Sustainable infrastructures for content services of special interest publishing houses

Svenja Hagenhoff (Presenter), Jörn Fahsel
Friedrich-Alexander-University Erlangen-Nürnberg
Institute for the Study of the Book, Professorship for E-Publishing and Digital Markets

emma conference, University of Hamburg, June 2015
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Publishing houses as content (service) provider

• Since hundreds of years Information, knowledge and entertainment is distributed to society by written and printed media as books, pamphlets, newspapers or magazines (e.g. Eisenstein 1979, Burke 2011, Gitelman 2014)

• If content shall fulfill this important function
  o it needs to be presented in the right form depending on the reader’s reception situation and
  o readers (or users) need to find the right content and to get access to it
Information technology is needed

• For long times in the publishing industry technology was only needed at the reproduction state (›printing‹) but not at the stage of collecting, editing, bundling and archiving of content (›publishing‹ mainly as intellectual work)

• Nowadays technology is needed at all stages (e.g. Knaf/Hünemörder 2012; Koblinger 2002, Hill 2010)
  o typing and editing the content
  o bundling and formatting the content: especially for providing reader’s individual content services
  o reproducing the content (printed as well as digital)
  o distributing the content and the media (storage management and logistics)
  o accessing and using the content

Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Data as part of IT

• Data used for description of
  o certain entities of the real world like products, customers, authors, rights
  o the relationship between them
  o and states like stock of products in the warehouse
  → utility data, to work with

• Data used to steer processes
  o Use case 1: select all articles of one author to bundle a reader
  o Use case 2: select all articles bought by one specific customer
  o Use case 3: deliver all new papers to a certain customer
  → control data

• Data as part of the product
  → raw material, assets

Data is a crucial resource
It needs to be managed
This is a ‹book›

Data as raw material and part of the book

Needed for production

Data to describe the book

Needed for distribution
Software applications as part of IT

• Software is a tool

• It supports workflows within a business, e.g.
  o ›from manuscript to book‹
  o ›from license request via contract conclusion to license in-payment‹

• Software works with data

analyzing the tasks within an organization, their sequences and dependencies as well as the needed and produced data is the key factor concerning having the right software support
Data & software for producing and distributing content

Application programming interface

Which **person** or **service** is allowed to access which **data** at what **conditions**?

- bundling
- formatting
- delivery
- invoicing

**Document**
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Book-On-Demand: Entwicklung eines Konzepts zur Integration der Buchweiterverarbeitung in einen digitalen Workflow</td>
<td>Thielen</td>
<td>2003</td>
<td>Monograph (Chemnitz)</td>
</tr>
<tr>
<td>Change Management in Fachverlagen: Am Beispiel der Einführung eines Redaktionssystems</td>
<td>Heinold, E.; Hagenhoff</td>
<td>2010</td>
<td>Brancheninformationen der Deutschen Fachpresse</td>
</tr>
<tr>
<td>Management kreativitätsintensiver Prozesse Theorien, Methoden, Software und deren Anwendung in der Fernsehindustrie</td>
<td>Becker, J.; Schwaderlapp, W.; Seidel</td>
<td>2012</td>
<td>Monograph (Berlin)</td>
</tr>
</tbody>
</table>
## Publications with focus on data management

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifikation und technische Bewertung von integrierten Datenverteilungsvarianten für eine effiziente Mehrfachnutzung multimedialer Medieninhalte</td>
<td>Benlian</td>
<td>2004</td>
<td>Working paper (München)</td>
</tr>
</tbody>
</table>
# Publications with focus on application systems

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>IT-Verhalten und Defizite in KMU</td>
<td>Meyer/Tirpitz/ Koepe</td>
<td>2010</td>
<td>Monograph (Köln)</td>
</tr>
<tr>
<td>Softwareunterstützung für die Bereitstellung klassischer Medienprodukte und -dienstleistungen</td>
<td>Hess/Dörr</td>
<td>2012</td>
<td>Working paper (München)</td>
</tr>
</tbody>
</table>
# Publications with focus on ‚reference models‘

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Year</th>
<th>Published in</th>
</tr>
</thead>
<tbody>
<tr>
<td>Referenzmodellierung für Buchverlage</td>
<td>Tzouvaras</td>
<td>2003</td>
<td>Monograph (Göttingen)</td>
</tr>
<tr>
<td>Generische Bücher - ein graphentheoretisches Modell zur logischen Strukturierung von Büchern in on-Demand-Publikationsprozessen</td>
<td>Kreulich</td>
<td>2002</td>
<td>Monograph (Chemnitz)</td>
</tr>
<tr>
<td>Referenzmodell für technische und organisatorische Abläufe bei der international verteilten Medienproduktion</td>
<td>Engelbach/Delp</td>
<td>2006</td>
<td>Working paper (Stuttgart)</td>
</tr>
<tr>
<td>Ein Referenzmodell für die Herstellung von Fachmedienprodukten.</td>
<td>Delp</td>
<td>2006</td>
<td>Monograph (Heimsheim)</td>
</tr>
<tr>
<td>Referenzmodellierung technologischer Hauptprozesse der grafischen Industrie</td>
<td>Reiche</td>
<td>2008</td>
<td>Monograph (Chemnitz)</td>
</tr>
</tbody>
</table>
Synopsis

- Literature review continuously done (English and German media)
- Selection is done very strongly (not everything connected with ›technology & media‹, focus is data management, workflow modelling and application landscapes)

- Not very much literature concerning the topic ›IT and IT infrastructures in the publishing industry‹
- Lack of reference models (workflows, data, system landscape)
- Lack of basic literature about industry specific views and therefore a lack of precise wording, esp. concerning ›data management‹

<table>
<thead>
<tr>
<th>Topic</th>
<th>Number of publications</th>
<th>Published between</th>
</tr>
</thead>
<tbody>
<tr>
<td>Workflows and workflow management</td>
<td>9</td>
<td>2003-2010</td>
</tr>
<tr>
<td>Data and data management</td>
<td>6</td>
<td>2004-2014</td>
</tr>
<tr>
<td>Application systems: descriptions of certain systems, existence of systems, implementing processes</td>
<td>8</td>
<td>2001-2013</td>
</tr>
<tr>
<td>Reference models</td>
<td>7</td>
<td>2003-2014</td>
</tr>
</tbody>
</table>
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Basic data of the survey

• Focus of interest: state of the art of software support concerning *customer relationship management* at *publishing houses providing expert content* for branches, professions or hobbies

• Motivation: expert content publishers need to have close contact to their readers to deliver the content they need and to develop suitable new content services

• Qualitative survey in 2013, semi-structured interviews, category system for interpretation
  
  o Part A:
    Meaning of ›focusing on the customer‹ within the publishing industry
  
  o Part B:
    How is the topic ›customer relationship management‹ grounded concerning organization, instruments, and IT infrastructure
  
  o Part C:
    General relevance and progress of the topic within the publishing industry
# Respondents

<table>
<thead>
<tr>
<th>Position of the respondent</th>
<th>Company</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content strategist</td>
<td>Specialist publisher: &gt; 250 employees providing content in the field of law, economics, public administration, finance &amp; taxes, health, transportation</td>
</tr>
<tr>
<td>CEO</td>
<td>Specialist publisher: &gt; 250 employees providing content in the field of logistics, pharma, electronic, engineering, automotive</td>
</tr>
<tr>
<td>Assistant of CEO</td>
<td>Specialist publisher: &lt; 250 employees providing content in the field of law</td>
</tr>
<tr>
<td>CEO</td>
<td>Specialist publisher: &lt; 250 employees providing economic news</td>
</tr>
<tr>
<td>Senior manager</td>
<td>Business consultancy with media industry department</td>
</tr>
<tr>
<td>CEO</td>
<td>Business consultancy with focus on publishing industry</td>
</tr>
<tr>
<td>CEO</td>
<td>System provider</td>
</tr>
<tr>
<td>Product manager</td>
<td>System provider</td>
</tr>
<tr>
<td>Product Manager</td>
<td>System provider</td>
</tr>
<tr>
<td>Manager</td>
<td>System provider</td>
</tr>
<tr>
<td>CEO</td>
<td>Umbrella association</td>
</tr>
</tbody>
</table>
Selected results in a nutshell

• Data about customers (readers as well as advertising companies) are collected

• Collected data is about transactions (what was bought when by whom, who has birthday when);
• Collected data is not about
  o the readers content needs
  o what the reader does with the provided content
    (e.g. how does the attorney work with the paragraphs)
  o the customers complaints
  or this data is not utilized

• Publishers do appreciate that integration of customers knowledge into development of new content services is needed
• But anyhow new content services are ›armchair decisions‹

• Individualized content services are typically not provided due to missing information about the reader / user
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Basic data of the case study

- Focus of interest: *publication workflow* and *design of the IT infrastructure* in publishing houses
- Aim:
  - identifying which data is stored how and were?
  - identifying which working steps are done how?

- Publisher 1 (of about 12 to 15):
  - publishing house providing expert content for the beverage industry
  - about 20 employees
  - products
    - 4 magazines (1 of which is published in 3 languages)
    - 1 newsletter
    - about 120 books
Software application landscape (extract)
Selected results in a nutshell

• Data are stored redundantly
  o content as raw material is separated in offline and online production
  o product data (title of a publication, author,…) are in the production system as well as in the sales system

• There is no integrated databased storage of data

• There are a lot of legacy applications with a lot of interfaces in between

• Content management system as ›digital assembly line‹ is a critical resource
  o dependency from software providers is high
  o lifecycle costs of a new system would not be covered by extra revenues generated by ›digital content services‹
  -> investment is not financeable
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Synopsis of the situation

• Standard software applications are often too big for small and medium sized publishing houses

• Data pools as well as software applications did grow uncontrolled (patchwork approach), they are not result of modelling work

• Using standard software applications causes dependency from software providers regarding a critical resource (e.g. content management systems as ›digital assembly belts‹)

• A balance is need between
  o workflow and data management efficiency
    • employees need to concentrate on editing the content and not on managing data
    • total cost of ownership include cost of manpower and cost of infrastructure
  o Responsiveness towards (future) requirements in the content service industry

• There is a lack concerning reference models as a blueprint or model pattern
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Reference models

- are general models of a certain class of issues
- describe objects, their typical characteristics and the relationships between them

- can be used
  - as design patterns for creating individual solutions efficiently (blue print, recommendatory artifact)
  - as communication means to understand the issue better
  - as means of standardization within a branch → cutting back complexity, uncertainty and avoiding re-inventing the wheel

- It can describe
  - domain specific data, their structure and relationships
  - workflows
  - software application landscapes

See e.g. Fettke, P.; Loos, P. 2003 / Prieto-Díaz, R. 1990.
Model between real world and IT system

**Organisation:**
- People
- Physical objects
- Artifacts

**Model of the organisation:**
- Workflows
- Entities

**IT-System:**
- Data structures
- Functionality
- User interfaces

Similar Horstmann 2011, S. 80.
Example: data modelling

Example: workflow modelling

See for workflow modelling with UML e.g. Fettke, P. et al. 2007.
Outline

1. Motivation

2. IT as basic infrastructure in the digital world

3. IT in media industries: State of the art
   a. State of the art in literature
   b. Empirical analysis 1: survey
   c. Empirical analysis 2: case study

4. Sustainable IT infrastructures as a research topic
   a. Learnings
   b. Reference models as basis for sustainability
   c. Outlook
Remaining research

Case study 1  Case study 2  ...  Case study 15

Distilling reference models concerning publication workflows and data for the publishing industry

- Case study based analysis: spring to autumn 2015, each case 2 days locally, using standard notations
- Deduction of the abstract models autumn 2015 to spring 2016

PhD thesis of Jörn Fahsel
Literature


