

UOML

The best way for
Document Interoperation

Alex Wang
Founder/Chairman
Sursen Co.

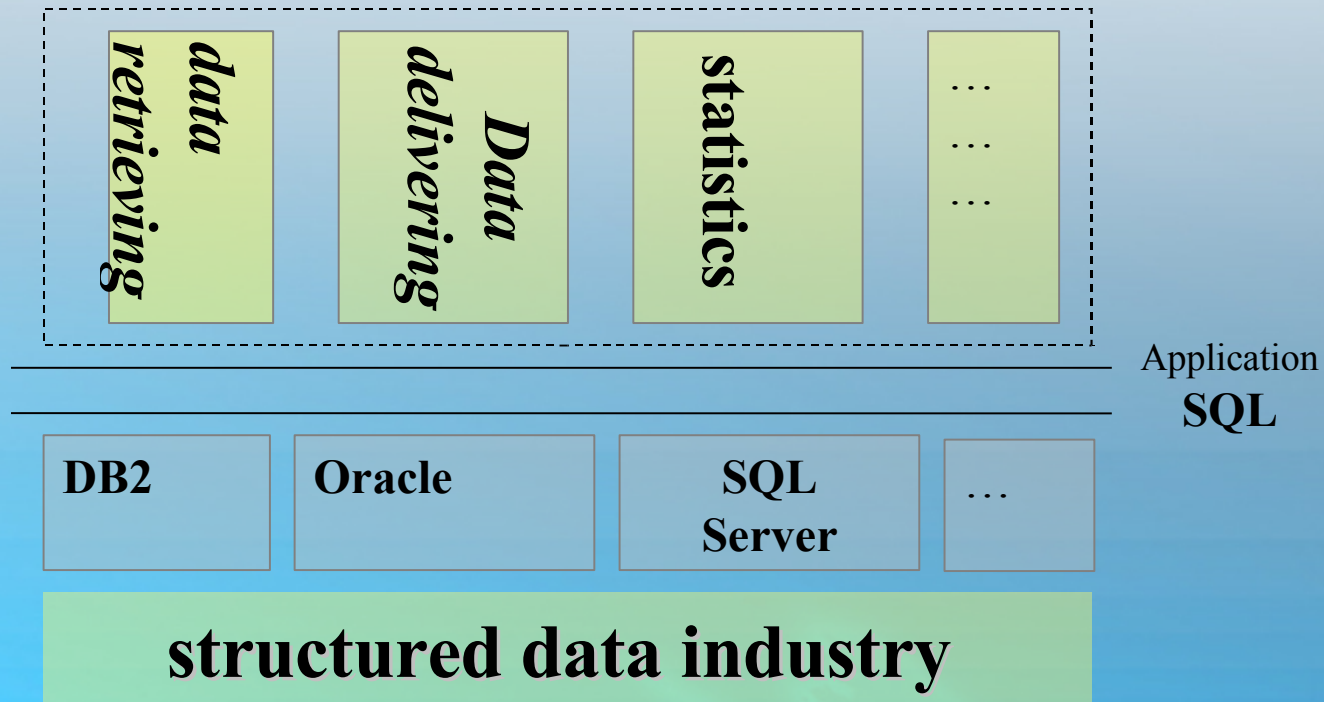
Puzzles on Document Interoperability

- Document hereby refers the content that can be printed on paper
- Paper is still the only interoperation platform till now
 - Different kinds of pen can write on same paper
 - Pencil, brush pen, black pen, red pen etc.
- People pay much money for computer & software, but document created by one software can't be processed by another one
 - Even MS Word can't process MS Excel file

Storage format standards not a solution

- Previous efforts are limited to document storage format standards
- Industry practices over the past decade proved it's not a solution
- Simple format, e.g. TXT, can't meet the need of various software
- For complicated ones, e.g. PDF, only a few companies among hundreds of thousands software companies all over the world, have enough resources and professional technology required to handle it
- Block innovation
- Affect performance optimization
- Impact portability and scalability

Study from structured data industry

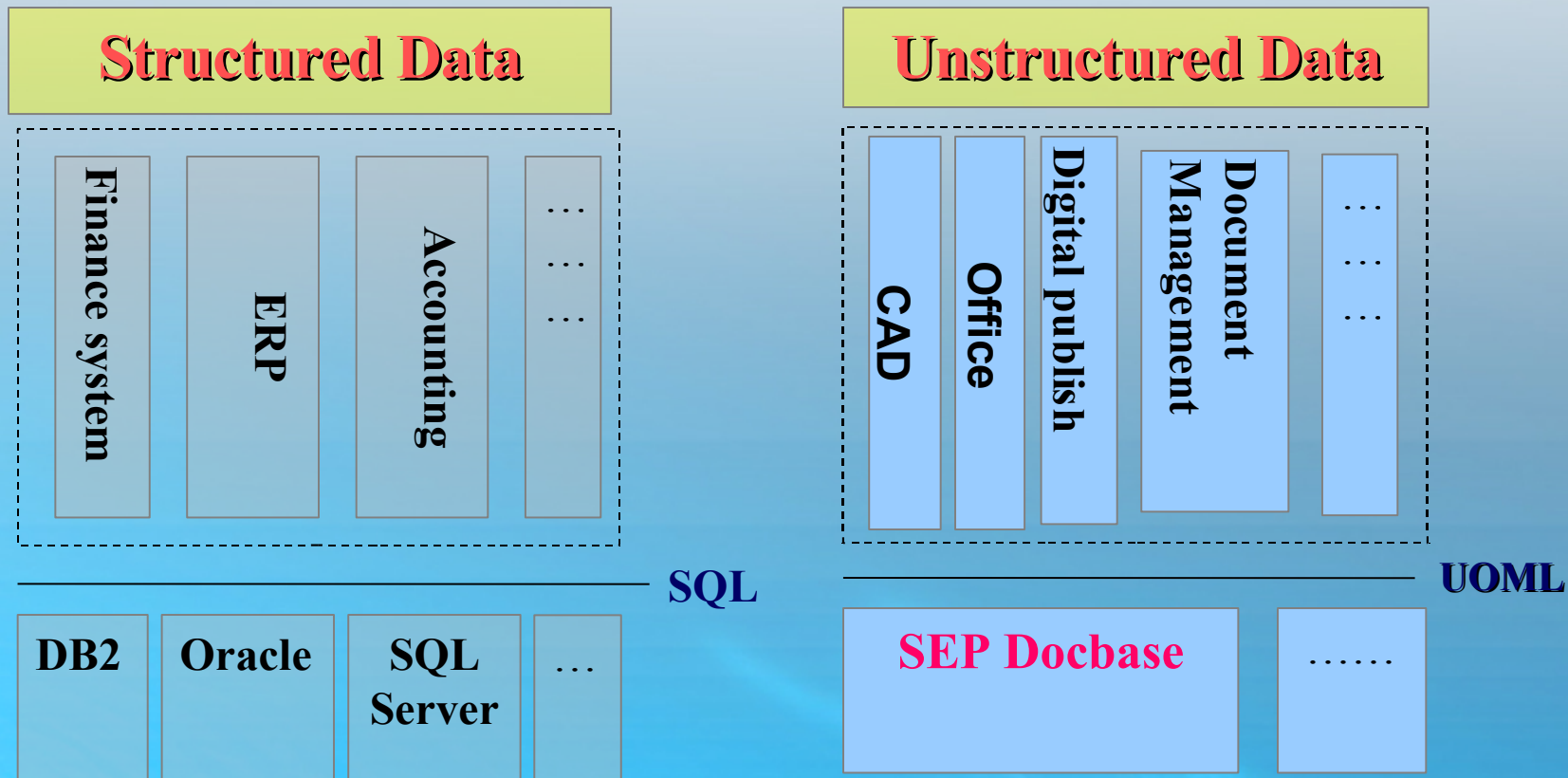


Storage format standards step into operation standard

SQL vs. DBF

- DBF was market standard of database storage format
- Why not continue to improve DBF format? Why SQL?
- High performance database(Oracle,DB2...) uses very complicated storage format, e.g. it is storage media related
- Is it possible that every application can read/write this format itself?
 - Few companies are able to parse such complicated format
 - It is certain that the data will be destroyed by bugs of some software when they save data
- SQL avoids this problem
 - Although to implement the operation is skillful and difficult, a request for the operation is straight forward to any programmer
 - The storage of data is maintained by professional company to ensure the integrity of data

From SQL to UOML



SQL ↔ UOML

DBF ↔ PDF

Database ↔ Docbase

The best way for document interoperation

About UOML

- Abbreviation of Unstructured Operation Markup Language
- Based on XML, cross-platform, programming language independent, application independent
- Define general document operating functions
- Various applications can do document-related operations through UOML call, including create, edit, delete, search, retrieve docbase, docset, document, page, layer, text, graphics, images, font, metadata, hyperlink, bookmark etc.

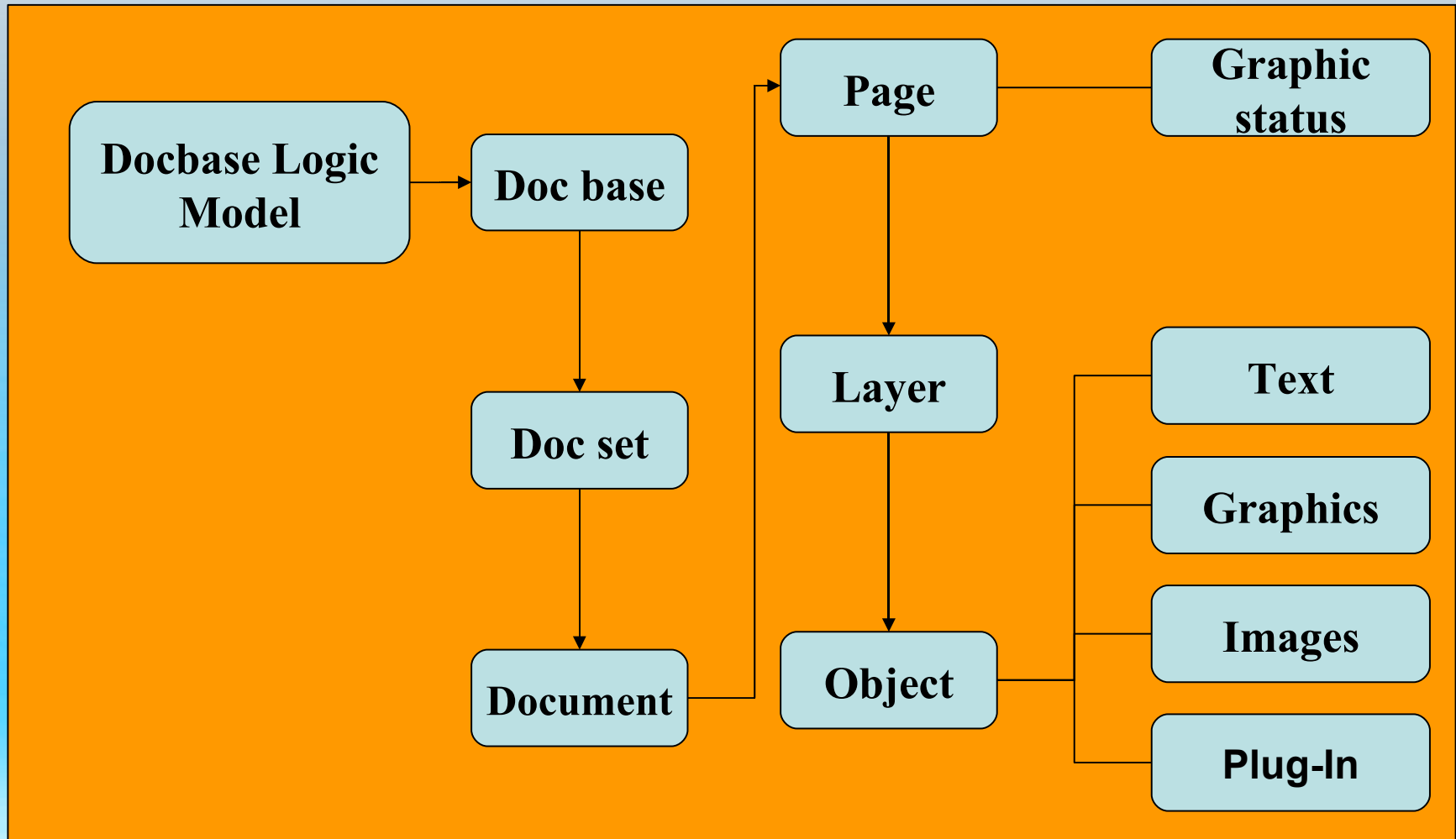
UOML vs. ODF

- UOML deals with information after typesetting, ODF deals with information before typesetting
- UOML limits functions to abstract from paper, ODF has rich functions
- UOML defines operation interface, ODF defines storage format
- UOML and ODF are designed for different purpose, resolve different problems

Docbase: implement of UOML

- Docbase is the foundational software platform that implements UOML
- Integrate functions such as document organization, page description, information security, index and search, content extraction, fonts management, storage management, plug-in mechanism, and script description etc.

Docbase Logic Model



Benefits and Value

- Different software can process the same document
- Information can flow smoothly through all stages
- Enable industry division and avoid repeated development
- Enable the editing and use of complicated documents
- Organization and management of multiple documents
- Open standard breaks monopoly
- Provide more opportunities for small and medium companies
- Nurture a new industry that is larger than database industry

UOML overview

- Expressed in XML
 - Independent of programming languages
 - Independent of platforms
 - Based on open standards
- Define the abstract structure of documents
 - The structure frame should have high universality
 - The structure is based on paper-abstraction, mature for use
- Define the abstract operations for documents
 - Only basic operations are defined
 - The structural restriction is defined by the document structure.
- There should be no limit to physical storage and implementation

Operations of UOML

- UOML_OPEN
- UOML_CLOSE
- UOML_GET
- UOML_SET
- UOML_INSERT
- UOML_DELETE
- UOML_USE
- UOML_RET
- UOML_SYSTEM

UOML Object (partial)

- UOMLO_POINT
- UOMLO_ARC
- UOMLO_BEZIER
- UOMLO_CIRCLE
- UOMLO_LINE
- UOMLO_RECT
- UOMLO_ELLIPSE
- UOMLO_ROUNDRECT
- UOMLO_COLOR_CMYK
- UOMLO_SUBPATH
- UOMLO_PATH
- UOMLO_COLOR_RGB
- UOMLO_IMAGE
- UOMLO_TEXT
- UOMLO_ARCSEG
- UOMLO_LINESEG
- UOMLO_BEZIERSEG
- ...

UOML Alliance

- Formed voluntarily by enterprises, units, organizations and individuals that abide by UOML standard
- Committed to document interoperability through open standards
- Documents can be interchangeable and interoperable among software developed by all members of UOML Alliance
- Provide free license and technical support to assure the wide and successful adoption of UOML standard

The logo for UOML, consisting of the letters 'UOML' in a bold, red, sans-serif font. The letters are outlined in black, and there are horizontal lines above and below the letters, creating a stylized, blocky appearance.

Recruit members to form UOML TC

- UOML Birds-of-a-feather Session at luncheon
- E-Mail: alexwang@sursen.com
- Tel: (408)416-1266 (before May 16th)
+86-13701193777 (after)
- MSN: dlwang@shusheng.net