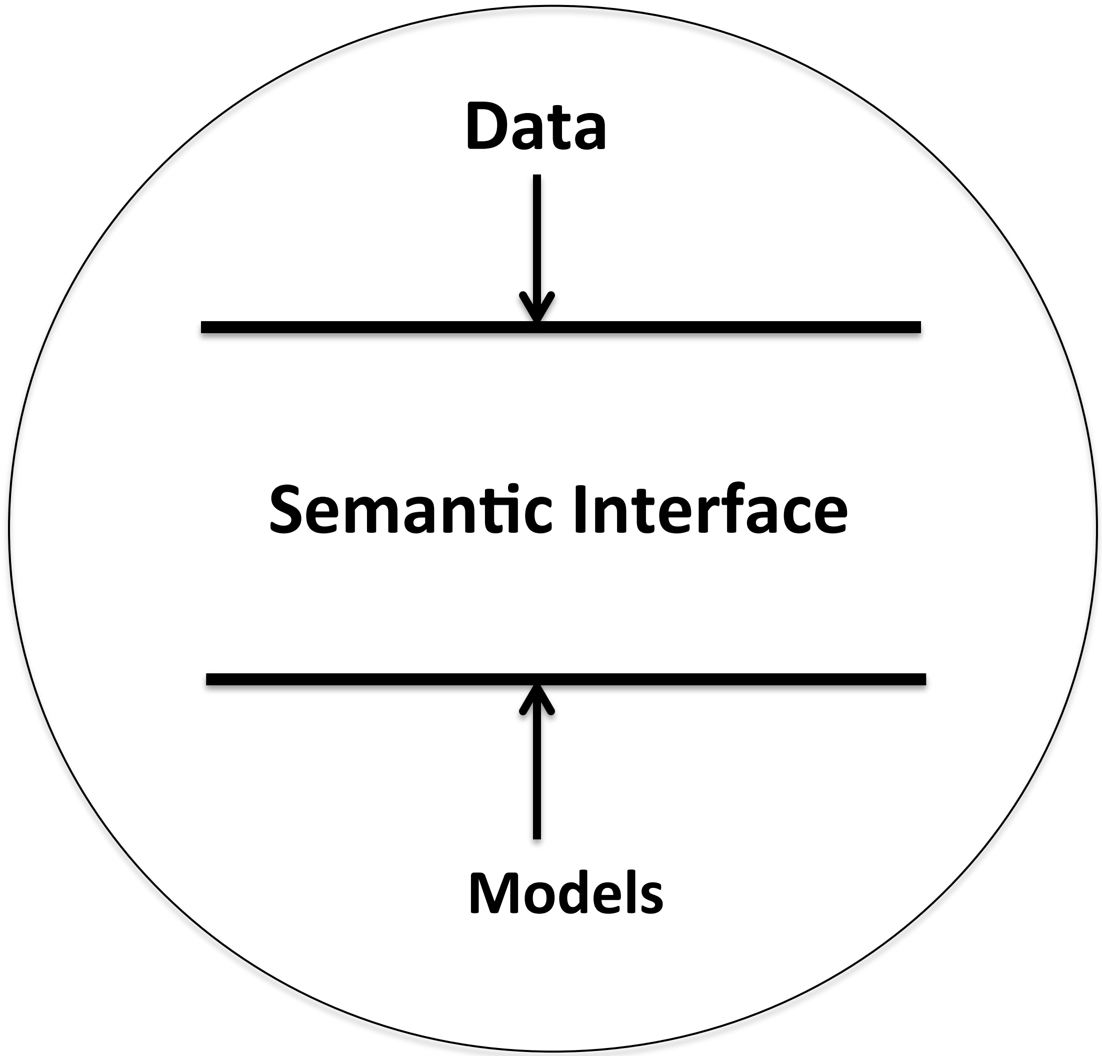
30 YEARS OF ENGINEERING THE REAL WORLD

# VISION


Eliminate the boundaries between the Internet and Enterprise computing.

# MATH MODELING (1998)

INTERNET



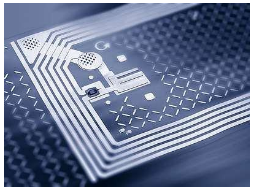
# The Danger of Incremental Thinking in Engineering

Space Station Design

Goodyear Aircraft Corporation, 1960’s

# INTERNET EVOLUTION

- **The Web of Information**
  - HTML, static web pages, www
- **The Web of Things**
  - Linking physical objects together, RFID
  - EPCglobal Network
- **The Web of Abstractions**
  - Interoperability, data and mathematical models
  - Computer languages and protocols for connections
  - Software as a Service (SaaS)

$$E(ACL) = \int_{L=-\infty}^{L=\infty} \left\{ \int_{H=RL}^{H=\infty} (H - RL) f(H) dH \right\} f(L) dL \quad (4)$$

# I. DATA AND MODELS

## Kratulos

## *Example:* Agricultural Data

# PROBLEM DEFINITION

Combine surface observation data of disease with temperature data from NOAA.

Both data sources are available via the Internet.

Form a set of data.

# SURFACE DATA

# AGRICULTURAL EXAMPLE

## Field Scouting

http://ingehygd.blogspot.com/2009/04/serveon-pda-demo.html

# DATA INTEGRATION

- **Two separate streams of data**
  - Observations from the field
  - Weather data from NOAA
- **Form an integrated data set for analysis**
  - Attach a logit model
  - Project disease growth rate
- **Weather data**
  - Point observation
  - Interpolation required

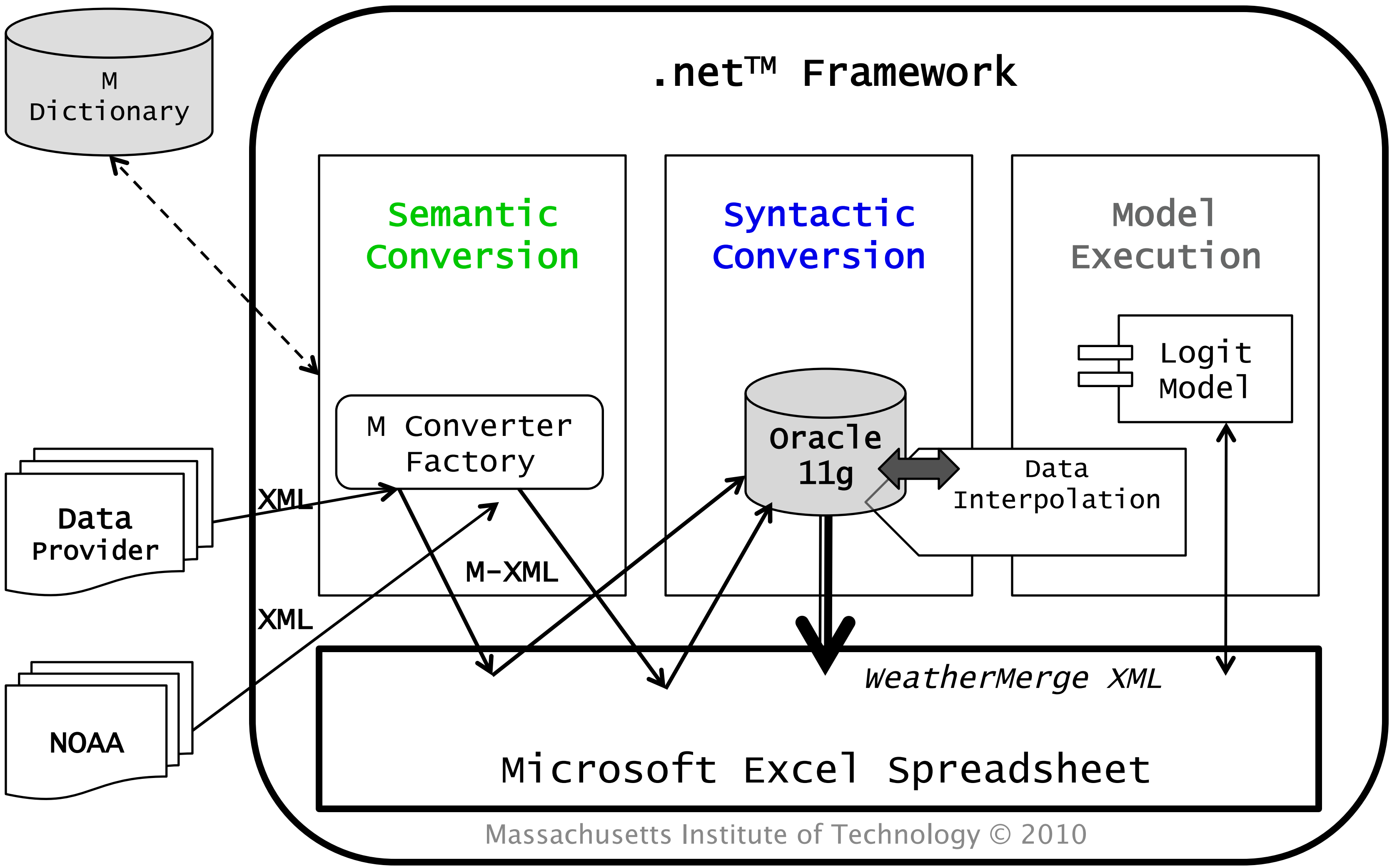
 

Kratulos is the underlying architecture and code

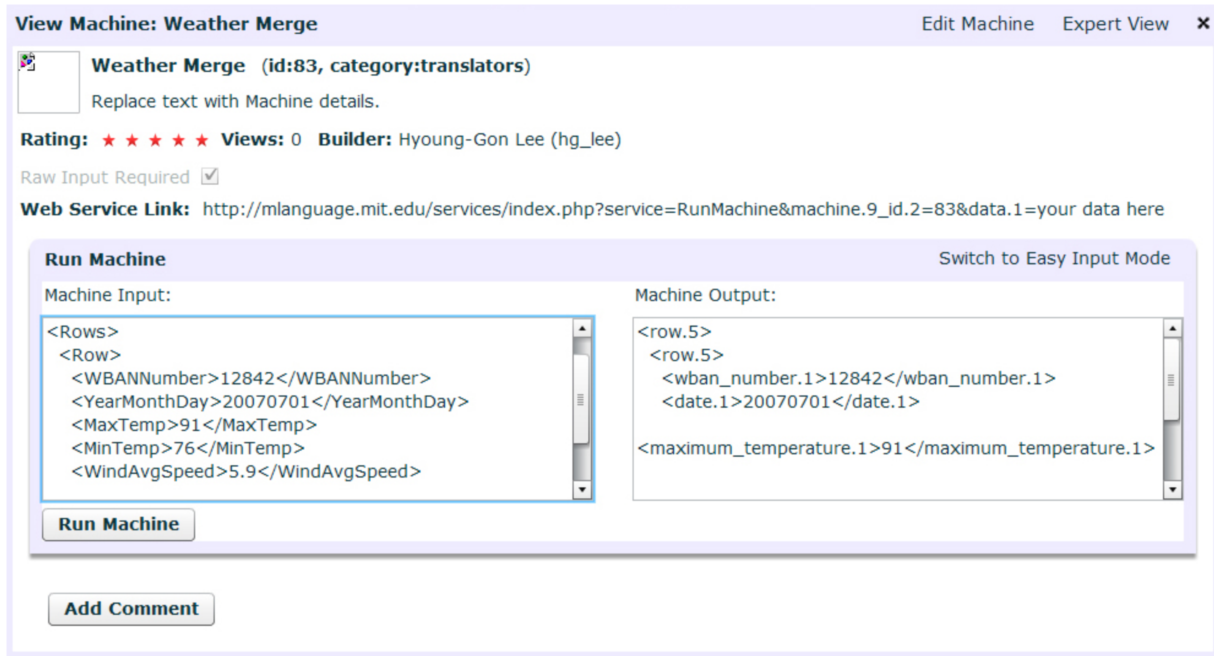
The core is Oracle 11g.

***Kratulos*** is a general approach for merging data and connecting models.

# KRATULOS


M
Dictionary
.net™ Framework
Semantic
Conversion
Syntactic
Conversion
Model
Execution
Logit
Model
M Converter
Factory
Oracle
11g
Data
Interpolation
Data
Provider
XML
M-XML
XML
NOAA
WeatherMerge XML
Microsoft Excel Spreadsheet

# SEMANTIC CONVERSION





```
- <Rows>
  - <Row>
      <WBANNumber>12842</WBANNumber>
      <YearMonthDay>20070701</YearMonthDay>
      <MaxTemp>91</MaxTemp>
      <MinTemp>76</MinTemp>
      <WindAvgSpeed>5.9</WindAvgSpeed>
      <PrecipitationWaterEquiv>0</PrecipitationWaterEquiv>
    </Row>
  </Rows>
```

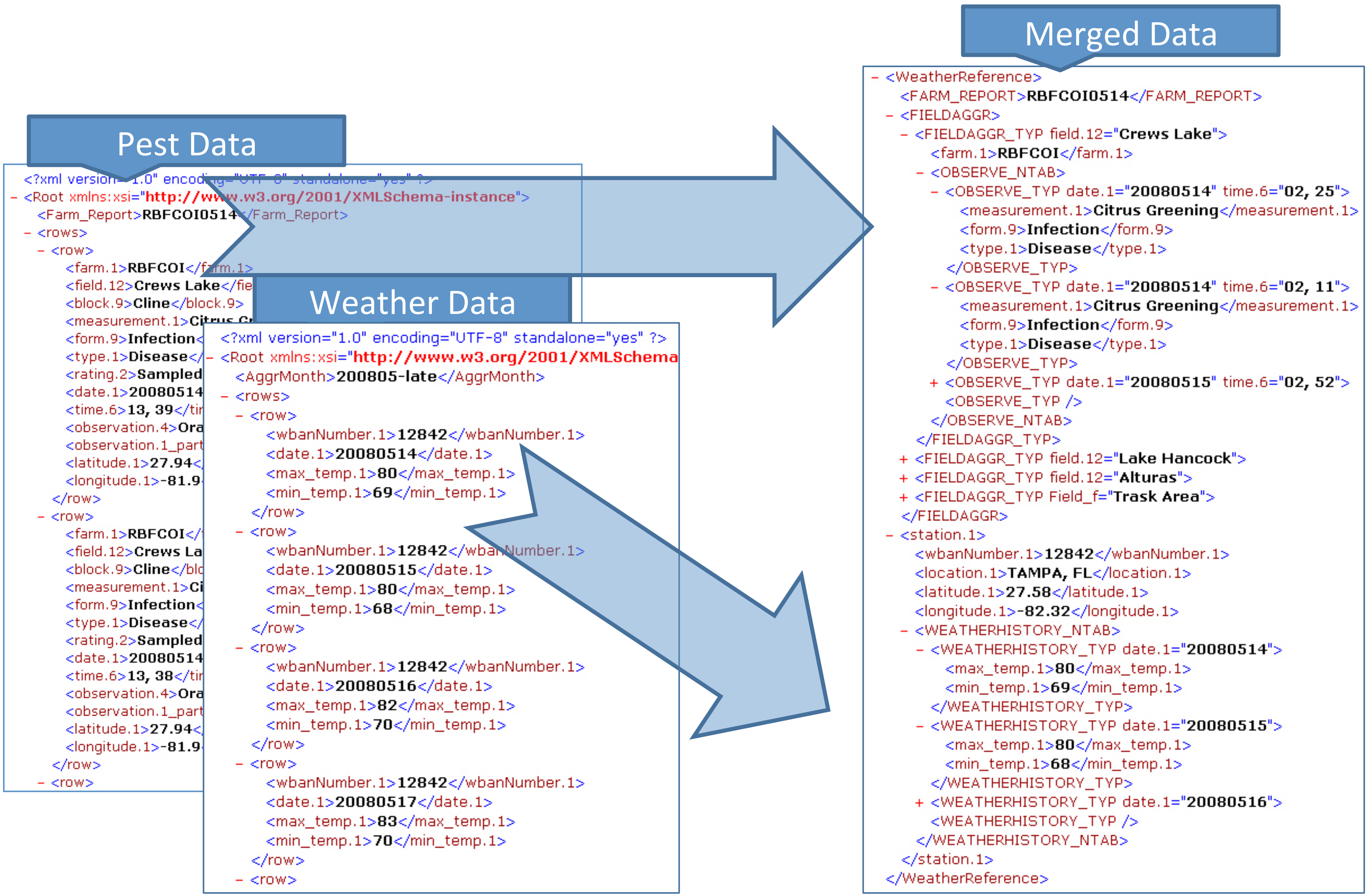


```
- <row.5>
  - <row.5>
      <WBAN_number>12842</WBAN_number>
      <date.1>20070701</date.1>
      <maximum_temperature.1>91</maximum_temperature.1>
      <mimimum_temperature.1>76</mimimum_temperature.1>
      <average_wind_speed.1>5.9</average_wind_speed.1>
      <precipitation.4>0</precipitation.4>
    </row.5>
  </row.5>
```

Laboratory for Manufacturing and Productivity

30 YEARS OF ENGINEERING THE REAL WORLD

# M-XML



Laboratory for Manufacturing and Productivity

30 YEARS OF ENGINEERING THE REAL WORLD

# MERGED DATA – EXCEL

WeatherMerge.xlsm - Microsoft Excel non-commercial use

| | Pest Data X Provider | | | | | | | Weather Data NOAA | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | field_name.1 | farm_name.1 | event_date.1 | pest_common_name.1 | pest_form_name.1 | pest_type_name.1 | date.1 | location.1 | latitude.1 | longitude.1 | max_temperature.1 | min_temperature.1 | averag |
| 3 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 4 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 5 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 6 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 7 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 8 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 9 | Lake Hancock | RBFCOI | 20070701 | Citrus Greening | Infection | Disease | 20070701 | TAMPA, FL | 27.58 | -82.32 | 91 | 76 | |
| 10 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 11 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 12 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 13 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 14 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 15 | Lake Hancock | RBFCOI | 20070702 | Citrus Greening | Infection | Disease | 20070702 | TAMPA, FL | 27.58 | -82.32 | 90 | 74 | |
| 16 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 17 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 18 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 19 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 20 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 21 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 22 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 23 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 24 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 25 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 26 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 27 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 28 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |
| 29 | Lake Hancock | RBFCOI | 20070703 | Citrus Greening | Infection | Disease | 20070703 | TAMPA, FL | 27.58 | -82.32 | 90 | 75 | |

Click on a word and the exact definition appears as a pop-up.

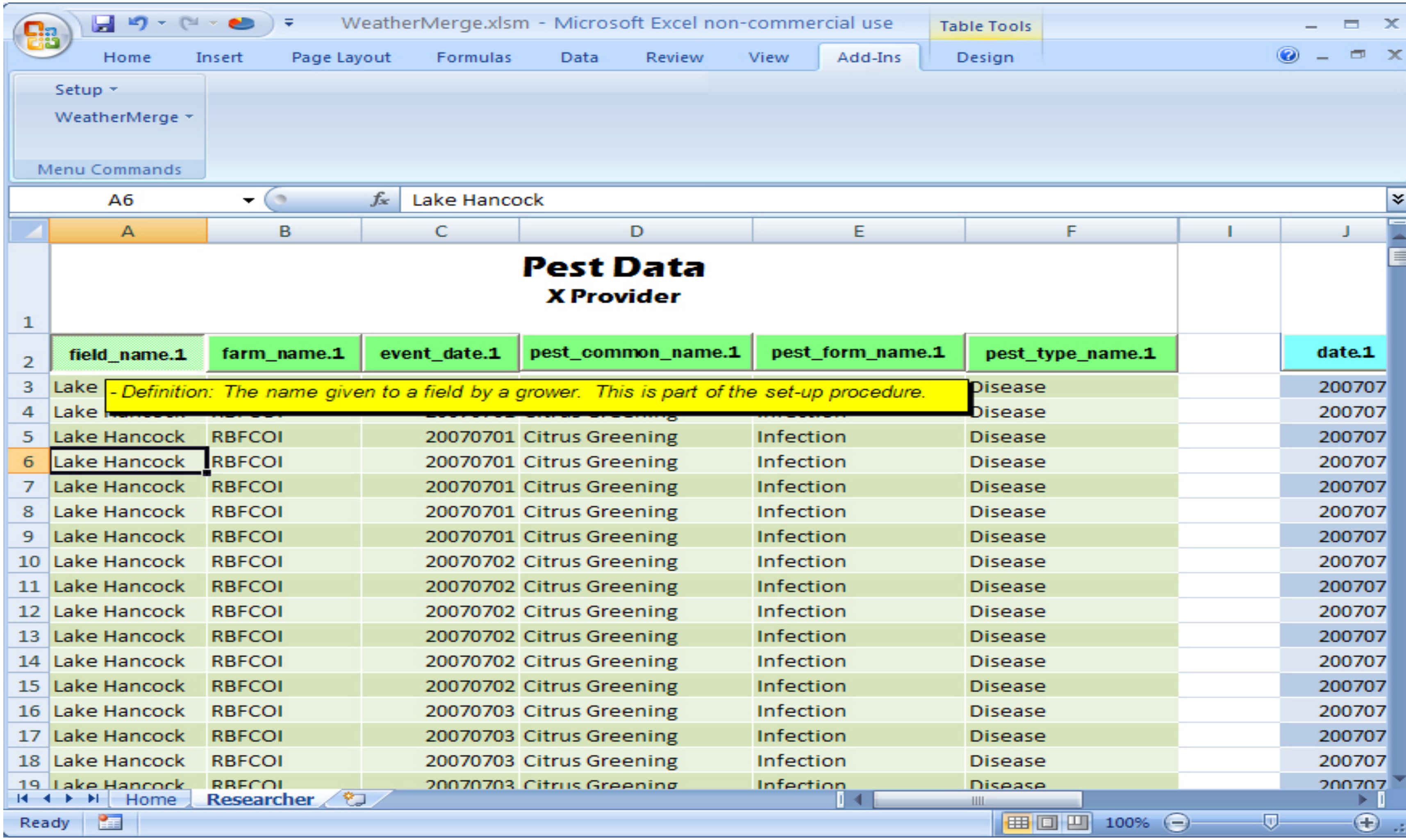
On the following slide, `field_name.1` appears as an embedded word.

It is linked directly to the M Dictionary, located on a remote server.

Laboratory for Manufacturing and Productivity

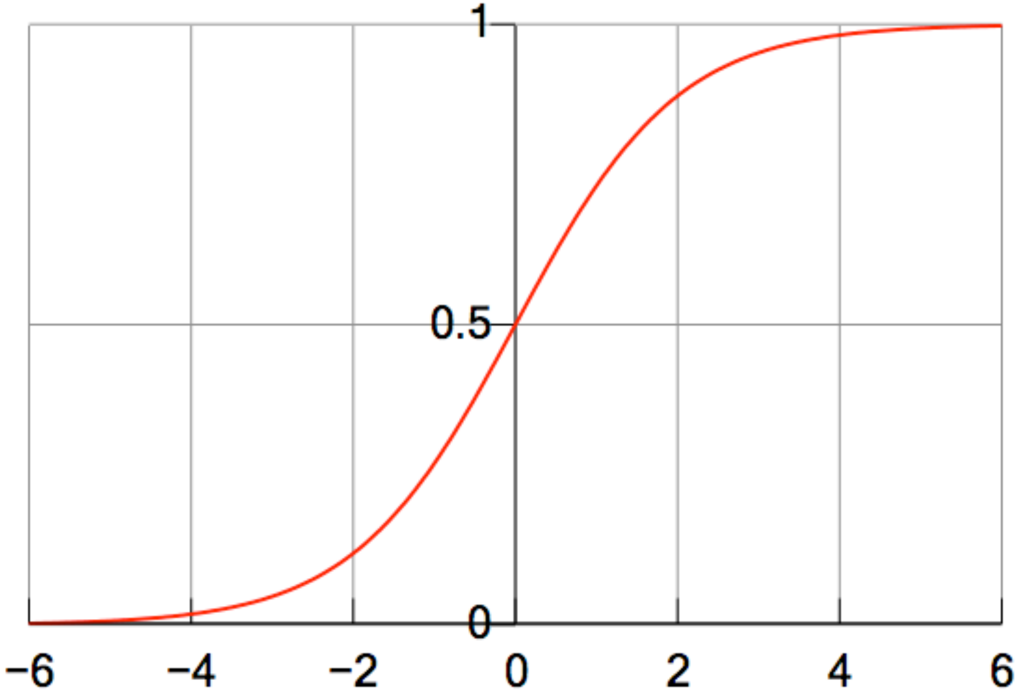
30 YEARS OF ENGINEERING THE REAL WORLD

# WORD DEFINITION



# LOGIT MODEL

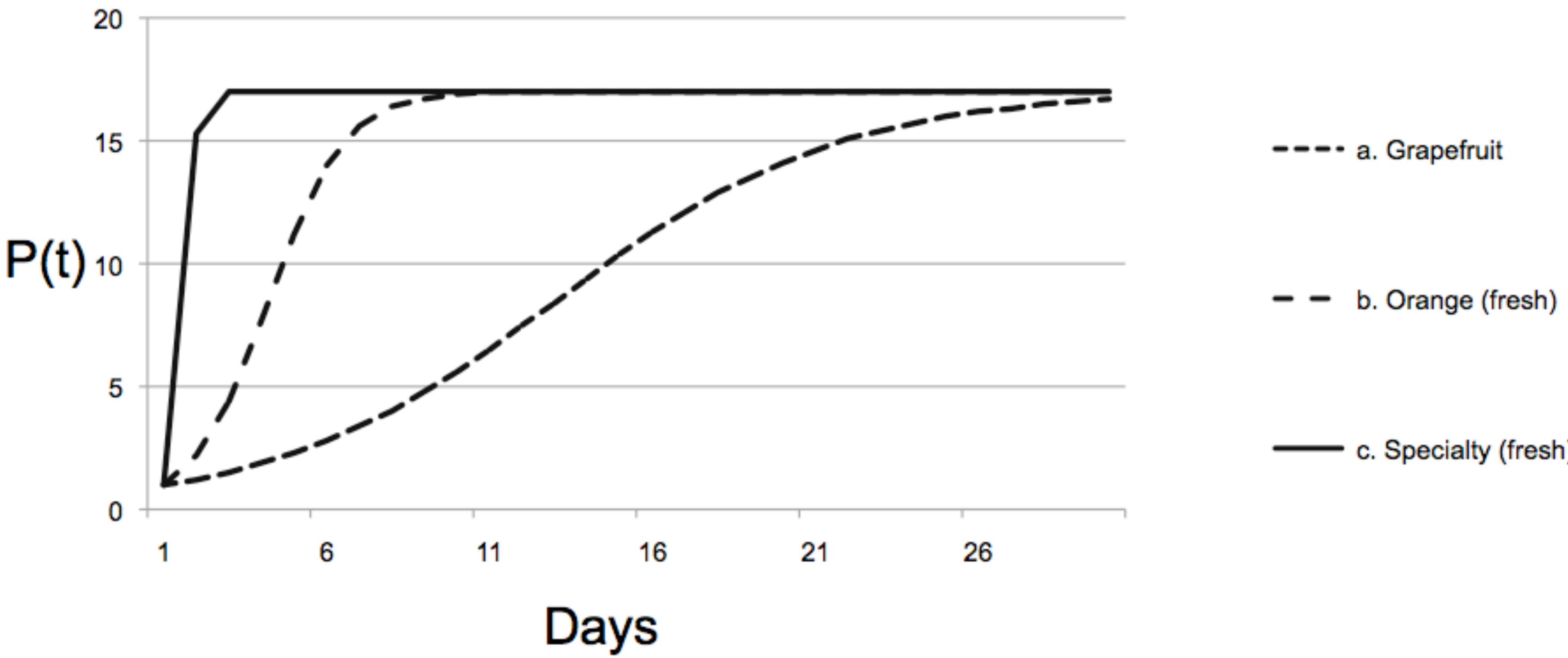
$$P(t) = 1/(1+e^{-t})$$



P(t) = K/(1+(B*Exp(-(GR) * t)))

Where:

K = maximum environmental capacity
GR = the modified Allen growth rate
B = initial pest density at t=0


20
15
10
5
0
P(t)
1
6
11
16
21
26
Days
a. Grapefruit
b. Orange (fresh)
c. Specialty (fresh)

# SUMMARY

- **New approach for projecting disease growth in agriculture**
  - Weather/surface observation data set created instantaneously
- **WeatherMerge™ intended as a form of ERP for agriculture**
- **Long-term, replace human scout with robot**
- **Broader applications in ecosystems services**

# II. INTERNET AND THE ENTERPRISE

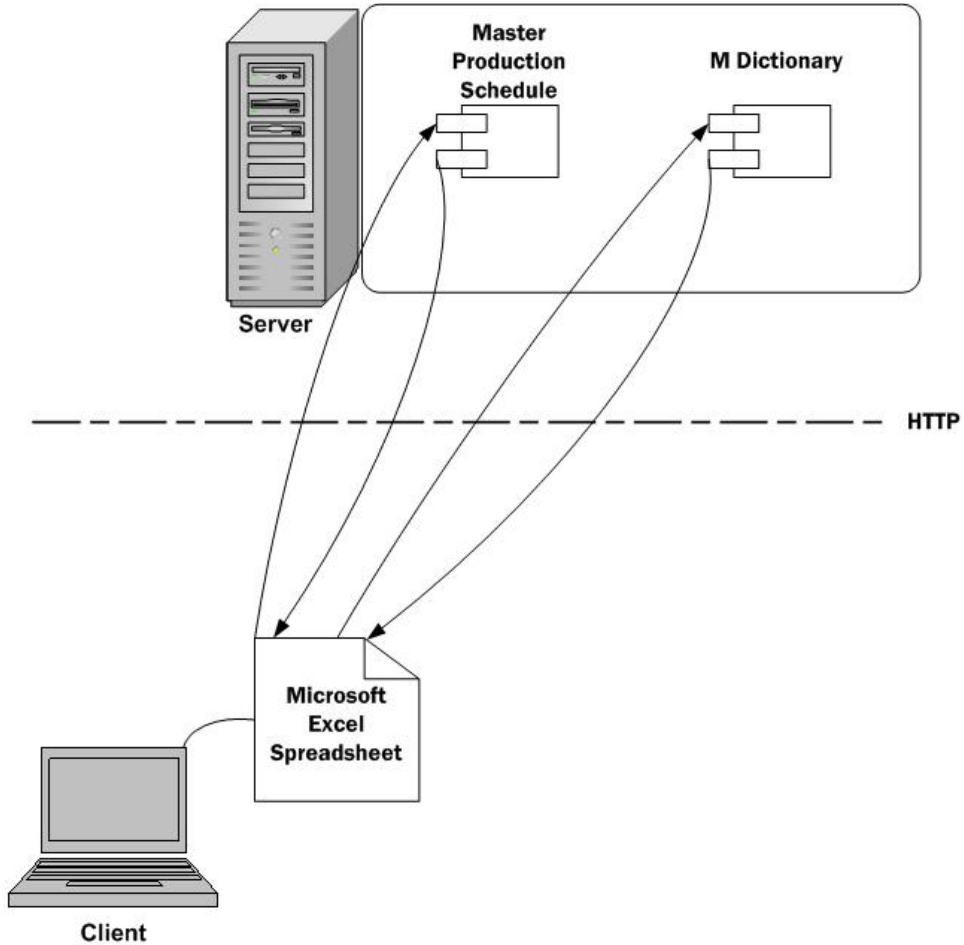
Lee-Schuster Semantic Enterprise Architecture

*Example:* The Open System for Master Production Scheduling (OSMPS)

# HIGHLIGHTS

- ***Open Systems***
  - Open source versus open systems
  - Powerful trend in the computer industry, replacement for ERP software
- **M Dictionary and other web standards**
  - mlanguage.mit.edu
- **Software as a Service**
  - Access a sophisticated model on a remote server using a spreadsheet interface that can reside on any microcomputer with Internet link
  - Match a specific model to a specific problem
  - In contrast to *Kratulos*, no database
- **No implementation of model on local computer, access is immediate, simple interface**

# ARCHITECTURE



***Internet* Based**

- **Boundary**
  - Concerns about security
  - M Dictionary becomes complex
  - Loss of control over words and noun phrases
  - Business model

***Enterprise* Based**

- **Intra firm IT connections**
  - Machine understandable
  - Quick and low cost
  - Flexible
  - Semantic intranet

# Semantic Web 3.0

- **Connections**
  - RFD and OWL
- **Reasoning**
  - Use Semantic web trees
- **Discovery**
  - Relationship between knowledge elements
- **Other**

***Limitation: tree semantics***

# Lee-Schuster Semantic Enterprise Architecture

Intranet (firm) or Internet

Example:
Open System for Master Production Scheduling (OSMPS)

*Overcomes semantic ambiguity*

A need exists for low cost scheduling software.

Typical ERP scheduling packages cost more than $100,000 including installation.

# THE PROCESS INDUSTRIES

- **Manufacturing operations that include:**
  - High speed manufacturing lines (example, bottle filling)
- **Major industries**
  - Food
  - Chemical
  - Pharmaceutical
  - Paper
  - Biotechnology
- **The sector is probably 50% of worldwide manufacturing**

# TRAITS

- **Wide range of environments**
  - Continuous vs. batch process
- **No universal solutions**
  - Wide differences between different segments
- **High customer service expectations**
  - 99% percent cases ordered vs. cases shipped
- **Dynamic demand**
  - Sudden changes in demand
  - Characterized as “lumpy”

# OTHER

- **Repetitive manufacturing**
  - Production of similar products using a common machine
  - Example; metal parts such as screws
  - High volume
  - Fixed capacity; multiple products
- **Non manufacturing**
  - Agriculture, defense, publishing, services
  - Delivery of mathematical models

# MODIFIED DIXON SILVER

- Make-to-stock manufacturing environment; no stock-outs or backorders permitted
- Multi-item, single level, dedicated production lines with finite capacity
- Setup times and cost are nonzero and sequence independent
- Safety stocks (buffers) are determined "outside" of the scheduling system
- MODS

# MANUFACTURING SYSTEM

- Enterprise Resource Planning (ERP); exclusive system for decision making in manufacturing firms
- Critical in asset management, highly complex
- Important role in delivering customer service
- ERP is a large data base combined with models
- Delivery by packaged software
  - Maturing technology

# PRODUCTION DECISIONS

Forecasts of Demand

Aggregate Plan

**Master Production Schedule**
**Schedule of Production Quantities by production and time period**

Materials Requirements Planning System
Explore master production schedule to obtain requirements for components

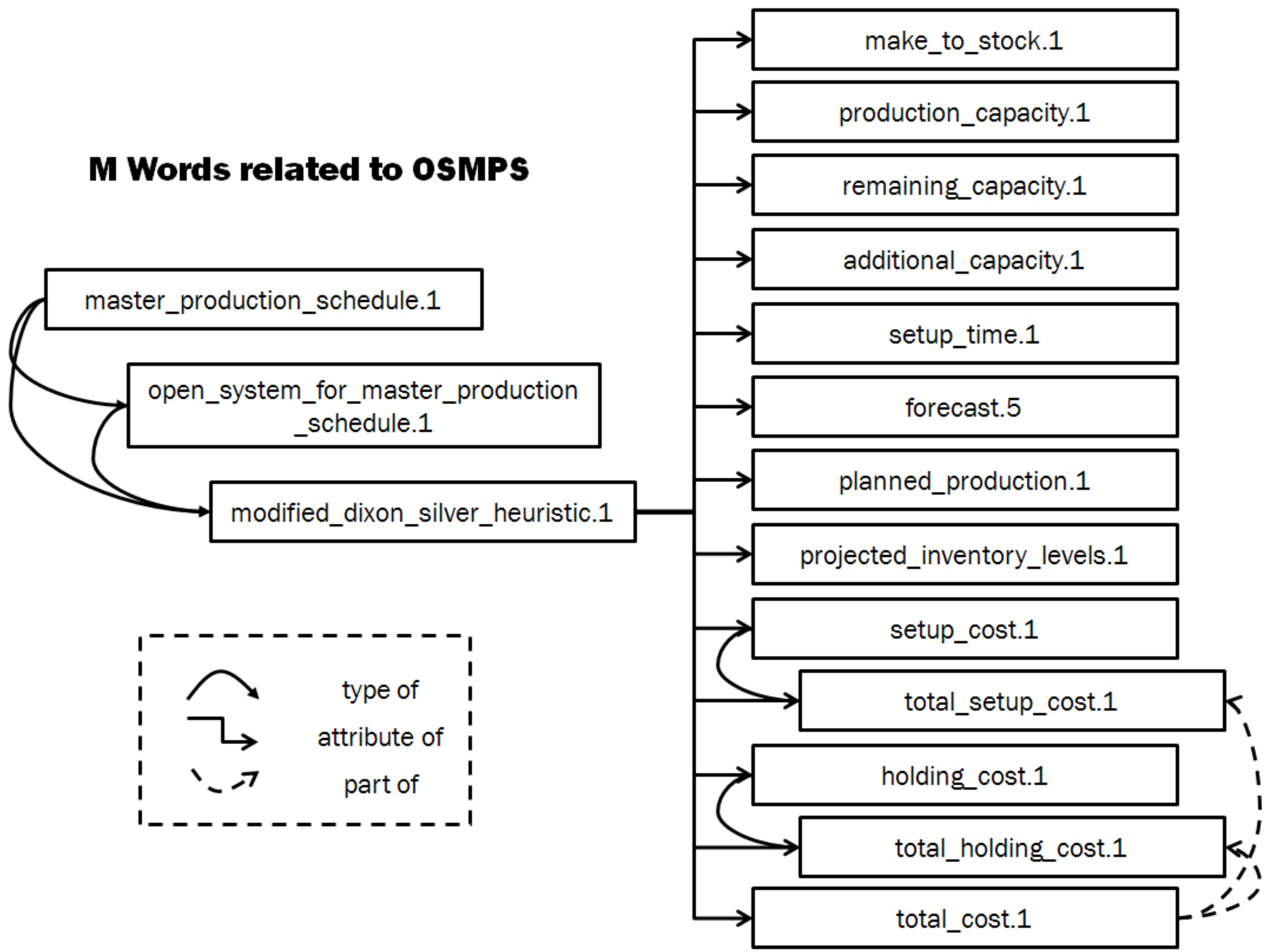
Detailed Job Shop Schedule
To meet specification of production quantities from the MRP system

Adapted from: Nahmias, S. (1993), *Production and Operations Analysis*, New York: Irwin.

# OSMPS – TWO PARTS

- **The model**
  - MODS
  - Highly sophisticated approach for make-to-stock production scheduling
  - Very fast calculation speed (seconds) as compared to traditional OR approaches
  - Comprehensive and robust, a candidate for a world standard
- **The delivery**
  - Excel spreadsheet interface
  - Use M Dictionary
  - New approach to Internet and Intranet programming

# OSMPS – ONTOLOGY



The ontology is contained within the M Dictionary.

Laboratory for Manufacturing and Productivity

30 YEARS OF ENGINEERING THE REAL WORLD

# EXCEL SPREADSHEET

Steps:

1. Enter data in <u>yellow</u> fields including **production_capacity.1**, **forecast.5**, **capacity_absorbed.1**, **holdi**
2. (optional) Check the meaning of the words by clicking on a word; re-click to close definition box.
3. Please make sure that macros are enabled. The spreadsheet will automatically interact with the MIT serve
4. Click on Run OSMPS.
5. Check the results (model outputs) by looking at **planned_production.1**, **projected_inventory_levels.**

<u>Note:</u> Forecast.5 is netted for beginning inventory. This version DOES NOT contain safety stock.

| | |
|---|---|
| total_holding_cost.1 | $65,316 |
| total_setup_cost.1 | $274,200 |
| total_cost.1 | [illegible] |

**Run OSMPS!!**

*- Definition: The capability of a system to perform its expected function. 2) The capability of a worker, machine, work center, plant, or organization to produce output per time period. Capacity required represents the system capability needed to make a given product mix (assuming technology, product specification, etc.). As a planning function, both capacity available and capacity required can be measured in the short term (capacity requirements plan), intermediate term (rough-cut capacity plan), and long term (resource requirements plan). Capacity control is the execution through the I/O control report of the short-term plan. Capacity can be classified as budgeted, dedicated, demonstrated, productive, protective, rated, safety, standing, or theoretical.*

*- Attribute of: modified_dixon_silver_heuristic.1*

| | 8 | 9 | 10 | 11 |
|---|---|---|---|---|
| production_capacity.1 | 100 | 100 | 100 | 100 |
| remaining_capacity.1 | 2 | 0 | 0 | 24 |
| additional_capacity.1 | 0 | 0 | 0 | 0 |

| | ITEM # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| forecast.5 | 154 | 0 | 0 | 0 | 6.3 | 6.3 | 6.8775 | 6.8775 | 6.8775 | 6.8775 | 8.841 | 8.841 |
| | 155 | 0 | 5.88 | 5.88 | 5.88 | 5.88 | 6.4575 | 6.4575 | 6.4575 | 6.4575 | 8.82 | 8.82 |
| | 156 | 17.052 | 4.5675 | 4.5675 | 4.5675 | 4.5675 | 4.9875 | 4.9875 | 4.9875 | 4.9875 | 4.557 | 4.557 |
| | 157 | 35.28 | 9.555 | 9.555 | 9.555 | 9.555 | 10.21125 | 10.21125 | 10.21125 | 10.21125 | 6.678 | 6.678 |

(Period #)

# INPUTS

- `Forecast.5`: **by item: the demand for each period netted for beginning inventory, by item (cell C35 to BB36 – anticipated units sold per week).**

- `Production_capacity.1`: **units of capacity available (cell C29 to BB29 – total hours available for the manufacturing line or machine)**

- `Capacity_absorbed.1`: **units of capacity required for production, by item (BH35 to BH66 – hours to produce 1,000 units)**

# INPUTS (CONTINUED)

- `Holding_cost.1`: **the cost of holding inventory, by item (BJ35 to BJ66 – Dollars per 1000 units per month)**

- `Setup_cost.1`: **the cost of a setup, by item (BL35 to BL66 – Dollars per setup)**

- `Setup_time.1`: **the time to setup, by item (BN35 to BN66 – hours per setup**

# OUTPUTS

- `Remaining_capacity.1`: **the amount of surplus capacity per week (C30 – BB30, hours)**

- `Additional_capacity.1`: **the amount of capacity needed over standard capacity (C31 – BB31, hours)**

  **NOTE: the MODS algorithm makes every effort to fit production into available capacity, however, sometimes an over capacity situation exists.**

- `Planned_production.1:` **the production schedule by week (C35 – BB66, units per week by item)**

# OUTPUTS (CONTINUED)

- `Projected_Inventory_Levels.1`: **the amount of inventory remaining at the end of each week (C105 – BB136, units per week by item)**
- `Total_holding_cost.1`: **the sum of the holding cost for the 52 week period, Dollars**
- `Total_setup_cost.1`: **the sum of the setup cost for the 52 week period, Dollars**
- `Total_cost.1`: **total holding cost plus total setup cost, Dollars**

# SMALL SCALE EXAMPLE

| | |
|---|---|
| total_holding_cost.1 | $65,316 |
| total_setup_cost.1 | $274,200 |
| total_cost.1 | $339,516 |

Run OSMPS!!

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| production_capacity.1 | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| remaining_capacity.1 | 20 | 43 | 21 | 11 | 18 | 7 | 3 | 2 |
| additional_capacity.1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

| | ITEM # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | capacity_absorbed.1 | holding_cost.1 | setup_cost.1 | setup_time.1 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| forecast.5 | 154 | 0.00 | 0.00 | 0.00 | 6.30 | 6.30 | 6.88 | 6.88 | 6.88 | 0.44 | $50 | $400 | 1 |
| | 155 | 0.00 | 5.88 | 5.88 | 5.88 | 5.88 | 6.46 | 6.46 | 6.46 | 0.44 | $50 | $400 | 1 |
| | 156 | 17.05 | 4.57 | 4.57 | 4.57 | 4.57 | 4.99 | 4.99 | 4.99 | 0.56 | $50 | $200 | 1 |
| | 157 | 35.28 | 9.56 | 9.56 | 9.56 | 9.56 | 10.21 | 10.21 | 10.21 | 0.56 | $50 | $200 | 1 |

| | ITEM # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| planned_production.1 | 154 | 0.00 | 0.00 | 0.00 | 12.60 | 0.00 | 13.76 | 0.00 | 6.88 |
| | 155 | 0.00 | 11.76 | 0.00 | 11.76 | 0.00 | 12.91 | 0.00 | 12.91 |
| | 156 | 21.62 | 0.00 | 9.14 | 0.00 | 9.56 | 0.00 | 9.98 | 0.00 |
| | 157 | 35.28 | 9.56 | 19.11 | 0.00 | 19.77 | 0.00 | 10.21 | 10.21 |

| | ITEM # | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|---|
| projected_inventory_-levels.1 | 154 | 0.00 | 0.00 | 0.00 | 6.30 | 0.00 | 6.88 | 0.00 | 0.00 |
| | 155 | 0.00 | 5.88 | 0.00 | 5.88 | 0.00 | 6.46 | 0.00 | 6.46 |
| | 156 | 4.57 | 0.00 | 4.57 | 0.00 | 4.99 | 0.00 | 4.99 | 0.00 |
| | 157 | 0.00 | 0.00 | 9.56 | 0.00 | 10.21 | 0.00 | 0.00 | 0.00 |

# MICROSOFT ISSUES

- **OSMPS requires MSXML 6.0 Service Pack 1 from Microsoft.**
- **A firewall might block the connection to MIT**
- **This spreadsheet does not work in Microsoft Excel for Apple computers**

# III. INTRODUCTION

## Details of the Theory and Technology

# DATA

Each year, the amount of data grows by as much as **40 – 60** % for many organizations.

In 2004 alone, shipments of data storage devices equaled four times the space needed to store every word ever spoken during the entire course of human history.

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

Lyons, Daniel (2004), "Too Much Data," *Forbes*, December 13.

# BUSINESS PROBLEM

“data, data everywhere but not a byte to use.”

Sunil Gupta of SAP paraphrasing Samuel Taylor Coleridge during *Smart World 2004*, sponsored by the MIT Industrial Liaison Program.

# THESIS

Industry needs a new form of organization for data to speed search and make connections quickly.

Models are the means of analyzing data.

Brock, David L., Edmund W. Schuster, Stuart J. Allen, and Pinaki Kar (2005), “An Introduction to Semantic Modeling for Logistical Systems,” *Journal of Business Logistics* 26:2, pp. 97 – 117.

# PROOF

Research, design, and implement a system for data and model integration that solves practical problems.

The focus is manufacturing, agriculture, defense, and other industries.

# FUNDAMENTAL PROBLEM

"One cannot step twice into the same stream …not even once."

Heraclitus and Cratylus (Kratulos), late 5th century BC

Words are the basis for anything intellectual.

The Greeks thought it impossible to have a system of logic because the meaning of words constantly change.

# SEMANTIC AMBIGUITY

- **A single word has several different meanings**
- **Difficult to achieve "machine understandable" semantics**
- **The next two slides show examples from my blog for a search on "apple."**

# EXAMPLE 1 – HTML

## Entry Title: “Apple Chip Design Plans”

```
<div><span class="Apple-style-span"  style="font-
family:verdana;”></span></div><div><span class="Apple-
style-span”style="font-family:verdana;">Several years ago,
I worked with a group of students from EECS on a project to
apply ultra low power circuit technology to RFID tags…</
span></div><div><span class="Apple-style-span"
style="font-family:verdana;”></span></div><div><span
class="Apple-style-span"  style="font-family:verdana;”>It
appears one of the reasons that Apple is doing chip design
internally relates to the desire for technology that
reduces iPhone power consumption.  I very much believe this
is a good idea.  MIT is a leader in this type of
technology.
</span></div>
```

Laboratory for Manufacturing and Productivity

30 YEARS OF ENGINEERING THE REAL WORLD

# EXAMPLE 2 – HTML

## Entry Title: “My Mom and May”

```
<div><span class="Apple-style-span"  style="font-
family:verdana;”><span class="Apple-style-span"
style="font-size:medium;”></span></span></div><div>
<span class="Apple-style-span"  style="font-family:verdana;
"><span class="Apple-style-span"  style="font-size:medium;”
>It was a general farm including grapes, dairy, raspberries,
vegetables, stone fruits, apples, and grain crops.
```

# EXAMPLE (CONTINUED)

- The two blog posts have no relationship, yet both appear as search results for "apple"
- HTML code is specific; content contains ambiguous words
- For XML, no uniform semantics or syntax exists
  - Limit to machine understanding
- "Leading analysts have estimated that 35-65% of system integration costs are because of semantic issues*

*2006 Semantic Technology Conference, San Jose, CA.

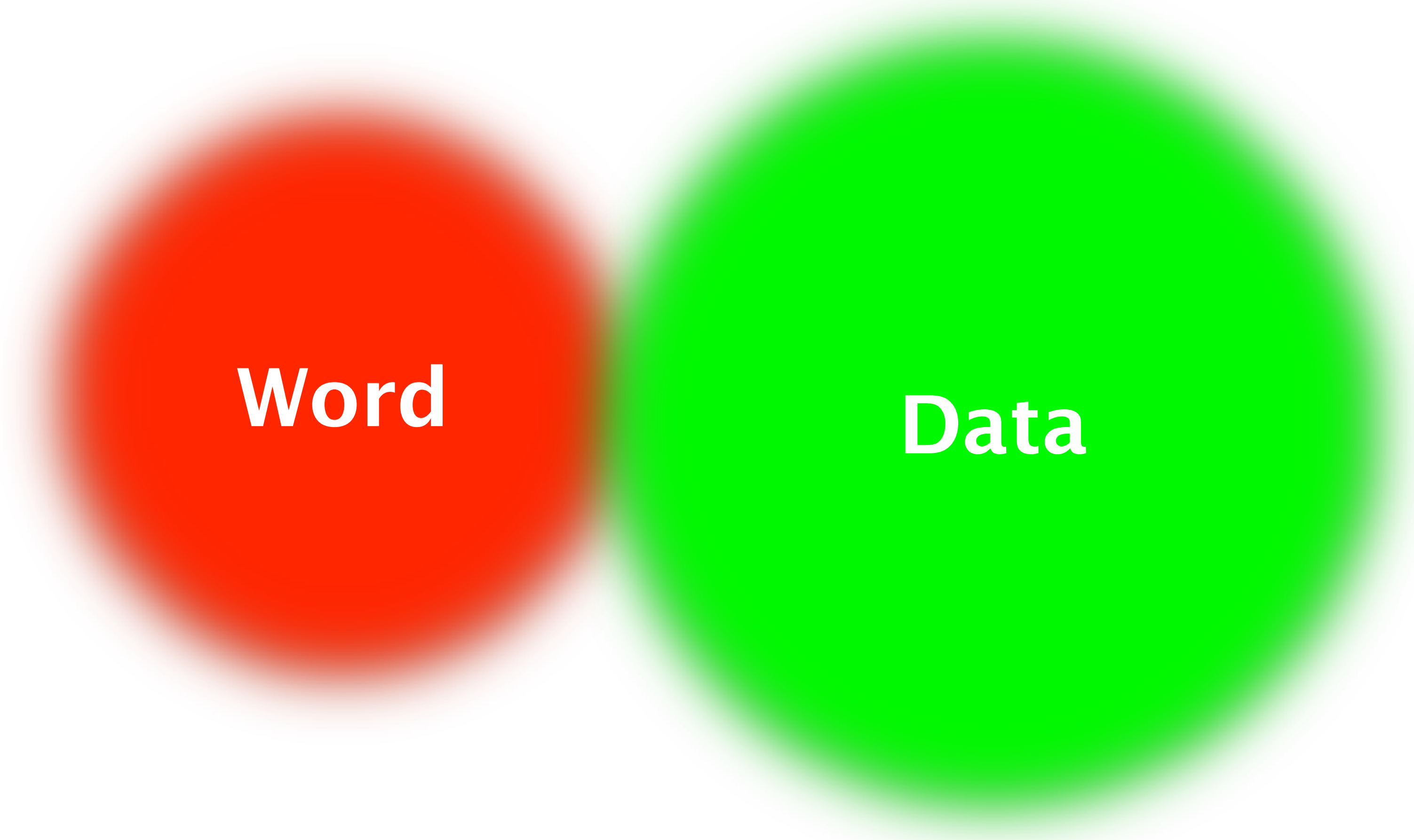
# RESEARCH GOALS

- **Solve the issue of semantics and syntax for XML**
- **Achieve interoperability for data and mathematical models**
- **Create an auxiliary language to integrate models/data**
- **Apply to industry**

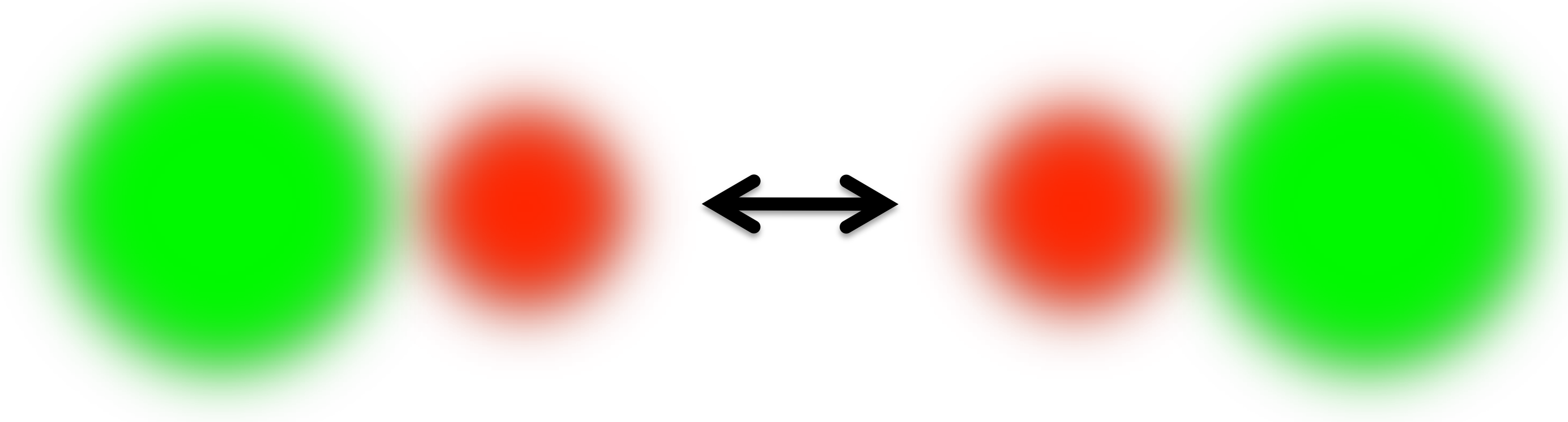
# PHILOSOPHY

- ***Integrate*** **IT standards with innovations**
  - XML, Web Services, legacy data, and M Dictionary
- ***Inductive*, examine a specific technical problem then solve**
  - Industry orientation
- ***Separate*** **computer code from Enterprise package software**
  - User could be anywhere
  - Connect through a simple interface
- ***Replication*, get model to exact application**
  - Low implementation cost
  - Remove the friction of modeling

# DATA MOLECULE



# INTERNET CONNECTION



Internet transfer of data using XML requires prior agreement on *semantics* and *syntax* between the sender and the receiver of the data.

This is a major limitation for XML.

No universal standard exists, resulting in many forms of XML.

# STANDARDS

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

Adapted from D.L. Brock

# SPECIALIZATION

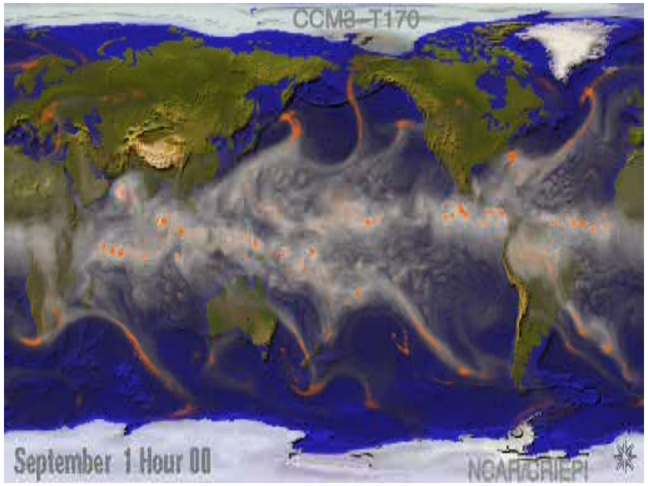
Multiple standards make it difficult to *merge* data.

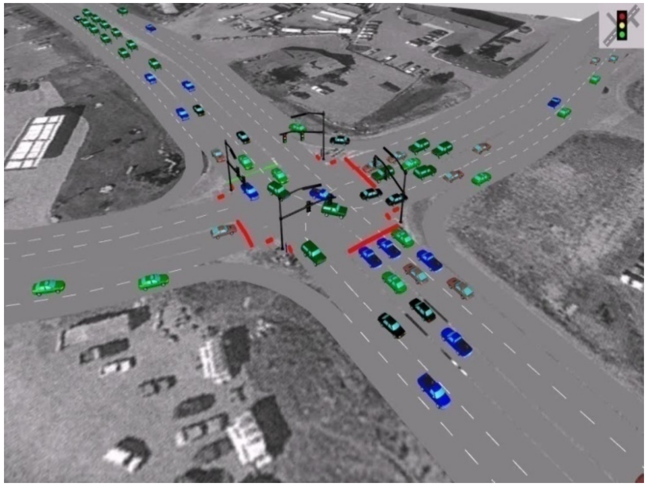
*Versioning* occurs within the same standard.

Over time, versioning is a significant problem.

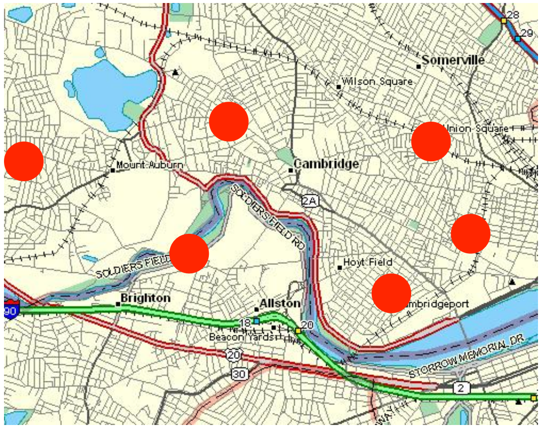
# BOSTON COMMUTE

**Live Traffic Data**



**Live Weather Data**




**Construction Location**

Adapted from D.L. Brock

# DATA INTEGRATION

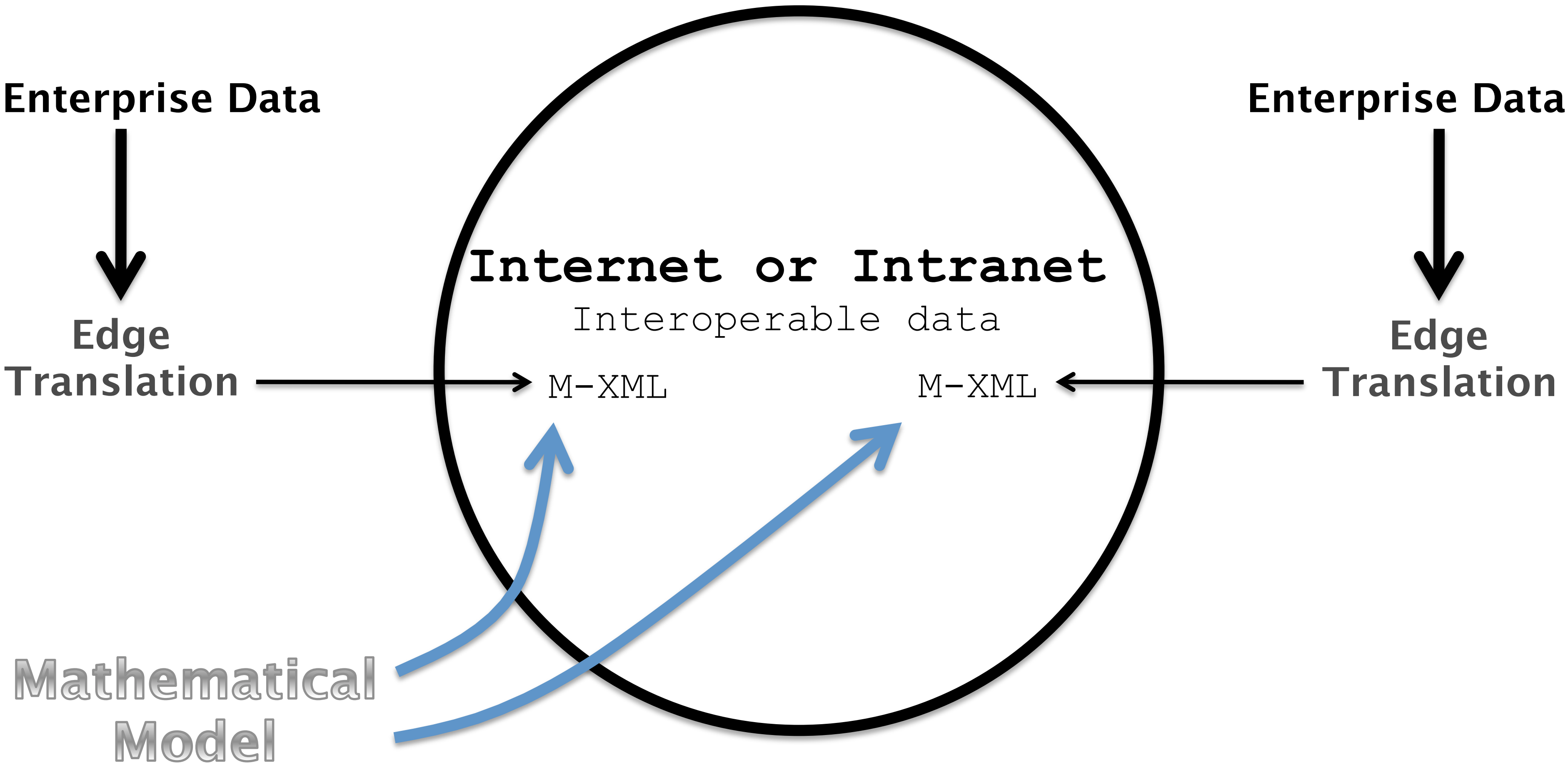
Merging XML data requires a “hub translator.”

This is a non real-time process.

The number of “many to many” combinations is polynomial, as a function of the number of nodes.

An auxiliary language reduces the combinations.

# AUXILIARY LANGUAGE



# PURPOSE

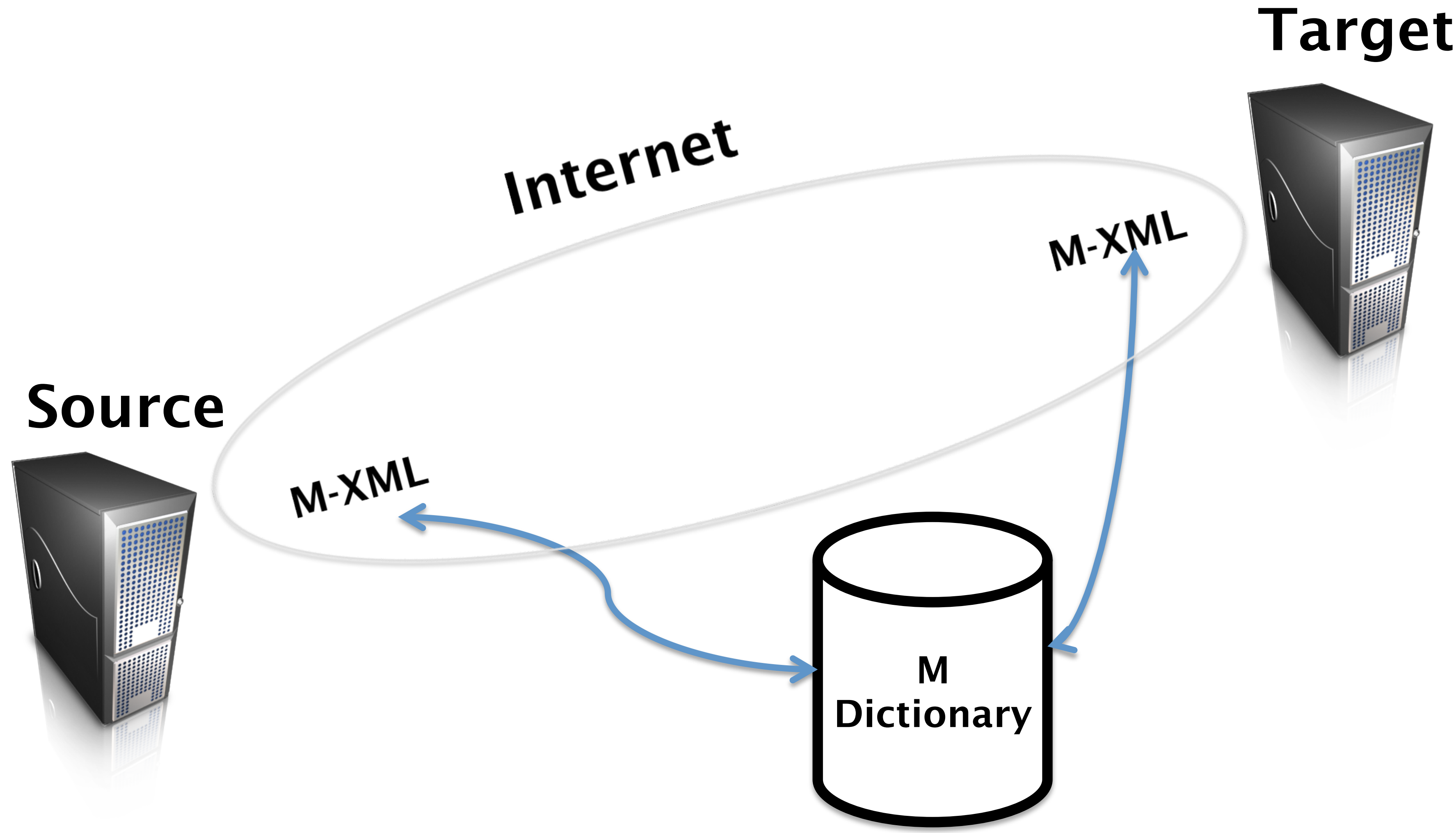
An auxiliary language is the glue that holds things together.

It is not a formal code like Java or C++.

The purpose is to make XML more effective.

Laboratory for Manufacturing and Productivity
30 YEARS OF ENGINEERING THE REAL WORLD

# INTERNET ARCHITECTURE



# ADVANTAGES

- **Translation to M-XML**
  - Achieve communication when the target is unknown
- ***Intelligent Data* that self identifies, much like RFID**
- **Addresses the "many to many" problem**
  - No need for a hub translator
- **A standard way to describe and connect models**
  - Web Services and Software as a Service (SaaS)
- **The *M Dictionary* represents an innovation**

# WORDS AND SEMANTICS

Cell n. – a manufacturing cell, in which a group of workers and/or machines work together as a team to produce dedicated set of products or assemblies.

Cell n. – usually microscopic structure containing nuclear and cytoplasmic material enclosed by a semi-permeable membrane and, in plants, a cell wall; the basic structural unit of all organisms.

# M DICTIONARY

date.1 n. – particular day specified as the time something happens. July 4, 1776 was the date of the signing of the Declaration of Independence.

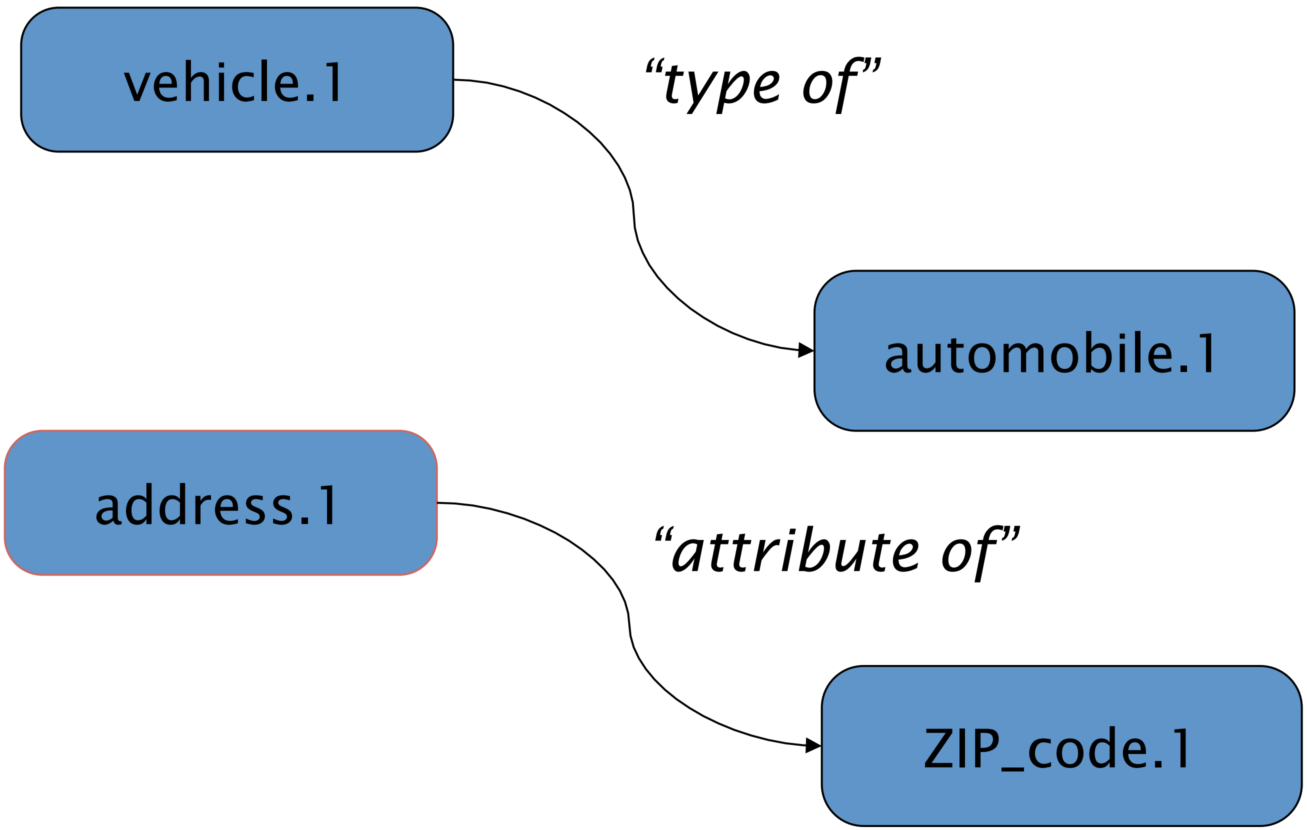
***Data Format***

**ISO 8601 (string) –** the international standard for date and time issued by the International Organization for Standardization (ISO).

**pattern: ([0-9]{4})(-([0-9]{2})(-([0-9]{2})(T([0-9]{2}):([0-9]{2})(:([0-9]{2})(\.([0-9]+))?)?(Z|(([-+])([0-9]{2}):([0-9]{2})))?)?)?)?**

# ONTOLOGY

## Relationships between words



# EXAMPLE

**car.1**, noun. A motor vehicle with four wheels; usually propelled by an internal combustion engine.

**"PART OF"**

Air_bag.1

Car_door.1

Car_mirror.1

→ Car.1 →

**"TYPE OF"**

Motor_vihicle.1

Automotive_vihicle.1

# INNOVATIONS

- **One definition per word**
- **Relationships between words (ontology)**
- **mlanguage.mit.edu**
- **Web Services connection**
  - `GetWord`
  - `TestRelation`
  - `Other connections also available`
  - `The dictionary becomes part of the Internet`

Laboratory for Manufacturing and Productivity

30 YEARS OF ENGINEERING THE REAL WORLD

# M DICTIONARY ENTRY

# SUMMARY

- **An improved method for XML semantics and syntax**
- **Base for interoperable data**
- **Exact search**

# IV. CONCLUSION

# SOFTWARE LICENSES, MIT

- **Kratulos**
  **(TLO Case No. 13752)**
  - Data integration
  - Machine understandable semantics
  - Connect models and data across the Internet
  - *Lee & Schuster*

- **MODS**
  **(TLO Case No. 13645)**
  - Finite schedule software for make-to-stock Mfg.
  - Heuristic that produces near optimal solution
  - Very fast, large problems take seconds to run
  - Other applications
  - *Schuster, Allen, Kar, & Lee*

# SOFTWARE LICENSES, MIT

- **Lee-Schuster Semantic Enterprise Architecture (TLO Case No. 13754)**
  - Intranet application
  - Low cost alternative for ERP
  - Minimal implementation
  - *Lee & Schuster*

For information on licensing, contact:


## Technology Licensing Office
Massachusetts Institute of Technology
Five Cambridge Center, Kendall Square
Room NE25-230
Cambridge, MA 02142-1493

Tel: 617-253-6966
Fax: 617-258-6790

Email: tlo@mit.edu

Edmund W. Schuster

Laboratory for Manufacturing and Productivity
Massachusetts Institute of Technology
77 Massachusetts Avenue, 35-135A
Cambridge, MA  02139

edmund_w@mit.edu

www.ed-w.info

http://ingehygd.blogspot.com


·GVTENBERG·WATT·ARKWRIGHT·
·WHITNEY·PERKINS·FVLTON·FAIRBAIRN·
·FROVDE·OTTO·DE LAVAL·WRIGHT·
ARCHIMEDES