XML in the News – Introduction and Overview
Tony Stewart
Director of Consulting – RivCom Ltd.

"XML in the News"

XML-UK

7 October 2003

- Current environment offers threats and opportunities
  - Newspapers face increasing competition in shrinking market
  - Changing customer profiles, delivery channels, business models
  - Survival requires ability to generate new revenue streams, improve processes and leverage new technologies
- These issues affect <u>all</u> news and media players
  - News agencies, ad agencies, PR firms, aggregators, production houses, printers, etc.

- XML and Internet technologies are part of the problem, and part of the solution
- The "XML space" is in transition
  - Established technologies based on the first set of XML standards are delivering real value
  - Newer, competing standards are confusing and in conflict
- Newspaper publishers must make appropriate decisions about how to respond
  - Business <u>and</u> technical issues
  - Impact on strategies, tactics, processes

- Overview
  - Introduce XML
  - Provide a high-level framework
- XML specifications for news and media
  - Three XML specifications in varying degrees of completeness
- Strategic and practical issues
  - Time Out case study
  - National and Regional newspapers
- Pulling it all together
  - International perspective
  - Panel discussion

What is this XML thing?

- Extensible Markup Language
- Core standard (W3C) provides a remarkably simple mechanism for encoding data/information

```
<presentation>
    <author>Tony Stewart</author>
    <presenter>Tony Stewart</presenter>
    <event>XML UK – XML in the News</event>
    <slide number="1">
        <title>Welcome<title>
        <body>…</body>
        …
    </slide>
</presentation>
```

- XML lets you define "markup languages" for specific purposes
  - e.g. an XML "language" for slide presentations
- An XML language is
  - A set of placeholders (tags) for specific types of information
  - Agreement between those who will use it
    - Simple to reach agreement if just internal
    - Complex if many parties involved
- XML is becoming the format of choice for <u>all</u> classes of information
  - Data, text, graphics, music, video, etc.
  - Standalone or as a metadata wrapper for non-XML info
    - Text, data – always
    - Graphics – depends (SVG vs. PDF etc.)
    - Music, video, etc. – probably not
- There are many XML formats, for many different purposes

What are all these things?

- XML
- XSLT
- XSL-FO
- XML Schema
- Ad/Venture
- CREST
- ebXML
- NewsML
- NITF
- PRISM
- SPACE/XML
- SportsML
- XPRL
- …

- AdsML
- IfraTrack
- JDF
- PROSE/XML
- SOAP
- RSS
- SVG
- XML Topic Maps
- RDF
- (and…)
- ASP/JSP
- HTML/CSS
- JavaScript
- …

- Two broad categories of XML languages
  - **Content** such as news stories, business documents, graphics
    - This document is a <presentation>; this line is an <author>
  - **Processing rules** about how a system should deal with XML information
    - If the document is an Invoice, acknowledge it and route it to the Accounts Payable system
    - If the text is a <speaker>, boldface and center it
    - If the object is a Circle, make it round
- Using these XML languages plus relatively generic tools we can
  - Define a type of information and how it should be processed
  - Exchange the information (internally or externally)
  - Control the processing

- **Types of processing**
  - **Define** the content and processing rules
  - **Validate** the documents/files/messages
  - **Store** them
  - **Search** them (much more powerful than plain text)
  - **Exchange** them (Internet oriented)
  - **Transform** them (combining multiple sources; for multiple purposes)
  - **Present** them (in multiple ways, in multiple media)

- **Types of applications**
  - **UI** (human interface presentation, display, print, etc.)
  - **EAI** (Enterprise application integration)
  - **eCommerce**/EDI
  - **Web Services** (distributed computing)
  - **BPM** (business process management, workflow)

Categories of XML standards

Foundation XML (standards to define, validate and transform XML)

XML, XML Schema, XSLT

Business information

General: ebXML, cXML…
Industry specific: AdsML,
Ad/Venture, CREST,
NewsML, NITF,
IfraAdConnexion, PRISM,
SPACE/XML, SportsML,
XPRL…

Business control

General: BizTalk,
BPEL, ebXML…
Industry specific:
AdsML, Ad/Venture,
IfraTrack, JDF…

Presentation

XML: PDF/X, RSS,
SVG, XSL-FO, XML
Topic Maps, XSLT…
Related: ASP/JSP,
HTML/CSS,
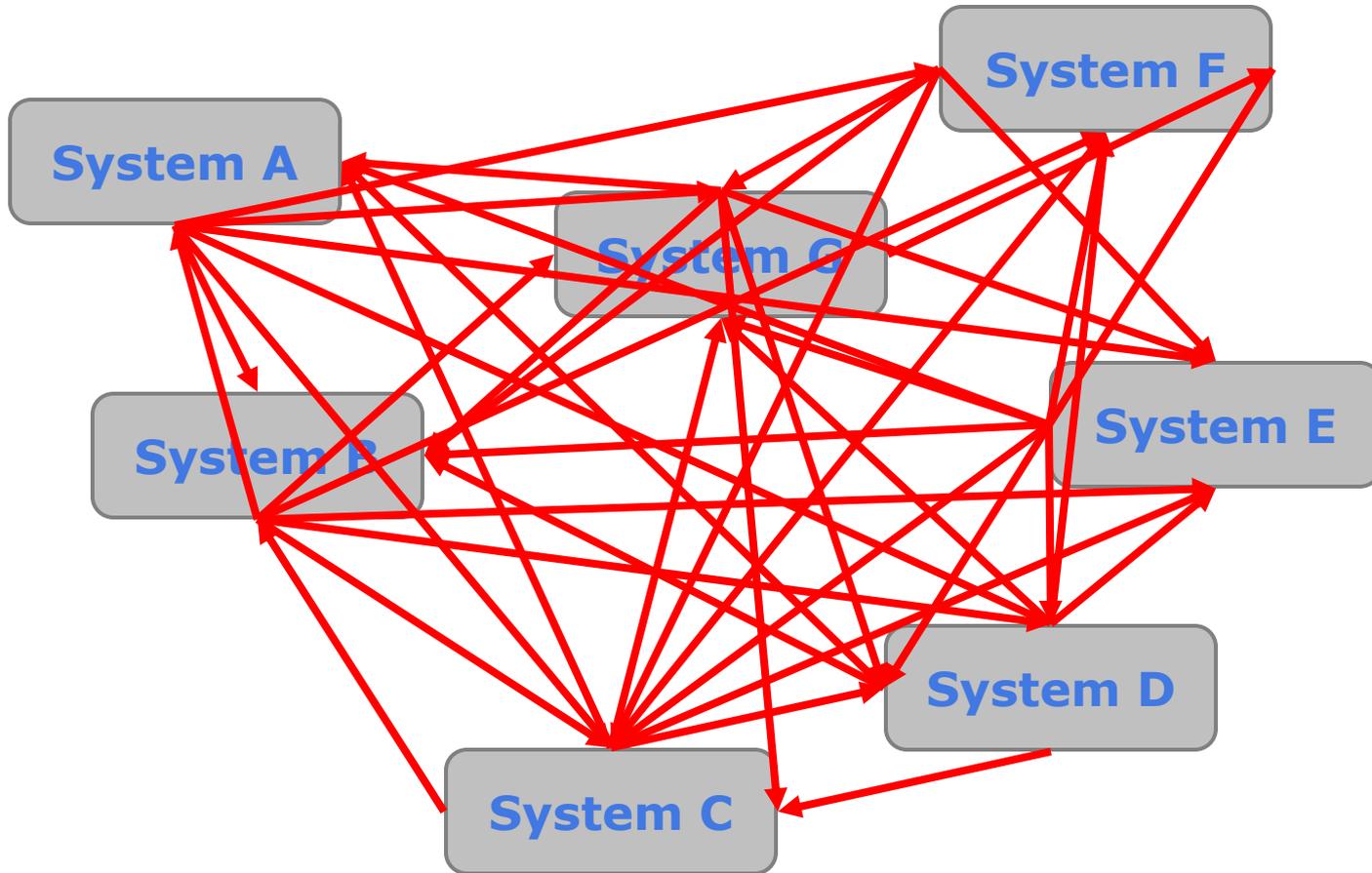JavaScript…

Reference information
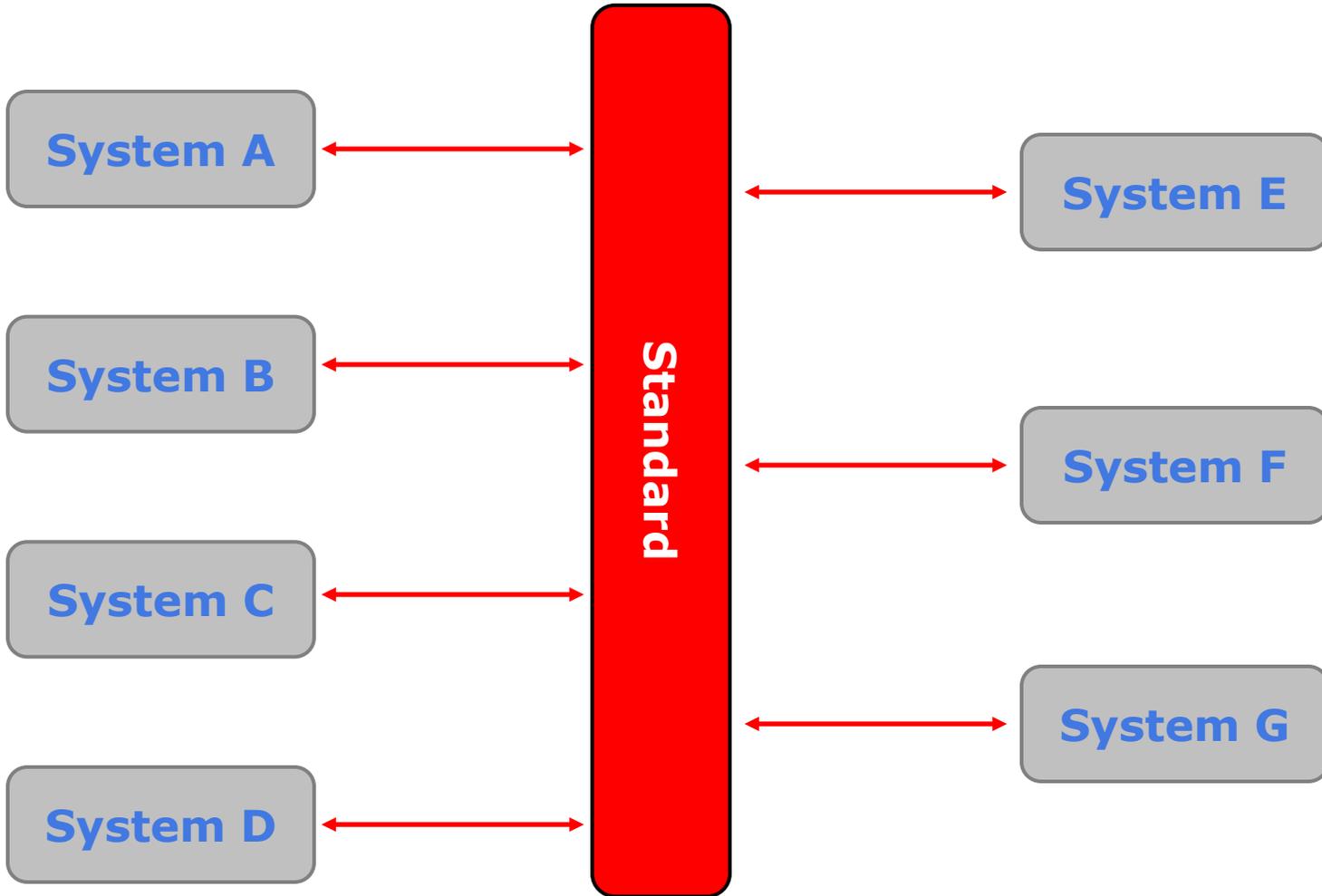RDF, RSS, XML Topic Maps, XMP…

Distributed computing
SOAP, Web Services…

RIVCOM
enabling business process improvement

- XML for presentation
  - Capabilities are well underway and understood
  - Many tools to apply formatting and styling to XML
    - Aiming for "write once, use many" and other holy grails
  - Vendors are delivering
- XML for business information and business control
  - Together these enable system-to-system communications
    - Exchange of defined business information following defined rules ("choreography")
    - More flexible and less expensive than non-XML "EDI"
  - This area is in flux
    - But it will lead to even greater business benefits…

A world without standards

With standards...

System A ↔ Standard
System B ↔ Standard
System C ↔ Standard
System D ↔ Standard
Standard ↔ System E
Standard ↔ System F
Standard ↔ System G
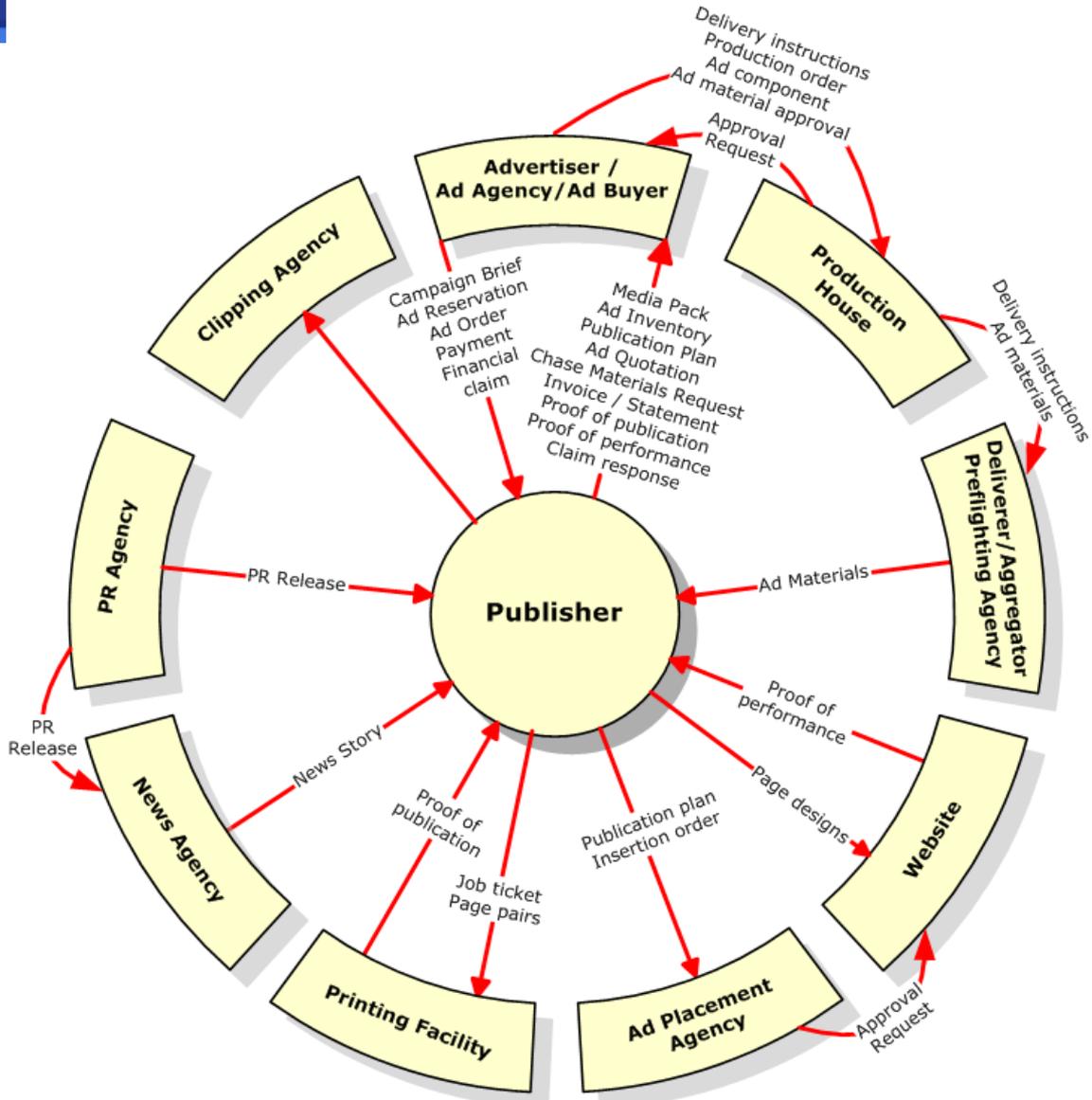
Publishing supply chains

- There are several major news-specific supply chains
  - News
  - Ads
  - Press releases
  - Publication (to print)
  - Publication (to Web)
- Each of these is a complex process involving multiple players…
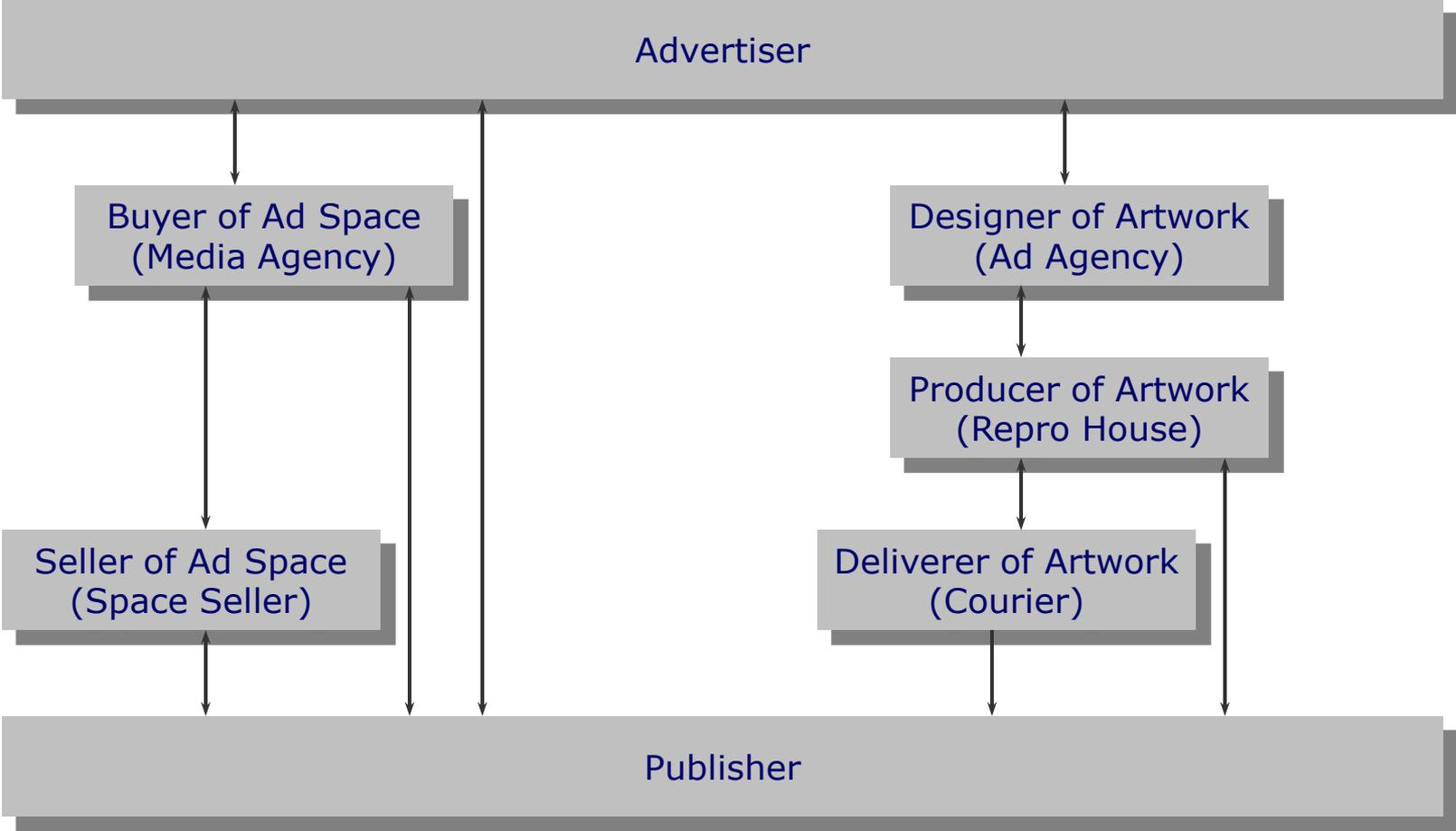
- Consider just some of the supply chains that publishers participate in
  - News
  - Ads
  - Press releases
  - Publication (to print)
  - Publication (to Web)
- Each of these is a complex process involving multiple players…
- All of which are active simultaneously…

Publisher-centric information flows

- Consider just some of the supply chains that publishers participate in
  - News
  - Ads
  - Press releases
  - Publication (to print)
  - Publication (to Web)
- Each of these is a complex process involving multiple players…
- All of which are active simultaneously…
- The ROI argument for standardisation is very powerful.

How do standards come about?

RIVCOM
enabling business process improvement

- A business community defines a standard
  - Types of information
  - Rules governing their movement, validation, processing
  - (assumes compelling reason to cooperate)
- Vendors create software
  - Business case?
    - Speed depends on industry drivers, participation
    - Pilot projects often driven by one or more "big hitters"
- Time passes…
  - Many proposed standards never find their audience
  - Critical mass becomes apparent in vendor offerings
- Standards become embedded in application software
  - Users often change their processes accordingly

- Foundation, Reference and Presentation standards are used in developers tools
  - You only need what you need
- Business Information/Control
  - General level, competition and confusion
    - Big players fighting turf wars
  - Industry level, relatively little competition
    - "Bubbling up" of regional solutions
    - Groups like IPTC, Ifra, NAA cooperating to resolve overlaps

Should I wait for the dust to settle?

- Use XML tools today for
  - Presentation
  - Storage (in some cases)
  - Application integration within your organisation
  - Information exchange where standard is stable and tools exist
    - (e.g. NewsML)
- Use for more purposes tomorrow when the tools support it
- Inter-organisation communications require community agreement
  - Ifra, NAA, AdsML, XPRL aspire to global solutions
    - Have to take into account work that is already in place
- Standards and standards bodies need your support

Why participate?

- "Standards are the basis of competition going forward. If you can set the standard or drive the standard, that enables you to get to market faster and have a broader market … Think of standards as an accelerator to your business."

  Tony Scott, CTO, General Motors

- "Our participation in standards groups comes from the fact that we're up front about creating an architecture strategy for our organization. We want to understand where they are going early so we can plan."

  **Adrian Kunzle, Co-Head of the Architecture Group,
  JP Morgan Chase Investment Banking Division**

Any questions?

- ?

Categories of XML standards

Foundation XML (standards to define, validate and transform XML)
XML, XML Schema, XSLT

Business information

General: ebXML, cXML…
Industry specific:
Ad/Venture, CREST,
NewsML, NITF,
IfraAdConnexion, PRISM,
SPACE/XML, SportsML,
XPRL…

Business control

General: BizTalk,
BPEL, ebXML…
Industry specific:
AdsML, Ad/Venture,
IfraTrack, JDF…

Presentation

XML: PDF/X, RSS,
SVG, XSL-FO, XML
Topic Maps, XSLT…
Related: ASP/JSP,
HTML/CSS,
JavaScript…

Reference information
RDF, RSS, XML Topic Maps

Distributed computing
SOAP, Web Services…

Categories of XML standards

Foundation XML (standards to define, validate and transform XML)

XML, XML Schema, XSLT

Business information

General: ebXML, cXML…
Industry specific:
Ad/Venture, CREST,
NewsML, NITF,
IfraAdConnexion, PRISM,
SPACE/XML, SportsML,
XPRL…

Business control

General: BizTalk,
BPEL, ebXML…
Industry specific:
AdsML, Ad/Venture,
IfraTrack, JDF…

Presentation

XML: PDF/X, RSS,
SVG, XSL-FO, XML
Topic Maps, XSLT…
Related: ASP/JSP,
HTML/CSS,
JavaScript…

Reference information
RDF, RSS, XML Topic Maps

Distributed computing
SOAP, Web Services…

Categories of XML standards
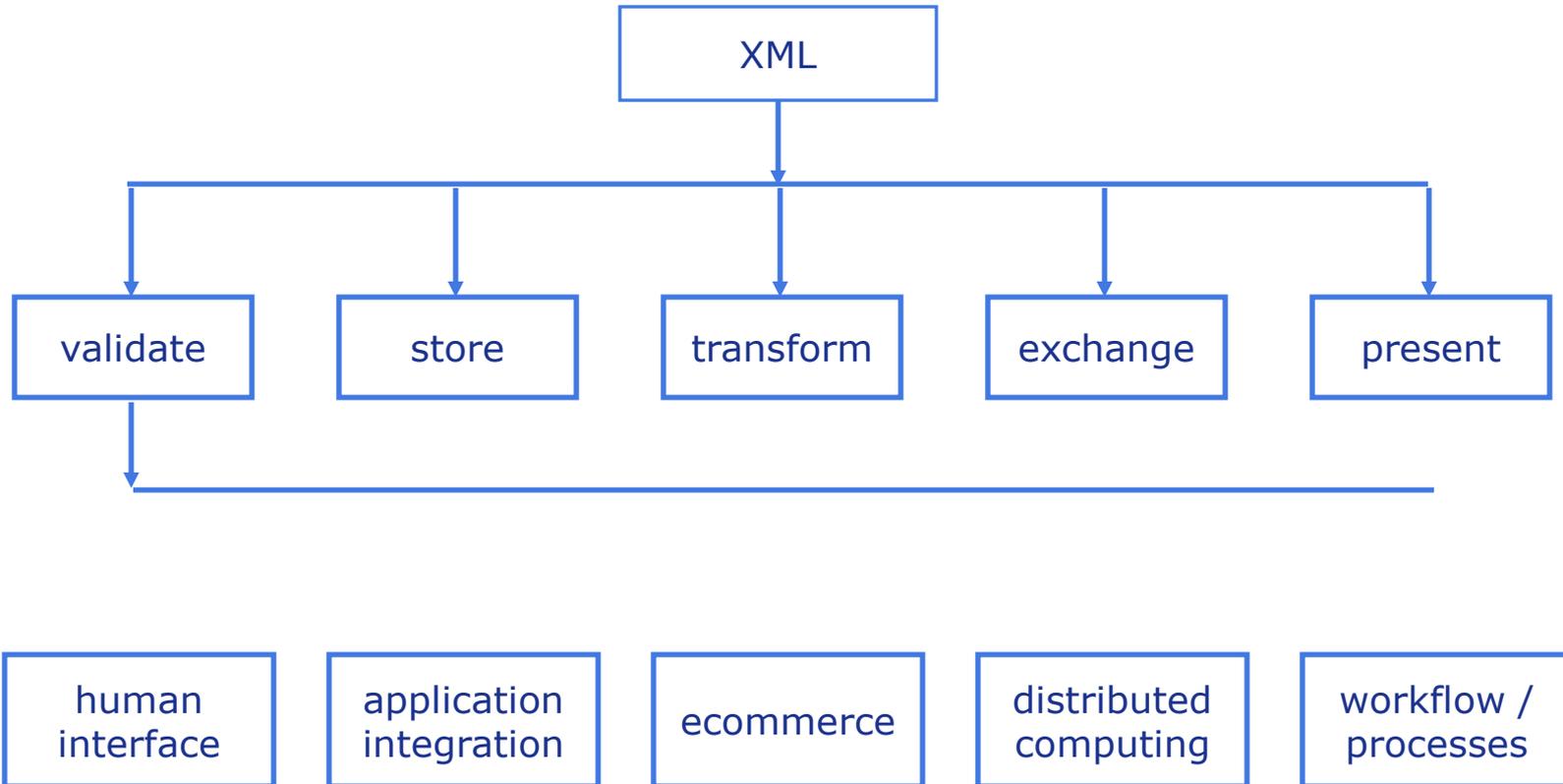
- Foundation
    - Standards to define and manipulate other XML formats
        - XML, XSLT, XML Schema
- Business information
    - Formats for representing specific types of business-level information
        - NewsML, SportsML, XPRL…
- Business control
    - Formats to control workflows and eBusiness choreography
        - AdsML, JDF, ebXML…
- Presentation
    - Formats and technologies used to "style up" XML information
        - HTML/CSS, XSL-FO, XSLT (again), JavaScript…
- Reference
    - Formats used to represent information models and networks
        - RDF, XML Topic Maps
- Distributed computing
    - Formats to enable Web Services and distributed computing (SOAP…)

WML Client

XML Client

Processing engine

HTML Client

XSLT

Database

What is XML?
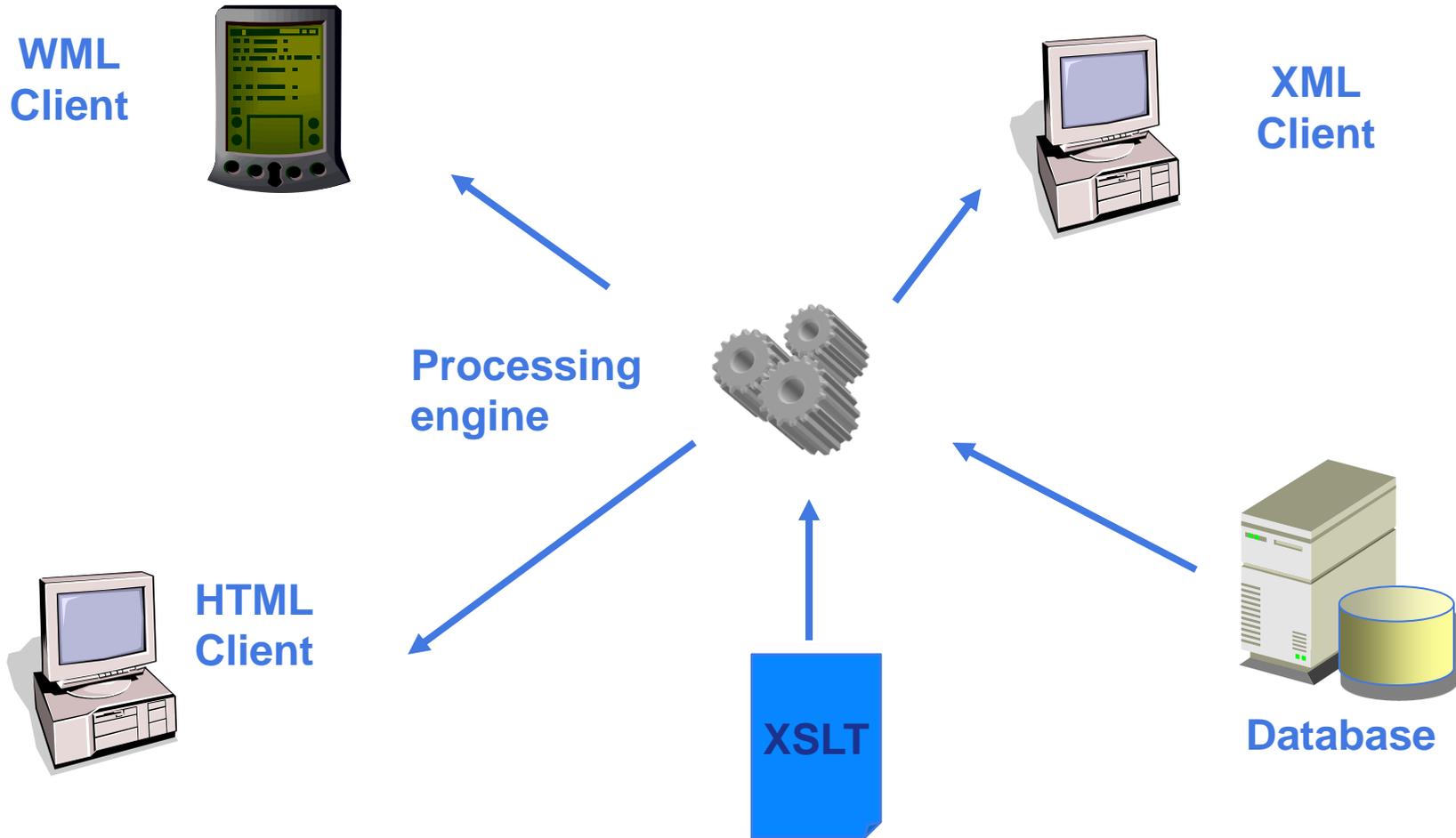
- "Extensible Markup Language"
  - Mechanism for creating computer-readable "languages" for specific purposes
  - An ever-growing set of languages that have been created
- XML aims to solve the general problem of making it easier for computers to communicate information
  - Between each other
  - Between a computer and a human being

- XML provides a standard mechanism to define "languages" or "formats" for specific types of information
  - Irrespective of the application or computer that will process it
  - Easy to change over time
  - Lots of ways to process it
- These are created for particular communities
- The members of the community still have to agree on what information they want in their XML documents,

Categories of XML standards

Foundation XML (standards to define, validate and transform XML)

XML, XML Schema, XSLT

Business information

General: ebXML, cXML…
Industry specific:
Ad/Venture, CREST,
NewsML, NITF,
IfraAdConnexion, PRISM,
SPACE/XML, SportsML,
XPRL…

Business control

General: BizTalk,
BPEL, ebXML…
Industry specific:
AdsML, Ad/Venture,
IfraTrack, JDF…

Presentation

XML: PDF/X, RSS,
SVG, XSL-FO, XML
Topic Maps, XSLT…
Related: ASP/JSP,
HTML/CSS,
JavaScript…

Reference information

RDF, RSS, XML Topic Maps

Distributed computing

SOAP…

Processes: newspaper ad supply chain

**Buyer of
Ad Space**
- Advertiser
- Media Agency
- Ad Agency
- Newspaper
  Ad Space
  Buying
  Company

**Seller of
Ad Space**
- Newspaper
- Media Agency
- Ad Agency
- Newspaper
  Ad Space
  Selling
  Company

Media Data ←

Ad Order →

Confirmation of
Ad Order on request ←

Artwork →

Proof / Fax Proof →

Invoice ←