R I V C O M

XML and Industrial Data: Experiences and Observations

Tony Stewart

Director of Consulting, RivCom

8 December 1998

RivCom

- Publishing services company
  - Specialising in structured information
  - Both print and electronic
- A small multinational with major industrial clients
- Active in standards development
  - W3C XSL and XML-Schemas
  - ISO STEP and STEP/SGML harmonisation
  - OASIS, PISTEP
- One of the first to use XML for web and print publishing (1997)
  - The first to display XML plus styling in an industry-standard browser

This Presentation

- Why industry is ready for XML
- What do I mean by industrial data?
- Issues and activities
  - Examples of current and recent projects
  - Often proprietary, and therefore unnamed
  - Most are in early stages
- Observations and conclusions
  - Related demonstrations during the breaks

Why industry is ready for XML

- HTML provided a fantastic ability to publish images of information
  - Static documents and data snapshots
  - Intended for humans
- Firewalls enabled a robust infrastructure
  - An undervalued component
  - Allows industry to use the web
  - But most systems remain proprietary and non-interoperable
- XML provides the remaining piece of the puzzle

Enter XML

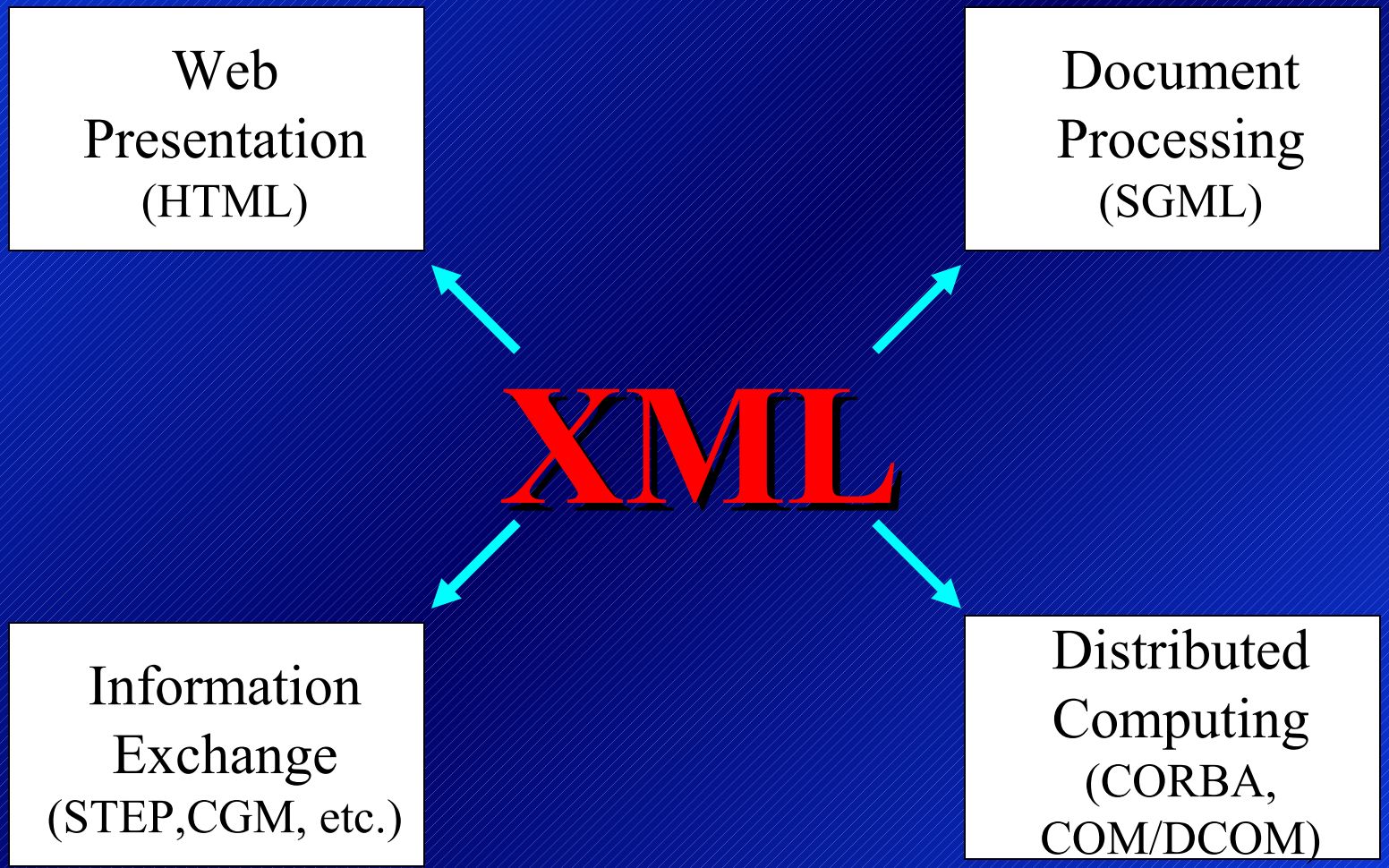
```
<person ID="123" location="ny">
 <first-name>Tony</name>
 <last-name>Stewart</last-name>
 <comments>Tony is a good speaker but has an
annoying habit of preparing his slides at the last
minute.</comments>
</person>

<location ID="ny">
 <name>New York Office</name>
 <address>945 West End Avenue</address>
</location>
```

XML gives us...

- Separation of content from format and behaviour
  - Information can be processed as well as presented
- The ability to transmit arbitrarily complex data structures
  - Extensible information systems through the use of links
- System-independent, interoperable data exchange
  - Easy to transform to/from other formats
- Multiple ways to apply semantics to the data
  - Style sheets, schemas, namespaces...

XML enables communications



Industrial Data

- Industries:
  - Processing (oil, nuclear, household detergents)
    - Requires processing plants
  - Engineering
    - Builds the processing plants
  - Aerospace
    - Complex products with long life cycles
  - In short, major enterprises
- Phases:
  - Design
  - Manufacturing
  - Lifecycle management
  - Enterprise operations

Phase 1: Design

- Inputs (corporate knowledge)
  - Data models
  - Research reports
  - Engineering standards and guidelines
- Outputs (the designs)
  - Drawings & diagrams (data)
  - Design notes, parts lists and data worksheets (documents)

XML projects

- Project: Linking research reports to the design process
  - OCR the research reports
  - Use XML for metadata and simple structural tagging
  - Looking at best ways to link to data models and design tools
- Project: Improving delivery and maintenance of engineering guidelines
  - Using XML for structural tagging and possibly XML/XSL delivery
  - Looking at best balance of structure vs. authorial freedom

Phase 2: Manufacturing

- Exchanging design information across the enterprise
  - Pricing and procurement
  - Production
- Developing documentation
  - Technical and end-user materials
  - Sharing information between engineers and technical writers
- Recording design revisions
  - and updating the documentation

XML Projects

- Project(s): Creating and delivering technical documentation
  - Updating existing systems to be more structured
  - Rethinking and rebuilding "failed" SGML projects
    - Reasons for failure usually stem from SGML rigidity
      - Requirement that the entire document conform to a DTD
      - Engineers don’t like tagging, can't maintain the documents
    - XML provides more flexibility
      - DTD optional
      - Alternate schema encodings
      - Can use XML islands (metadata within a non-well-formed document)
        - » Experimenting with Office 2000 "save as HTML" format

XML-related standardisation

- Goal: Use XML to transmit product designs
  - XML representation of EXPRESS schemas and EXPRESS-driven data (ISO TC184/SC4 NWI)
    - Demonstrations next month at ISO meetings in San Francisco
  - XML Schema efforts (including RDF, DCD, SOX...)

Phase 3: Lifecycle Management

- Maintenance
  - Instantly deliver appropriate documentation
    - Decision trees, IETMs
  - Update maintenance records (data warehouse)
- Customised instances of products
  - Create and maintain unit documentation
  - Update maintenance records
- Abnormal situations
  - Generate operating instructions on the fly
  - Notify internal and external authorities
  - Log all actions

XML developments

- Technology: Improved SGML/IETM delivery tools
  - Trickle-down benefits of XML technology development
  - Products from CITEK, SÖRMAN, DPSL using IE components and Mozilla code
- Technology: Flurry of repository and workflow vendor activities
  - POET and ObjectStore bulding XML storage layers
  - Omnimark demonstration delivering Aerospatiale manuals from relational database

XML-related standardisation

- Goal: Merge “documents” with “data” in major systems
  - STEP-SGML/XML harmonisation (ISO TC184/SC4 PWI "SGML and Industrial Data")
    - Enabling human-readable information (documents) and machine-processable information (product data) to be processed by the same tools
    - Working at an abstract level:
      - EXPRESS schema for SGML property sets
      - Property set for EXPRESS data

Phase 4: Enterprise Operations

- Gathering and distributing corporate knowledge
  - Procedures and guidelines
  - Best practices
  - Business models
  - Operational data (update and retrieve)
- Facilitating transactions
  - EDI

XML Projects (models)

- Project: Using XML to web-enable a data warehouse
  - Viewing and navigating the class hierarchy
- Projects: Using XML to transmit and navigate information models
  - PRIMA consortium
    - Prototype uses XML to view and transmit Business Process models based on the CIMOSA methodology
  - PISTEP Engineering Activity Model
  - Shell Downstream Business Activities Model
    - XML structure, HTML and XML delivery
  - XML for Express-driven data

Aside: XML empowers enterprise models

- XML makes the enterprise model directly accessible via the intranet
- Enables powerful viewing and navigational tools
- Links to other corporate information are easy to define and implement
- The model can become the heart of the enterprise information system
  - Similar to navigating a Piping & Instrumentation Diagram to locate a specific pump

XML Activities (EDI)

- Projects: XML for EDI
  - Demonstration system to be developed by consortium in the XML/EDI group within the CEN/ISSS Electronic Commerce Workshop
  - Numerous projects in the States (CommerceNet, Veo systems)
  - Commerce, medical informatics, EDIFACT vs. X12
- Technology: XML as an information interchange format
  - SAP, BAAN, ORACLE

Summary and conclusions...

- Structural and theoretical work is moving quickly
- Tools have started to appear
  - But most are still in beta!
- The most active projects are in the areas of data transmission and transformation...
  - Generated from a database or serialised from a model
- … more flexible document authoring & delivery...
- … and of course EDI
- But don’t overlook how XML is creating new roles for documents

Four documents, four paradigms

- Document as Information (traditional)
  - Shell's Guide to Procurement and Logistics Management
- Document as Navigable Map
  - PISTEP Engineering Activity Model
  - Product and enterprise models
- Document as database interface
  - RivCom's SGML/XML structures prototype
- Document as Application
  - Shell's Competence Gap Analysis Tool
  - (ideal for data worksheets)

Vision: XML at the heart of the system

- An overheating valve generates...
- Messages to people (documents)
  - Changes the appearance of the operating instructions
- Messages to machines (software objects)
  - Reduces flow to the valve; shuts down operations
- Messages to the outside world (email/phone/etc.)
  - Notifies headquarters, calls the fire brigade
- Messages to the data warehouse (data)
  - Logs all readings, displayed information, automatic and operator actions, etc.

Thank you for your attention

Tony Stewart
Director of Consulting, RivCom

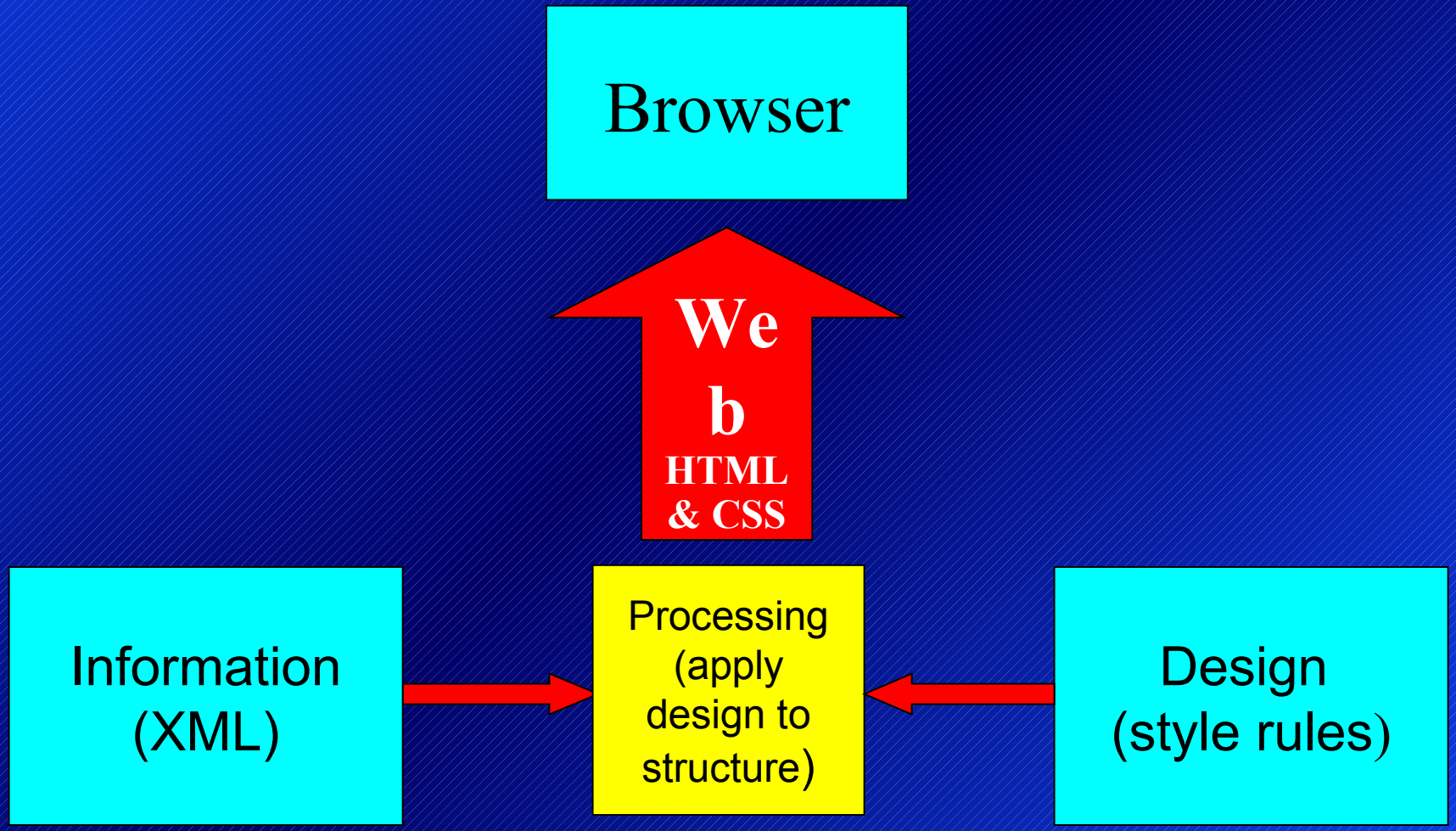
Tel: +1 212 222-4332
Fax: +1 212 662-6900
UK Tel: +44 1793 792000
UK Fax: +44 1793 792001

tony.stewart@rivcom.com
http://www.rivcom.com

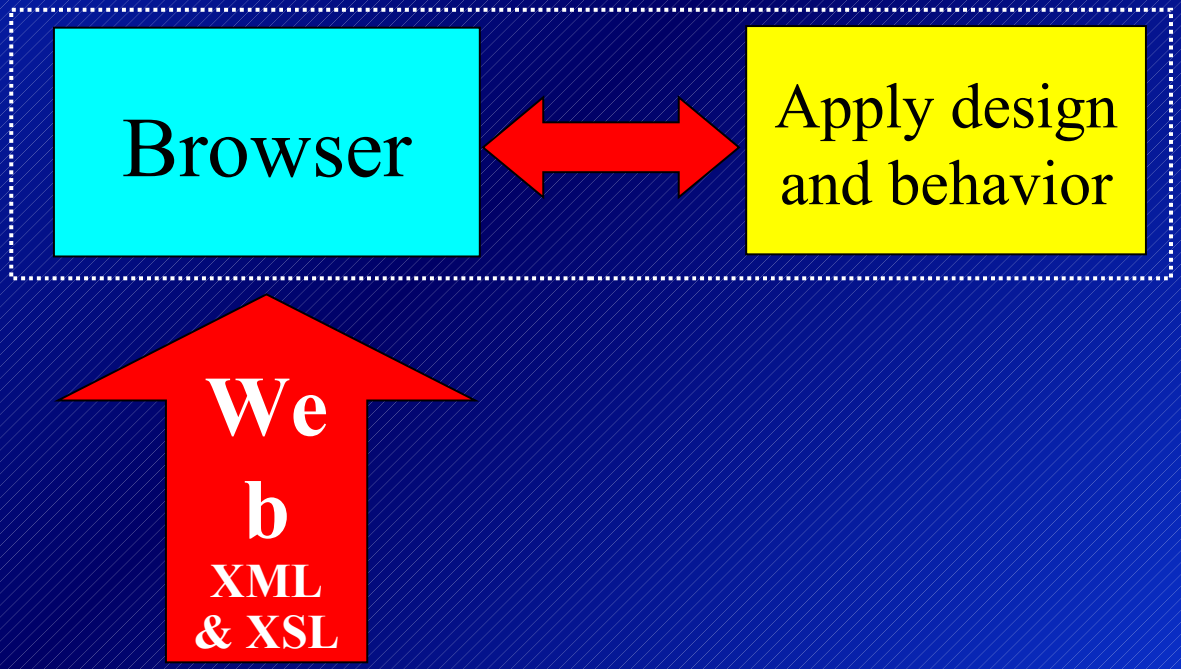
XML gives us...

- Separation of data from format and behaviour
  - Information can be processed as well as presented
- The ability to transmit arbitrarily complex data structures
  - Extensible information systems through the use of links
- A system-independent encoding for interoperable data exchange
  - Easy to parse and transform to/from other formats
- Multiple ways to apply semantics to the data
  - Style sheets, schemas, namespaces...

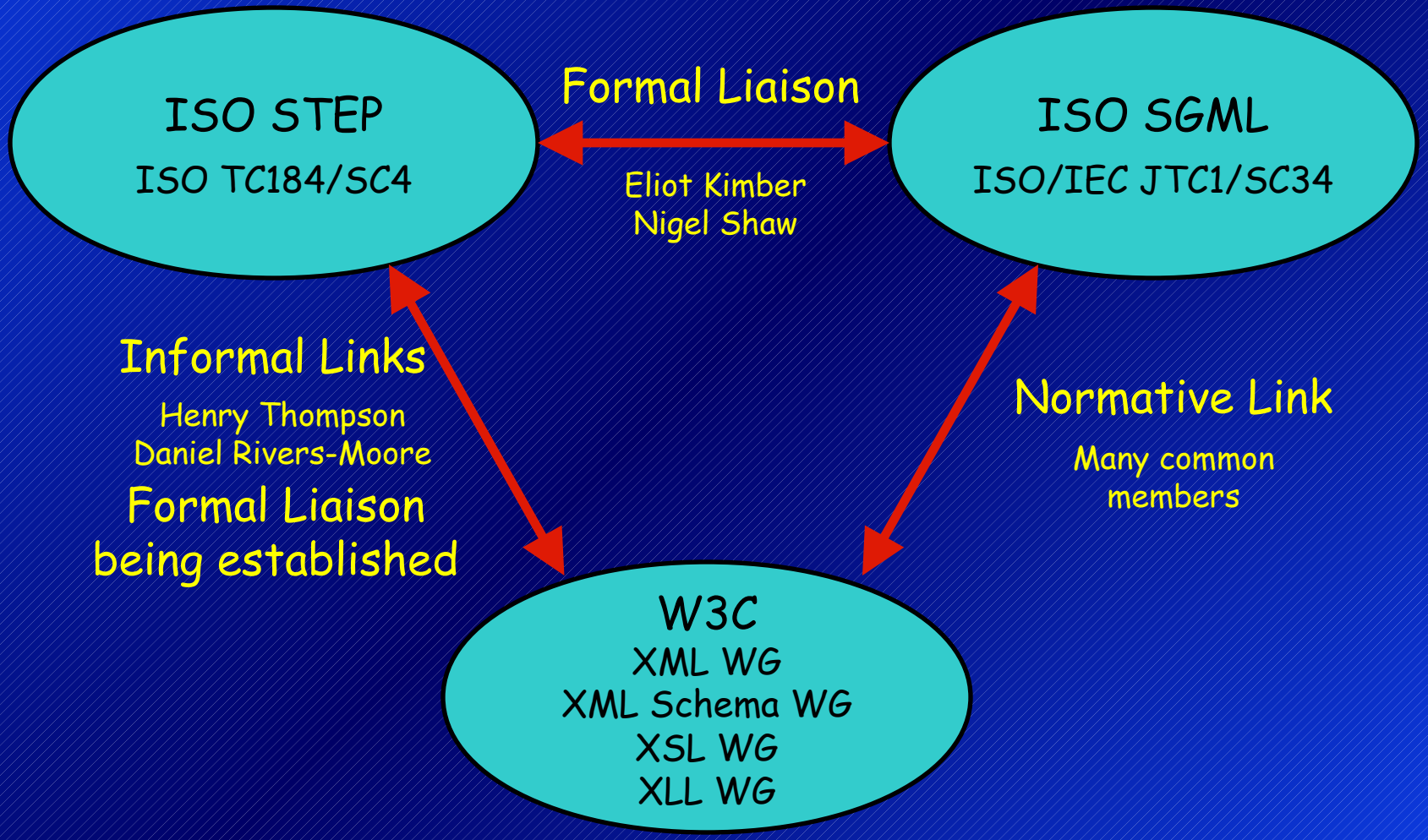
Apply Design to Information - On the Server



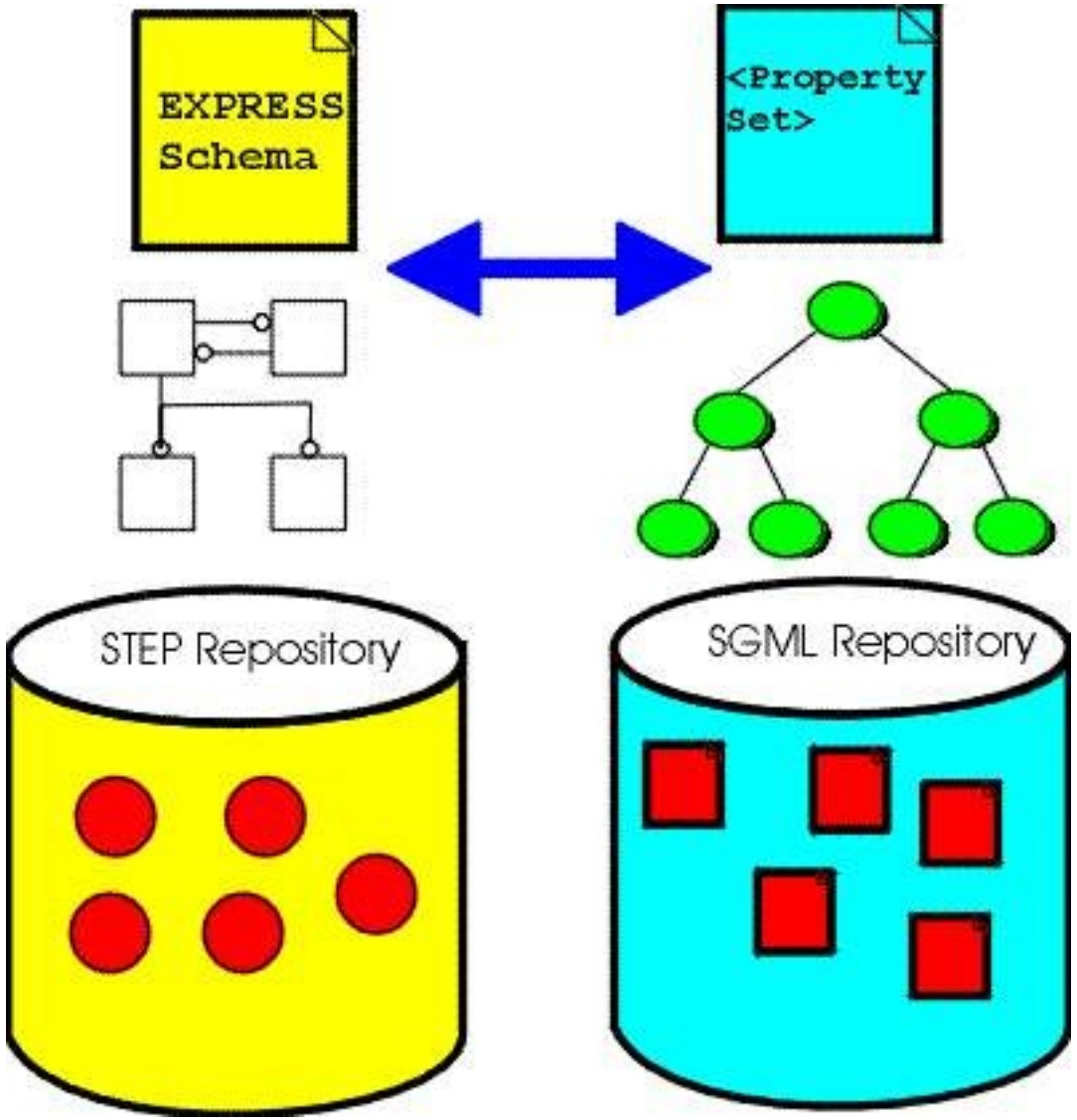
Apply Design to Information - In the Browser



STEP/SGML harmonization



Document/data integration (1)



Document/data integration (2)