Does the world need more XML standards?

Francis Cave

Chair, ISO/IEC JTC 1/SC 34

Document description and processing languages

ISO/… what?

- ISO International Organization for Standardization
- IEC International Electrotechnical Commission
- ISO/IEC JTC 1
  - The one (and only) Joint Technical Committee of ISO and IEC
  - Formed in 1987 by the merger of separate ISO and IEC Technical Committees, to collaborate on ICT standardization
- ISO/IEC JTC 1/SC 34
  - A Sub-committee of JTC 1
  - Document Description and Processing Languages

It all started fifty years ago…

1967 Early ideas on separation of document content from document format (William W. Tunnicliffe, Graphical Communications Association)

1969 Goldfarb, Mosher and Laurie invent GML at IBM

1971 GCA established the “System X” committee under Tunnicliffe, which eventually led to the development of GenCode®

1973 Initial release of GML as an IBM product

1978 IBM published the DCF GML User’s Guide, written by Goldfarb, which became the basis of the first working paper submitted for standardization of what would become SGML

History of SC 34: https://jtc1historyblog.wordpress.com/sc-34/

1980s: SGML becomes ISO 8879

1982 International STM Group hosted the first conference on SGML in Amsterdam; Goldfarb and the ANSI SGML committee sought European support for the development of SGML as an International Standard

1983 GCA publish Standard 101-1983
GenCode and the Standard Generalized Markup Language

1984 SGML submitted to ISO;
an international Working Group was formed to work on drafting, convened initially by Tunnicliffe, then by Dr James D Mason:
ISO TC 97/SC 18/WG 8

1986 Publication of SGML as ISO 8879

1987 US DoD MIL-STD-1840A became the first major adoption of SGML
ISO TC 97 subsumed within ISO/IEC JTC 1

1990s: A time of seismic change in the ICT standards development world

1993	Internet Engineering Task Force starts to function as a standards body
1994	World-wide Web Consortium founded
1997	XML 1.0 published by W3C
1997	JTC 1 re-organized and SC 18 ‘Text and office systems’ disbanded
1998	JTC 1/ SC 34 ‘Document description and processing languages’ established, taking ownership of past and current projects of the former SC 18/WG 8

1990s: Business as usual

1991 ISO/IEC 9541-1: Font Information Interchange
– Extension to font resource architecture

1995 ISO/IEC 10180 Standard Page Description Language (SPDL)

1996 ISO/IEC 10179 Document Style Semantics and Specification Language (DSSSL)

1997 ISO/IEC 10744 Hypermedia/Time-based Structuring Language (HyTime)

1999 Correction made to ISO 8879 to align with W3C XML

2000 ISO/IEC 15445 HyperText Markup Language (HTML)

2001 ISO/IEC 13240 Interchange Standard for Multimedia Interactive Documents (ISMID)

2000s: SC 34 looking for a new focus…

Topic Maps

2000	ISO/IEC 13250 Topic Maps Revised in 2003; replaced by ISO/IEC 13250-2 and 13250-3
2006	ISO/IEC 13250-2 Topic Maps – Data Model
2007	ISO/IEC 13250-3 Topic Maps – XML Syntax Revised in 2013
2009	ISO/IEC 13250-4 Topic Maps – Data Model
2010	ISO/IEC 13250-6 Topic Maps – Compact Syntax
2011	ISO/IEC 19756 Topic Maps – Constraints Language (TMCL)
2015	ISO/IEC 13250-5 Topic Maps – Reference Model

2000s: SC 34 looking for a new focus…

Document Schema Definition Languages (DSDL)

2003 ISO/IEC 19757-2 **RELAX NG**
Revised 2008

2006 ISO/IEC 19757-3 **Schematron**
Revised 2016, currently under further revision

2006 ISO/IEC 19757-4 **Namespace-based Validation Despatching Language (NVDL)**
Corrected in 2008

2008 ISO/IEC 19757-8 **Document Semantics Renaming Language (DSRL)**
Corrected in 2011

2008 ISO/IEC 19757-9 **Datatype- and namespace-aware DTDs**

2009 ISO/IEC 19757-7 **Character Repertoire Description Language (CREPDL)**
Corrected in 2015

2011 ISO/IEC 19757-5 **Extensible Datatypes**

2011 ISO/IEC 19757-11 **Schema Association**
Developed in association with W3C; the <?xml-model …?> PI

2000s: SC 34 learns to collaborate…

Office document file formats – the main published standards:

2006 ISO/IEC 26300 OpenDocument Format (ODF) v1.0

2008 ISO/IEC 29500 Office Open XML File Format (OOXML), in four Parts:
ISO/IEC 29500-1 OOXML – Fundamentals
ISO/IEC 29500-2 OOXML – Open Packaging Conventions (OPC)
ISO/IEC 29500-3 OOXML – Markup Compatibility and Extensibility (MCE)
ISO/IEC 29500-4 OOXML – Transitional Migration Features

2012 ISO/IEC 26300/Amendment 1 - to align with ODF 1.1

2015 ISO/IEC 26300 OpenDocument Format (ODF) v1.2, in three Parts:
ISO/IEC 26300-1 ODF – OpenDocument Schema
ISO/IEC 26300-2 ODF – OpenFormula
ISO/IEC 26300-3 ODF – Packages

ISO/IEC 29500-3 OOXML – Part 3: MCE (major revision)

Recent work in SC 34

Other recently-published standards

2014 ISO/IEC TS 30135 EPUB3, in seven Parts (aligned with EPUB 3.0)

2015 ISO/IEC 21320-1 Document Container File – Part 1: Core

2016 ISO/IEC 19757-3 Schematron (new edition, now being revised further)

Current SC 34 structure and work highlights

WG 4

- Maintenance of ISO/IEC 29500, especially revision of Part 2 OPC
- Development of ISO/IEC 30114-2, using ISO/IEC 19757-7 CREPDL to specify character repertoire constraints on OOXML documents

WG 6

- Maintenance of ISO/IEC 26300 in collaboration with OASIS ODF TC

JWG 7

- Development of ISO/IEC TS 22424-1/2 EPUB3 Preservation
- Revision of ISO/IEC TS 30135 EPUB3

WG 8

- Revision of ISO/IEC 19757-3 Schematron and ISO/IEC 19757-7 CREPDL

Future work for SC 34…

2018 Study Group on Document Semantics Support:

Looking at gaps in the standards available for describing documents

Some of the use cases that are being explored:
- collaborative authoring of documents
- document publishing workflows
- recording document evolution
- recording patterns of document usage
- document archiving and discovery
- documents in a machine learning environment

Why do we need XML standards?

- Responding to demand from implementers/users…
  - To improve the quality of implementations
  - To encourage (consistent) implementation
  - To promote interoperability between implementations
- In other words…
  - To protect users from poor quality, inconsistent implementation, lack of interoperability
  - To protect implementers from poorly-defined specifications, leading to increased costs and higher risks of failure in development

Why do we need international standards?

- To promote equality of experience for both standards implementers and users across international boundaries
  - Standards should be implementable and usable anywhere
  - Implementers and users should not be penalised by reasons of geography, geopolitical status, language or culture
  - Implementers should not be penalized for not having been able to participate fully in the standards development process
  - As well as being useful, standards need to be intelligible to implementers everywhere

What standards *should not try to* do… (in my opinion, and as a generalization)

- Create new technologies…
  - for which there is no known demand
  - for which there is a demand, but alternatives already exist
  - for which there is a demand, but the technology is unproven

Thank you!

Francis Cave

francis@franciscave.com