MASSACHUSETTS INSTITUTE OF TECHNOLOGY
SLOAN SCHOOL OF MANAGEMENT

15.565 Integrating Information Systems:

Technology, Strategy, and Organizational Factors

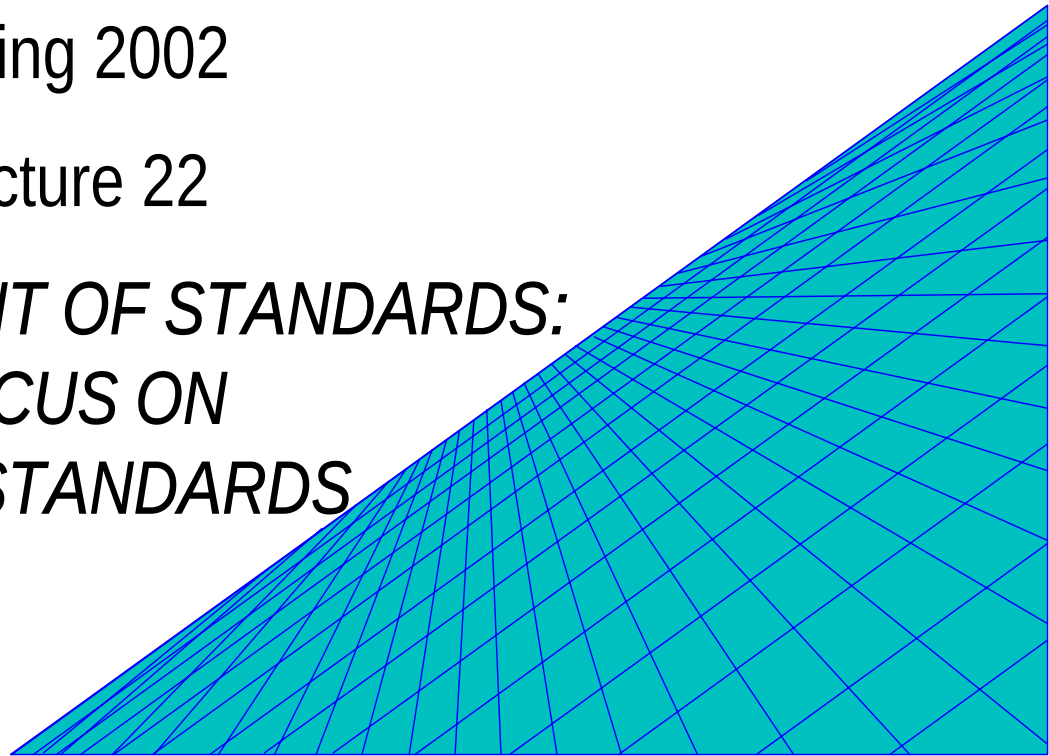
15.578 Global Information Systems:

Communications & Connectivity Among Information Systems

Spring 2002

Lecture 22

DEVELOPMENT OF STANDARDS:
FOCUS ON
DATA STANDARDS



ISSUES

- WHY NOT JUST ONE BIG ORGANIZATION?

  -- BENEFITS OF LOOSELY-COUPLED ORGANIZATION

  -- BOUND RATIONALITY <----> UNCERTAINTY / COMPLEXITY

- THUS, NOT ONE BIG PIECE, LOTS OF PIECES.

  -- HOW DO WE COORDINATE?

- STANDARDS PLAY A KEY ROLE

STANDARDS

- WHAT TYPE OF ISSUE?

  -- TECHNICAL

  -- STRATEGIC

  -- ECONOMIC [1]

  -- ORGANIZATIONAL[2] (POLITICAL)

- HOW ARE INTER-ORGANIZATIONAL STANDARDS DEVELOPED?

  -- INDUSTRY-WIDE

- WHAT ARE INSIGHTS FOR INTRA-ORGANIZATIONAL STANDARDS?

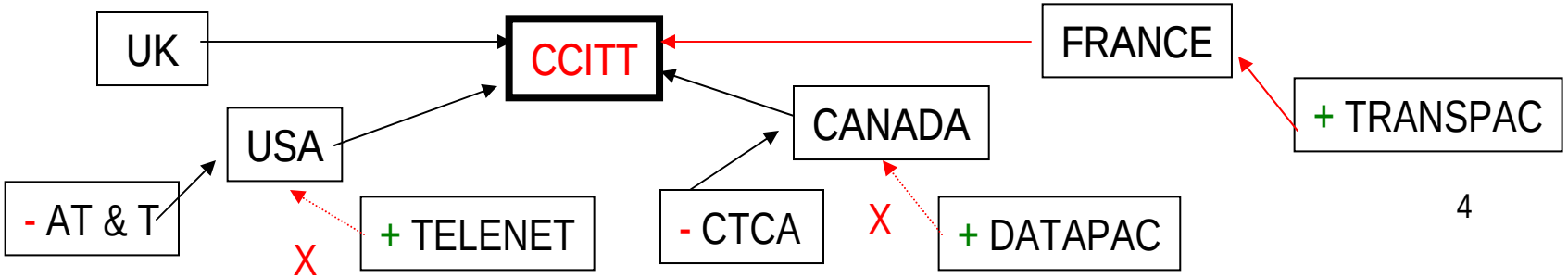
  -- INTERNAL TO A COMPANY

[1] Marshall Van Alstyne, Erik Brynjolfsson, and Stuart Madnick, "Why Not One Big Database? Principles for Data Ownership", *Decision Support Systems*, Vol 15, December 1995, pp. 267-284. [SWP #3695, CISL #94-03]

[2] Marvin Sirbu and Steven Stewart, “Market Structure and the Emergence of Standards,” CMU Working Paper, April 1985.

DATA COMMUNICATIONS STANDARDS

- FIRST DATA COMMUNICATIONS STANDARD
  -- PARIS, 1865, INTERCONNECTION OF TELEGRAPH SYSTEMS
- X.25 PACKET SWITCHED NETWORK -- SIRBU CASE STUDY
- PACKET SWITCHED NETWORK DEVELOPMENTS
  -- USA TELENET (1974); FRANCE TRANSPAC (1973); CANADA DATAPAC (1974), ...
- CONTROVERSIES / COMPROMISES (recall ATM 53 bytes standard)
  --X.25 CONTROVERSY OVER DATAGRAM VS VIRCUIT CIRCUIT
    -- RESOLUTION: PROVIDE BOTH (DATAGRAM RARELY IMPLEMENTED)
  -- PACKET SIZE DIFFERENCES:
    GERMANY (128, 1024), FRANCE (64, 128), JAPAN (32, 64, 128, 256), ETC.
    -- PACKET SIZE (128 BYTES COMMON), RETRIES, ETC.



MODEL FOR STANDARD DEVELOPMENT PREDICTION

- TYPES OF COMPATIBILITY
  - -- PEER-TO-PEER: TWO FACSIMILE MACHINES
  - -- INTERFACE: TV AND VCR

- VENDOR
  - -- CENTRALIZED (INTEGRATED) -- SUPPLIES ALL COMPONENTS
  - -- DECENTRALIZED (FRAGMENTED) -- NO ONE SUPPLIER PROVIDES ALL
    - -- DIFFER IN EXPERTISE,
    - -- PATENT/ LICENSING
    - -- GEOGRAPHY
    - -- CAPITALIZATION

- BUYER
  - -- CENTRALIZED -- ONE PARTY BUYS ALL COMPONENTS
  - -- DECENTRALIZED -- UNRELATED PARTIES
    - -- MODEMS, AM TRANSMIT/RECEIVER

SITUATIONS

- MAJOR FACTOR IS COORDINATION COSTS
  1. CENTRALIZED PURCHASE, CENTRALIZED PRODUCTION
     -- MULTIPLE STANDARDS CAN CO-EXIST (NO STANDARDS)
     -- NO TRANSACTION COSTS
     -- CANNOT SWITCH SUPPLIERS
  2. DECENTRALIZED PURCHASE, CENTRALIZED PRODUCTION
     -- STANDARD IMPORTANT
     (FACSIMILE TO ARBITRARY PEOPLE)
  3. CENTRALIZED PURCHASE, DECENTRALIZED PRODUCTION
     -- NO ONE FIRM SUPPLIES ALL COMPONENTS
     -- STANDARDS, BUT NOT UNIQUE (MULTIPLE STANDARDS)
     (CASSETTE AND CD; MULTIPLE OPERATING SYSTEMS)
  4. DECENTRALIZED PURCHASE, DECENTRALIZED PRODUCTION
     -- STANDARDS MOST IMPORTANT
     (X.25, TV)

SUMMARY OF KEY SITUATIONS

PRODUCTION DECISION (VENDOR)	PURCHASE DECISION (BUYER)	
	CENTRALIZED	DECENTRALIZED
CENTRALIZED	1. NO STANDARDS - CPU/DISK - LOCAL E-MAIL	2. STANDARDS IMPORTANT - DIAL-UP MODEMS - FACSIMILE
DECENTRALIZED	3. MULTIPLE STANDARDS - PBX - LAN	4. STANDARDS VERY IMPORTANT - TV - PACKET NETWORK

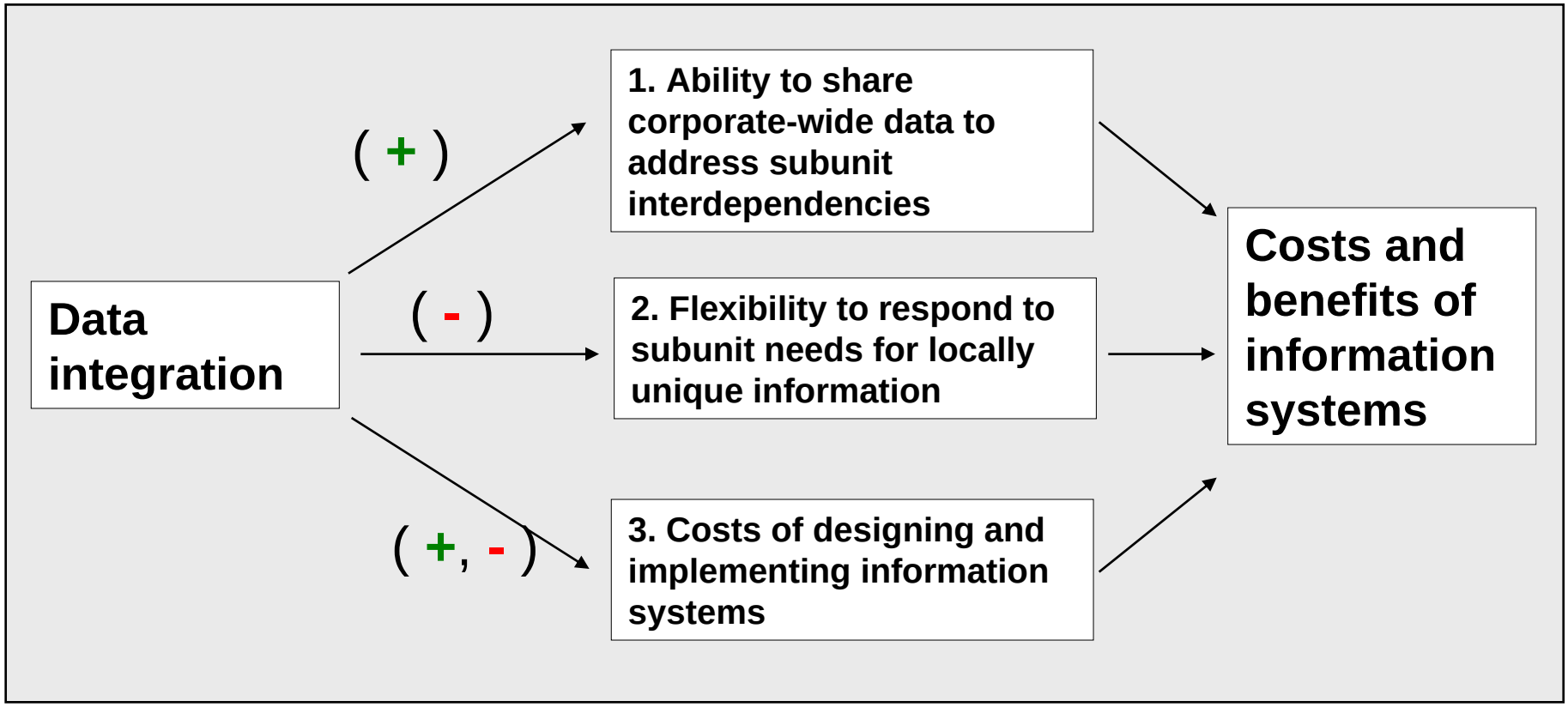
DATA STANDARDS / INTEGRATION: EXAMPLE AND COUNTER-EXAMPLE

	Integrated		Non-Integrated	
	Div A	Div B	Div A	Div B
1. PART_NUMBER for 3/4" SCREW	115899	115899	115899	718
2. CUSTOMER_ID Codes for IBM	42765	42765	42765	42675, 49345, etc.
3. Definition of SALES	Adjust returns	Adjust returns	Adjust returns	Not adjust returns

Example: Venezuela Nationalized Oil Companies – drilling pumps

Data Standards and Integration Research

The Impact of Data Integration on the Costs and Benefits of Information Systems



- Note: tradeoffs might be different for each sub-unit.
- Most successful if sub-units: (1) interdependent and (2) not highly differentiated.

Source: Dale L. Goodhue, Michael D. Wybo and Laurie J. Kirsch, "The Impact of Data Integration on the Costs and Benefits of Information Systems", *MIS Quarterly* , September 1992, pp. 300.

Lessons from the Field

Organization	Outcome
Ventura SSD	Textbook success story
Ventura Finance	Turning point in divisional support for IS
LSA	Still struggling to get organizational buy-in
Cedar	De-emphasized effort
Derrick	Shelved after reorganization
Nat. Tech.	Shelved
Consumers	Shelved
Foothill	Shelved

Source: Dale L. Goodhue, Michael D. Wybo and Laurie J. Kirsch, “Strategic Data Planning: Lessons From the Field”, *MIS Quarterly* , March 1992, pp. 23.

Benefits from Integrated, Sharable Data

- Improve communications across subunits (view for top mgt)
  - Data integration facilitates the collection, comparison, and aggregation of data from various parts of the organization, leading to better understanding and decision making when there are complex, interdependent problems.

Devlin Electronics:

Issues: On-time deliveries (a critical competitive issue in the semiconductor market) fell to only 70 percent, a multi-disciplinary team at Devlin Electronics used organization-wide integrated scheduling data to track how production schedules were made, changed, and adjusted by the many different groups involved. They found a number of interlocking problems: some plants were not properly updating their inventory levels and equipment conditions, marketing was overriding the schedule without regard to plant capabilities, and plants were overriding the system without regards to plant capabilities, and plants were overriding the system without regard to critical order requirements.

Outcome: Organizational-wide integrated data allowed Devlin to understand its problems and take corrective action. ***On time delivery improved from 70 percent to 98 percent***

Merrill Lynch (ML) - Current State:

ML has identified data integration issues and has launched several initiatives to standardize a minimum set of data definitions that will leverage the benefits of data integration within their different systems.

Example: Venezuela

Benefits from Integrated, Sharable Data

cont.

+ • Improve operational coordination across subunits (between units)
  – As the interdependence between subunits increase, the benefits of data integration will increase, and the amount of data integration in organizations should also increase.

Burton Trucking Company (BTC):

Issues: BTC completely overhauled its information processing systems based on a single logical data model for the entire corporation.

Outcome: This allowed them to link not only across geography but also across functions. By using common, sharable data, they found their dispatch systems (the responsibility of operations) could be expanded with little effort to give them a much better shipment tracking system (the responsibility of marketing). ***Thus, data integration allowed them to capitalize on previously unrecognized interdependencies between dispatching and shipment tracking.***

Merrill Lynch (ML) - Current State:

ML understands that standardization on data definitions is important to achieve this objective. The multi-disciplinary steering committee has been leveraging from other initiatives in different subunits (i.e. CPR) to improve coordination between different units and to avoid "reinventing the wheel".

Example: Venezuela

Flexibility to Meet Subunits Information Requirements

- Compromises in meeting local needs
  - As the differentiation between subunits increases, data integration will –
  - impose more and more compromise costs on local units; therefore, the amount of data integration in (rational) firms should decrease.

Burton Trucking Company (BTC):

Issues: At BTC, IS and operations ha developed a common dispatching system for its freight terminals across the country. It integrated data about all customers, equipment, and shipments.

Salesperson at each local terminal argued that they needed to add additional fields such as permissible delivery hours, after-hours phone numbers, and special instructions for drivers. But the salespeople couldn't agree on exactly which additional fields should be added. Terminals with close-in satellites (trucks every 2-3 hours) had a very different needs from those with distant satellites (trucks every 5 hours).

Outcome*:* IS decided that trying to standardize at this level didn't make sense. They designed 10 extra fields that the local people could use as they saw fit and gave them search capabilities and screens to update and query whatever data they needed in those fields.

Merrill Lynch (ML) - Current State:

Unique business units' need have been taken into account as the steering committee is defining high-level entities. However, ML might also need to analyze what is the implementation cost incurred by trying to meet a unit's unique need.

Example: Manufacturing company

Flexibility to Meet Subunits Information Requirements

cont.

- Bureaucratic delays that reduce local flexibility
  - Firms with increased data integration will experience greater bureaucratic delay in getting approval for changes affecting the data models used by individual subunits.

Superior Manufacturing (SP)

Issues: A major effort to integrate corporate data was underway. Procedures for changes to the corporate data model were: (1) Requests first go to data resource management office, which decides which decided which subject areas are affected, and (2) requests then are passed to the relevant subject areas data stewards who review, analyze, and recommend action. This recommendation will then been reviewed by affected division data administrators and their data user groups, who will approve or modify the recommendation.

Outcome: Finally, the modified recommendation will be presented to the corporate data resource management policy and steering committee for review and approval. All together, five groups will conduct four separate reviews of the request.

Merrill Lynch (ML) - Current State:

ML Steering Committee has been empowered to define those critical data elements with the support of Senior management. However, implementation is much more time consuming (i.e. CPR design and implementation).

Example: Johnson & Johnson

IS Design and Implementation Costs

- Why are up-front costs higher?
  - As the number and heterogeneity of subunits information needs increase, the difficulty of arriving at acceptable design compromise increases, and the costs of resulting design will increase more than linearly. Thus, (rational) firms will integrate less extensively when there are many heterogeneous subunits involved.

Ventura Products, Service Division (VPSD):

Issues: After having identified the data requirements for the first of two critical systems, VPSD spent many person-weeks of effort coordinating its new data definitions with the several thousand data standards on the corporate data dictionary. Because of the number of corporate standards and the possibility of subtle differences in meaning, it was quite difficult to determine which, if any, of the division's desired data elements were exactly the same as those in the corporate data dictionary.

Outcome: When the division began to design and develop its second new system, it decided the time and costs of coordinating with the corporate standards were not justified by the benefits to them.

Merrill Lynch (ML) - Current State:

A good exercise for ML is to review why other initiatives have been abandoned (i.e. PME).

IS Design and Implementation Costs

cont.

+ • What about long-term costs?

  – In turbulent environments, firms with many heterogeneous subunits will be even less likely to integrate extensively, and firms with homogeneous subunits will be more likely to integrate extensively.

−

Rolling Freight, Inc.:

Issues: The business has for years been oriented around managing the equipment needed to continue smooth railroad operations. Under this view of the business, customers and shipments were a secondary focus. With hard competition from trucking and other railroads, a new perspective is forming that puts customer and shipments on the center stage.

Outcome: This will require a major rethinking of the way in which data is conceptualized, captured, and managed.

Merrill Lynch (ML) - Current State:

ML has recognized that their different business units are becoming more interdependent (i.e. GRID modification to provide not only debt securities but also equity derivatives information) and several sources of data are being shared among units to meet this purpose. This factor might need to be revisited as ML moves into the implementation stage of its data standardization.

Other Important Issues to Consider

- Interdependences
  - In an organization, the technologies employed, the structure, the strategy, individual roles, and management processes are all tightly interdependent. No one can be changed without having an impact on the others.
  - Those data resources management efforts (and data integration efforts) that succeed are those driven by business requirements **clearly understood and championed by top management.**
- Power and Politics
  - Even with large net benefits to the organization as a whole, data integration may distribute those benefits in an uneven way, reducing the local autonomy of some divisions, changing the level access of critical information, or changing the power balance in some other way.
  - Even the possibility of changes in the power balance can cause resistance to a data integration effort by those concerned they might lose out, regardless of the argument taking a total organizational benefit point of view.

Success Factors Learned from Field Studies

*** Implementation of Integrated Systems**

Proposition #1: SDP is a design methodology, as well as a planning methodology.

Proposition #2: Data Integration must be critical to the strategic goals of the organization, as perceived by top management.

Proposition #3: Sufficient control over the planning domain is needed.

Proposition #4: Efforts to implement data integration need to balance global integration and local flexibility.

[From *Strategic Data Planning: Lessons From the Field*]

Success Factors Learned from Field Studies (continued)

* Data Architecture

Proposition #5: It is not clear what the most appropriate form for a data architecture is.

Proposition #6: It is not clear that SDP is the most effective means to produce such architectures.

Proposition #7: Too large a scope spells trouble.

Proposition #8: To have an effect on data integration, architectures must be enforced.

Proposition #9: Architectures should be "stolen" not reinvented.

Proposition #10: SDP spends too much time bringing novice data modelers up the learning curves.

Success Factors Learned from Field Studies (continued)

*** Identify Systems Priorities**

Proposition #11: For a systems priorities goal, SDP does not narrow its scope fast enough before the time consuming process of modeling business entities.

Proposition #12: Creativity is swamped by the volume of detail.

Proposition #13: The time required by individuals may self-select the wrong participants.

*** Rethinking Business Processes**

Proposition #14: The new understanding seems difficult to communicate.

Proposition #15: The cost of an SDP can probably not be justified by education and communication alone.

*** Education and Communication**

Success Factors Learned from Field Studies

Conclusion:

- *Widespread data Integration is an expensive proposition.*
- *In general, it will not be cost-effective to integrate all of an organization's data.*

  -- Better to "partially integrate" to achieve the most important benefits and avoid the most burdensome costs.

- *"Partial Integration" suggestions:*

  -- Limit the scope to only certain subunits.

  -- Require all subunits to use a selection of "global" or organizational-wide application systems (e.g., payroll) - with discretion on other less critical application systems.

Success Factors Learned from Field Studies

cont.

Conclusion:

- *Develop selected organizational-wide databases (such as those related to customers or products).*

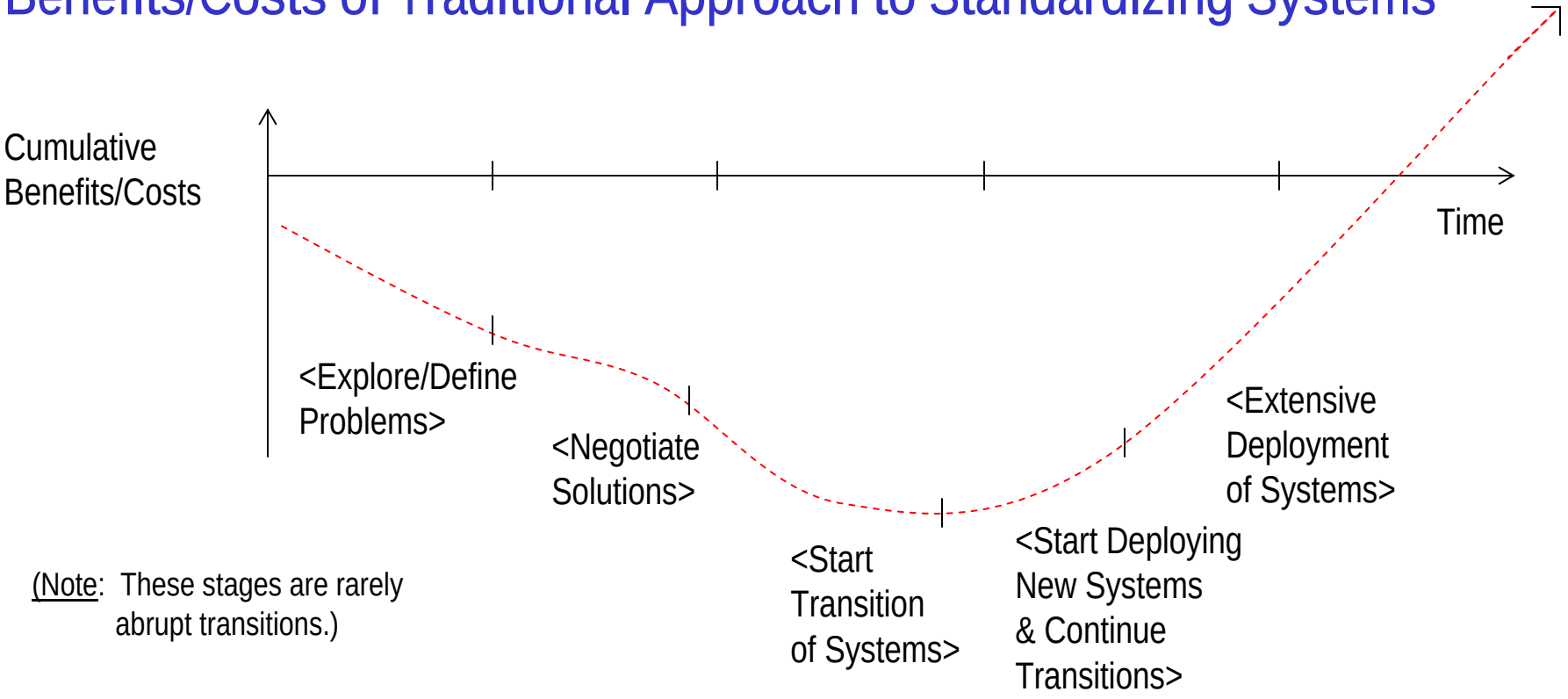
  -- require all application systems that use these entities to access and update the common databases.

- *Identify a selection of critical data elements.*

  -- hammer out agreed-upon definitions for those across the entire organization.

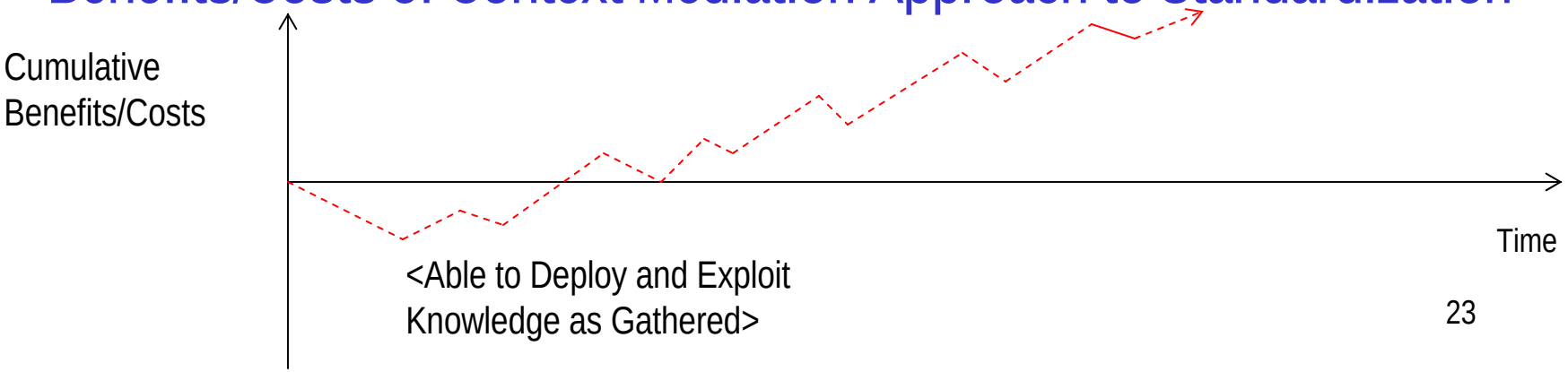
  -- These standard definitions can then be enforced in all systems development.

Benefits/Costs of Traditional Approach to Standardizing Systems



(Note: These stages are rarely abrupt transitions.)

Benefits/Costs of Context Mediation Approach to Standardization



CONCLUDING ISSUES

- DEFINITION OF CENTRALIZED VS DECENTRALIZED
- DEFINITION OF PURCHASER AND INTERCONNECTION (DIFFERENT OPTIONS)
  - -- DEPARTMENT
  - -- CORPORATION
  - -- INDUSTRY
- IMPLICATIONS FOR INTRA-ORGANIZATIONAL STANDARDIZATION
  - -- SAME CHALLENGES FOUND (E.G., MERRILL LYNCH EXAMPLE)
  - -- DIFFICULT TO ATTAIN: COST / BENEFIT TRADEOFFS, COMPLEXITY, TIME
  - -- "FOCUSED STANDARDS" APPROACH
- USE OF MEDIATOR TECHNOLOGY (CONTEXT MEDIATOR)
  - -- TO FACILITATE
  - -- TO DEAL WITH EVOLUTION