



OCLC Research

Private sector perspectives: Commentary

Brian Lavoie
Consulting Research Scientist
OCLC Programs & Research

DigCCurr 2007
Chapel Hill, NC
April 20, 2007

Themes

- Standards & standardization
- Openness & persistence
- “Private sector” & digital curation

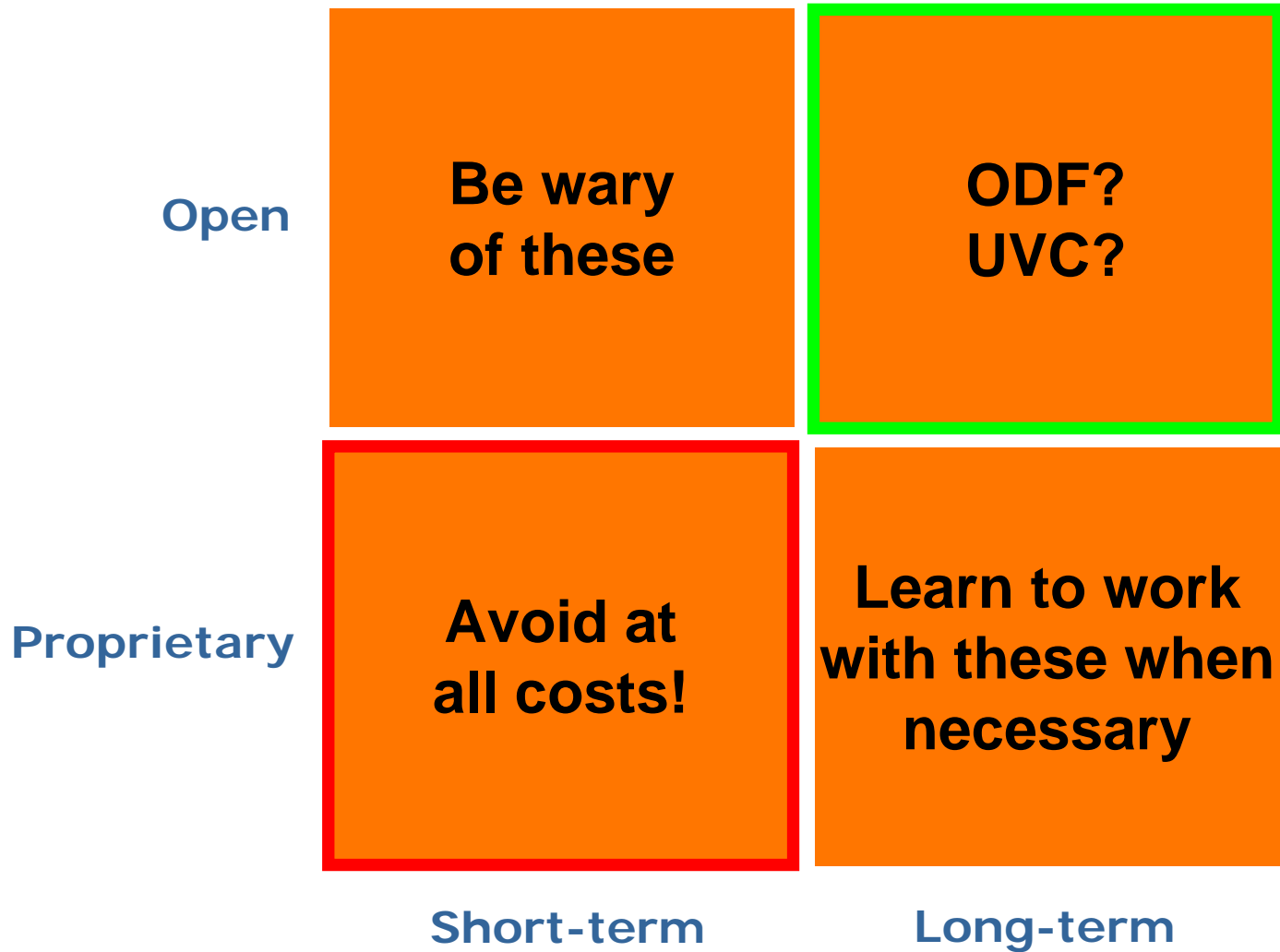
Standards

- Importance of standards for digital curation:
 - **OAIS**: standard view of preservation repository functionality, information packages
 - **ODF**: open format for representing Office-type materials (docs, spreadsheets, presentations)
- “Standards are blueprints” (Tom)
 - Analogy to house-building: “functional and safe house”
 - Standards: foundation for building functional and safe digital curation systems
- But standards suggest consensus:
 - Articulation of expected outcomes
 - Emergence of best practice

Openness and persistence

- Standards vital for reliable digital curation, BUT ...
- Tom: standards can be open or proprietary
 - **Proprietary:** information locked up in proprietary data formats; future control & curation of data ceded to vendors whose incentives & long-term existence might be uncertain
 - **Open:** preserve control of the data, which preserves ability to manage data over long-term (e.g., ODF)
- Raymond: technologies can become obsolete or inaccessible
 - **Transitory computing environments:** today's innovative technology is tomorrow's legacy system; migration/emulation can impact look, feel, and functionality of original object
 - **Persistent computing environments:** stable computing platform that can persistently support applications (decoders) for display and use of obsolete data formats (e.g., UVC)

Standards & technology space



Persistent, open standards & technologies: challenges

- Who will set directions, establish priorities, drive innovation?
 - Vendor-neutral standards bodies can be slow; need to synthesize range of perspectives
 - Digital curation: articulate features useful/essential for long term preservation (e.g., PDF/A)
- How can we avoid fragmentation?
 - E.g., UNIX: open standard in research community; many different vendor versions
 - E.g., TIFF: extensibility leads to many versions
 - Digital curation: fragmentation increases preservation burden
- Sustain involvement, interest (i.e., make it persistent)?
 - “Public ownership” may lead to underinvestment
 - Support for the standard: off-the-shelf implementations, technical support, enhancements, certification (e.g., Linux/RedHat)

“Private sector” and digital curation

- “Private sector”: third-party provision of digital curation services
 - Some or all of the digital curation process likely to be “outsourced” from local institution to external party
 - For-profits, non-profits, inter-institutional collaborations, systems for local implementation
 - Building and maintaining redundant digital curation capacity unrealistic, inefficient, and unaffordable
- Need trusted, well-understood third-party solutions
 - Based on convergence of demand-side expectations and supply-side capacity & commitments
- Convergence achieved through:
 - Consensus
 - Standards
 - Standardization } Openness & persistence

So ...

- Cultivate consensus across a range of digital curation requirements
 - Reference models, frameworks, case studies & synthesis
- Communicate consensus through engagement in standards **and** standardization efforts
 - Open and persistent standards are the goal
- Achieve third-party service provision that:
 - Matches demand-side expectations with supply-side commitments
 - Is well-understood, interoperable, widely applicable
- Robust, economical digital curation