



UNIVERSITY OF CALIFORNIA, BERKELEY  
SCHOOL OF INFORMATION

# INFO 290-5 | Fall 2013

## “The Future of EBooks”

### The Process Space for Creating eBooks

Robert J. Glushko  
[glushko@berkeley.edu](mailto:glushko@berkeley.edu)  
@rjglushko

Scott Abel  
@scottabel

17 September 2013



# Agenda

- Conversion of existing books – digitization
- Single-sourcing vs hand-crafting content
- Individual authors vs collaborative authoring
- Working effectively with publishers, vendors, consultants



# Conversion of Existing Books

- Digitization is the conversion of analog information into digital information
- Conversion is diagnosis
- Converting content
  - back catalogs vs legacy content
- Conversion models
  - in-house vs out-sourced
  - automated conversion
- Conversion challenges
  - file types, authoring decisions, versions



# Conversion of Existing Books

- Legacy conversion challenges
  - file types
    - so many file types, so little time
    - Quark, InDesign, MS Word
  - formats
    - so many formats, so little time
      - text vs PDF vs RTF vs HTML vs other
  - versions
    - wrong version, wrong content



# Conversion of Existing Books

- Legacy conversion challenges
  - authoring decisions
    - inconsistent formatting
      - using styles vs having fun with fonts
      - word spacing
    - page references – “see page 16”
    - direction references – “above”
    - hyphens, emphasis, special characters
    - hyperlinks



# Conversion of Existing Books

- Legacy conversion challenges
  - design decisions
    - complexity of source content
      - images
      - tables
      - sidebars
      - endnotes and footnotes
      - math calculations / computer code
      - number of complex elements



# Conversion of Existing Books

- Legacy conversion challenges
  - target destinations
    - platforms
      - iOS vs Android vs Kindle vs \_\_\_\_\_
    - device capabilities
    - screen real estate



# Conversion of Existing Books

- Legacy conversion challenges
  - automated conversion software
    - how clean do you want it?
      - medical texts vs white papers
  - how good is the conversion tool?
  - first pass automation





# Conversion of Existing Books

- Legacy conversion challenges
  - post-conversion cleanup
    - manual process
    - requires skilled clean-up staff
      - content review
      - device-specific review and testing
      - orientation testing



# Single-Sourcing vs Hand-Crafting

- Handcrafting content
  - 1985 desktop publishing and the Macintosh
  - Aldus PageMaker
    - typesetting and graphic design
    - writers became defacto designers
    - just because you have all those fonts...



# Single-Sourcing vs Hand-Crafting

- Handcrafted content is
  - Fun for some, but
    - Inefficient
    - Error-prone
    - Costly
  - Impossible to meet increasing demands
    - customers: on-demand information
    - company: doing more with less staff
  - Not scalable



# Single-Sourcing vs Hand-Crafting

- Single-source publishing is
  - content management method
    - supports content reuse
    - write it once, use it often
    - designed to tackle multi-channel needs
    - reduces potential for errors
    - supports consistency of content
    - speeds time-to-market
    - highly scalable



# Single-Sourcing vs Hand-Crafting

- Single-source publishing led way to
  - customized content
    - for individual customers
    - to meet output requirements
  - dynamic content
    - assembled on demand in response to
      - user requests, behavior, transactions
      - proximity/location
      - other factors



# Single-Sourcing vs Hand-Crafting

- Structure vs community
  - XML single-source vs wikis and community
    - DocBook and DITA
      - open standards
      - common structure for consistent creation, sharing, reuse of content
      - DocBook better for creation
      - DITA for reuse



# Single-Sourcing vs Hand-Crafting

- Structure vs community
  - XML single-source vs wikis and community
    - Benefits
      - WYSIWYG authoring experience
      - Supports crowd-sourcing
    - Drawbacks
      - content locked in; no easy way to get out
      - similar problems as legacy content



# Single-Sourcing vs Hand-Crafting

- Structure vs community
  - XML single-source vs wikis and community
    - Putting them all together
      - create structured XML content
        - modular, well-structured, separate of format
      - in a structured content tool or wiki
      - encourage customer feedback and improvement





# Individual vs collaborative authoring

- Collaborative authoring
  - Multiple contributors working collaboratively
  - Overseen by editor or team of editors
  - Permission-based access
  - Recursive process
    - each change prompts others to make additional changes/improvements
  - Messy if goals, rules and responsibilities are not clearly defined and communicated



# Individual vs collaborative authoring

- Collaborative authoring
  - Variety of strategies
    - single-author writing for a team
      - incorporated feedback from others
    - sequential single-author writing
      - writers write their portions
      - pass content to next writer, editor



# Individual vs collaborative authoring

- Collaborative authoring
  - Variety of strategies
    - parallel writing
      - group divides project
      - horizontal division – each writer works on their content at the same time as other writers are working on theirs
      - stratified division – responsibility is divided by role: editor, author, QA, etc.



# Continuous vs editions vs annotation

- Continuous authoring
  - wiki approach – Wikipedia
  - community/authors/editors constantly improve content
  - edits can be made in content or held for moderation
  - content is fresh and up-to-date faster than other approaches



# Continuous vs editions vs annotation

- Editions

- a form or version of a published text
- content is only as fresh as it was on the day it was published
- customers don't benefit from new discoveries, error detection, etc. until next edition is published
- not the best strategy if content is about rapidly-changing subject matter



# Continuous vs editions vs annotation

- Disciplinary
  - contributors can annotate content
  - Bob example Discipline of Organizing