XML and other serializations: Shoe-horning business realities into code

Jamie Clark, OASIS

PIDX Digital Transformation Series

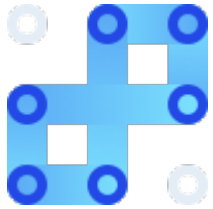
June 2023

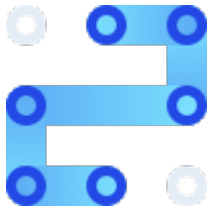
- OASIS is a global internet standards consortium, hosting open standards and open source projects since 1993
- We are accredited by ISO, ITU, JTC 1, ANSI and the European Commission, among others
- OASIS co-hosted the original ebXML project with the UN/ECE, the start of most modern e-commerce (PIDX was there too)
- OASIS hosts many successful standards for e-commerce, cybersecurity, egovernment and blockchain transactions

Today's topic: Using XML and similar tools

XML is a key encoding notation for transacting.

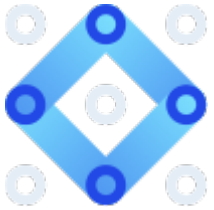
JSON (and YAML, UML, and various blockchain methods) may also work

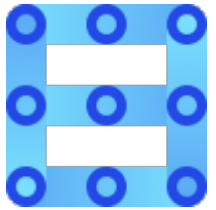


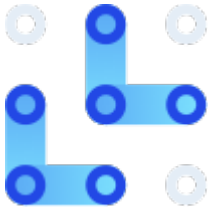


Your transactional data standards should be open, vendor neutral, widely-tooled, and interoperable

Look under the hood of your data standards, in *whatever* notation, to find their embedded business models







1

It's a set of choices, not a smackdown

XML is good for some things, JSON for others, and both work for some cases.

There are some other options like formal modeling (such as UML) and other markups (like YAML). Blockchain is just another way to represent that business data and those transactions.

All of these should be judged by the same business criteria.

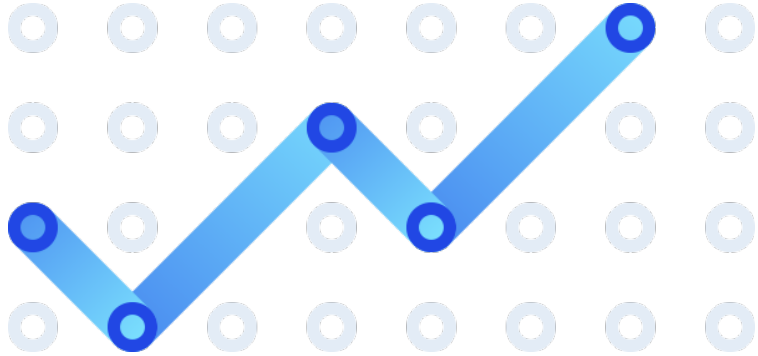
Different notations have different purposes

XML is the dominant notation for documents and business transactions across diverse systems. It describes data, and is more expessive.

JSON is used more for computational instructions, and is tighter. It assumes tight bindings between similar systems.

XML grew out of HTML and EDI (as markup for business content). JSON grew out of Javascript (the programming language).

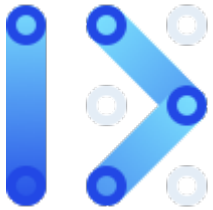
Either can be used in many cases, but they have different "sweet spots."



XML vs JSON relative virtues

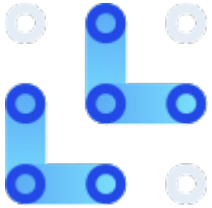
Holman: Horses for courses: A perspective on an XML vs. JSON discussion (2017)

https://www.xml.com/articles/2017/08/06/xml-vs-json-discussion/



Megathread on XML vs. JSON:

http://lists.xml.org/archives/xml-dev/201504/msg00021.html

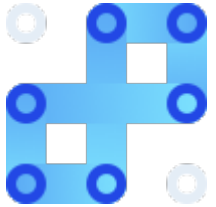


Parsing JSON is a Minefield:

https://seriot.ch/projects/parsing_json.html

It's not a smackdown

- At this writing, legislation, policies, business records and archival data tend to use XML; one-time exchanges conducted in a web browser, and identity tokens, tend to use JSON.
- Sure, there are tool preferences, but many popular data standards are expressed in both notations. It's also possible (with some limits) to transform from one notation to the other
- OASIS UBL TC, the most widely-used general e-invoicing standard, has both XML and JSON representations. PIDX, UBL, OAGi, and the UN/ECE core components library all grew out of EDI and the ebXML project.



2

Regulators and power users care about open, fair, vendor-neutral standards.

Once you implement a standard in your business processes, conformance is a sunk cost.

If you're locked in to a single vendor, and can't re-use the data across customers, transactions and platforms, it's lost sunk costs.

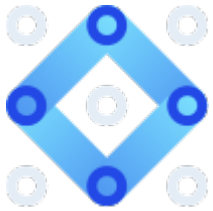
If you can't test it, you won't know if you're getting the right data.

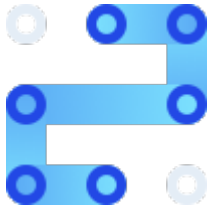
And often, you can't eat just one: you're probably using multiple standards in combinations.

One open standard → lots of implementations

- OASIS Universal Business Language (ISO/IEC 19845) is a common library of order-to-cash business documents (such as purchase orders, invoices, logistics and supply chain management for procurement). It's used in PEPPOL, the pan-European public procurement platform.
- Clear conformance clauses + no exclusive implementation= widely tooled.

What we we looking for in our standards?





- Modularity good
- Alternatives good
- Interoperability good
- Clear conformance clauses good
- Exclusive dependencies, usually bad

3

Look under the hood!

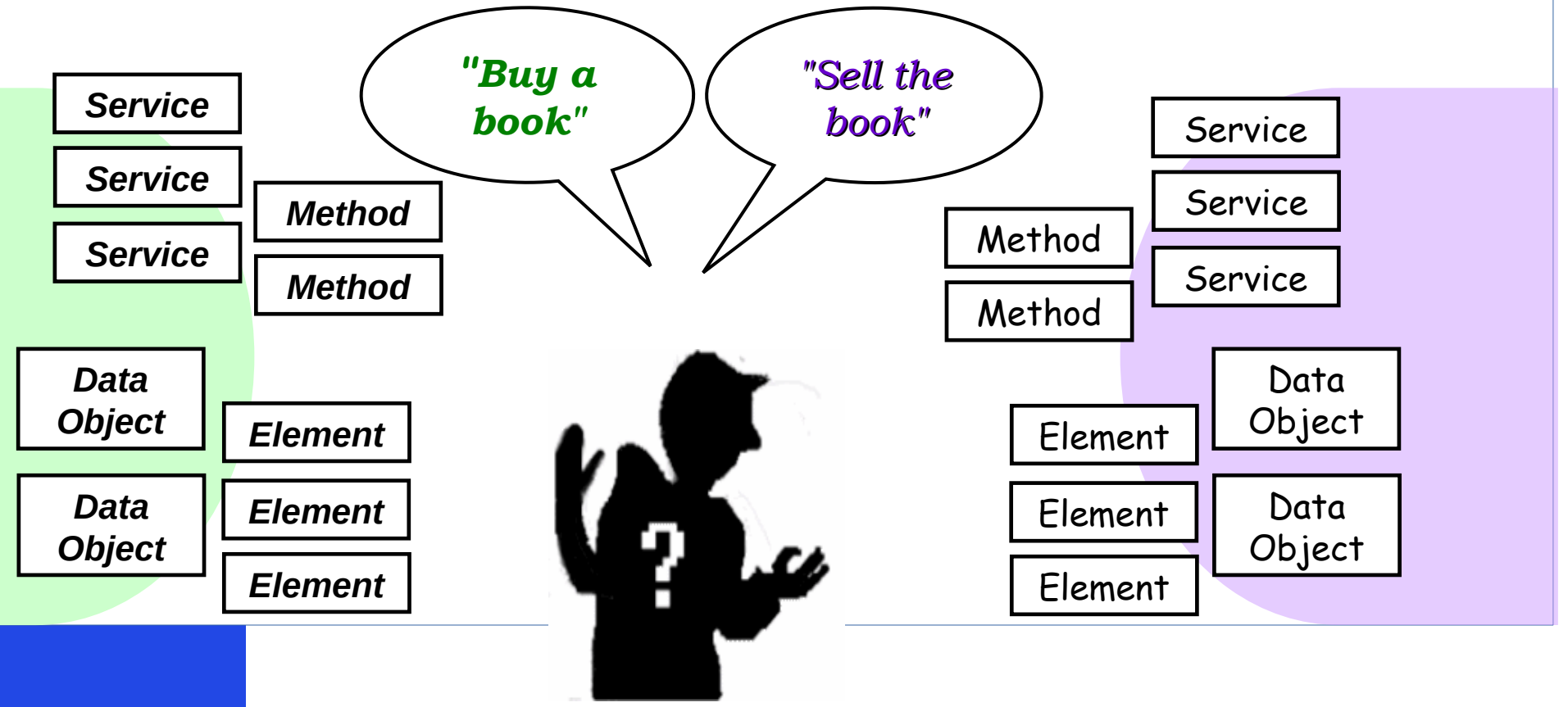
Data standards embed business choices and models!

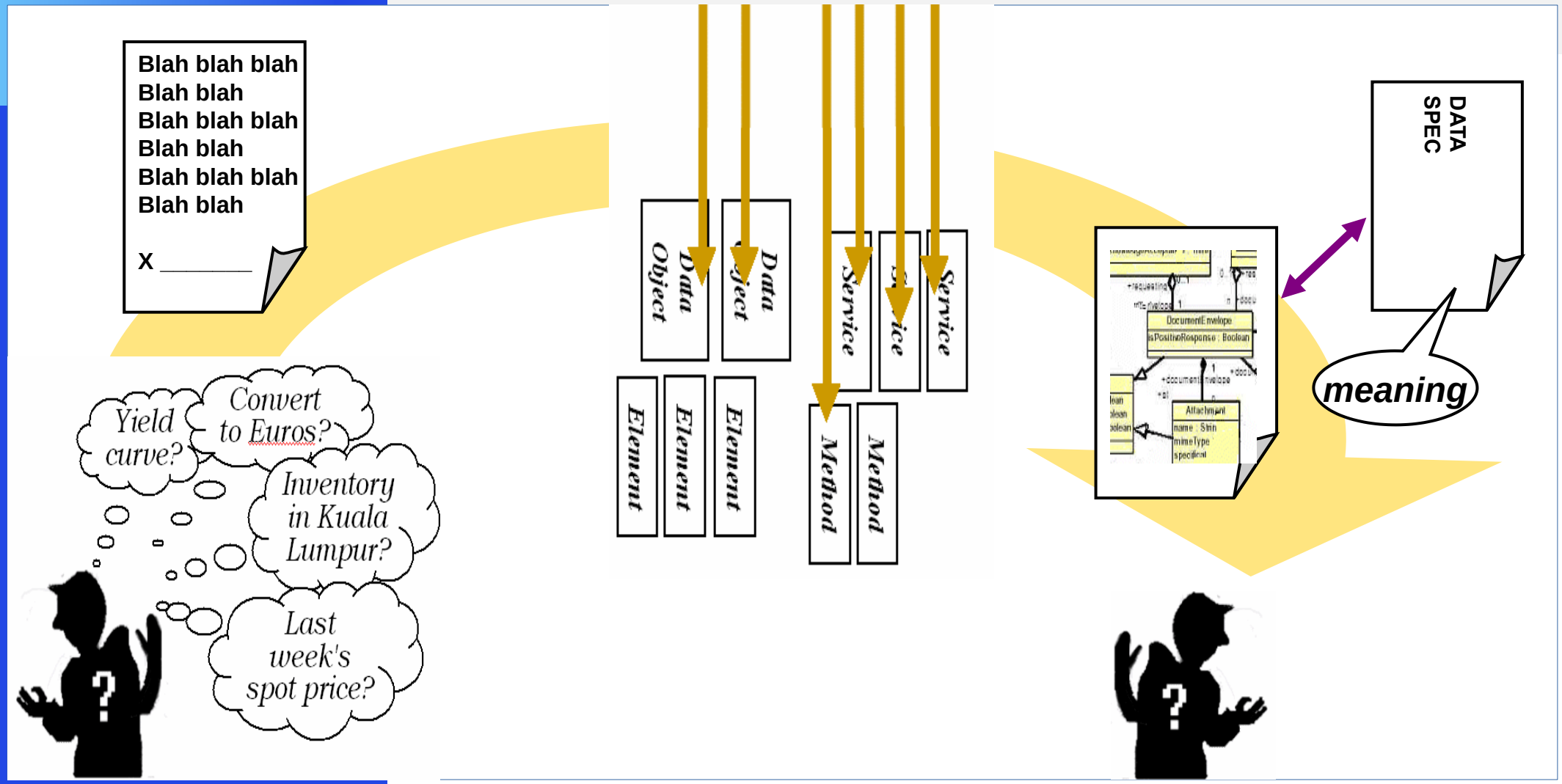
Look at the data structure, not just at the serialization in code.

The Boeing analysts and the Bag of LEGOs

The Banff experiment:

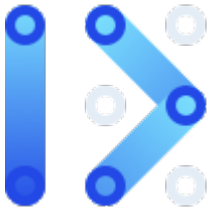
Amazon methods = eBay methods

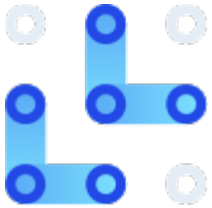



Blah blah blah
Blah blah
Blah blah blah
Blah blah
Blah blah blah
Blah blah
X _____
Yield curve?
Convert to Euros?
Inventory in Kuala Lumpur?
Last week's spot price?
Data Object
Data Object
Element
Element
Element
Service
Service
Method
Method
DocumentEnvelope
isPositiveResponse : Boolean
Attachment
name : Strin
mimeType
specificat
DATA SPEC
meaning

Think of the data standard as the stationary and the envelope for the transaction

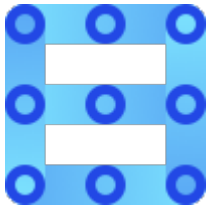
- Can this set of defined terms, and this set of structured data relationships, faithfully represent my transaction?
- Can I easily verify and audit that, when needed?
- Is it relatively easy to render and send a conformant set of messages?
- Does this method allow me to adequately secure the data (including our transactional history, the identities of counterparties, and any out-of-band data), in a way we can understand and prove up later?





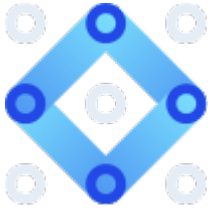
OK, just a bit of comparison to blockchains, for reliable business transacting

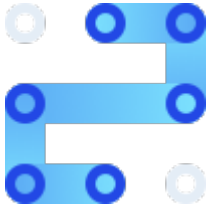
- It's early days.
- Multiple notations? Yup.
- Complex tool issues,
- Single-vendor lock-in? Often. But of the multiple successful notations, several (including Enterprise Ethereum Alliance and Hyperledger) are taking a standards-like, multi-vendor approach and avoid lock-in.
- I'm biased: EEA works with OASIS.
  https://entethalliance.org/EEACommunityProjects/



… just a bit: Blockchains in 2023

- Caution: Open source is not the same thing as open standards, from a pro-competitive point of view. Every 'free' tool is not multi-vendor. (See OMB Circular A-119 in the US, or the WTO Technical Barriers to Trade Agreement.)
- There's genuine progress on DLT standardization at ISO TC 307:
  - Vocabulary standard out: ISO 22739
    https://www.iso.org/standard/82208.html
  - Work on cross-chain interop: ISO/AWI TS 23516
    https://www.iso.org/standard/82098.html
- The Princess Bride rule still applies in blockchain.





Questions?

jamie.clark@oasis-open.org
@jamiexml
@jamiexml@infosec.exchange

www.oasis-open.org