Gap Analysis: Methodology, Tool and First Application

Dr. Moritz Gomm
FernUniversität Hagen

PARSE.Insight Workshop
21st - 22nd September 2009
Darmstadt
Agenda

I. Introduction

II. Gap Analysis Framework & Methodology

III. Gap Analysis Visual Tool (Live Demo)

IV. Example: Publishers Survey analysed
Introduction: Aim of the Gap Analysis

“The gap analysis determines ‘the space between where we are and where we want to be’, and serves as a means to bridge that space.”

“A gap analysis is conducted to determine the difference between the actual state and a desired future state of a system.”

Identify gaps in existing and planned infrastructures

- Background
  - Inspiring examples from NASA, eGovRTD2020, e-Science, etc.
  - Profound analysis of EU specific context for e-infrastructures is missing
Agenda

I. Introduction

II. Gap Analysis Framework & Methodology

III. Visual Gap Analysis IT-Tool (Live Demo)

IV. Example: Publishers Survey analysed
Gap Analysis Framework:
Type of Gaps

<table>
<thead>
<tr>
<th>Type of Gaps</th>
<th>Awareness-Gaps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffusion of concept</td>
<td></td>
</tr>
<tr>
<td>Awareness</td>
<td></td>
</tr>
<tr>
<td>Knowledge</td>
<td>=&gt; Knowledge-Gaps</td>
</tr>
<tr>
<td>Implementation</td>
<td>=&gt; Implementation-Gaps</td>
</tr>
<tr>
<td>Commitment</td>
<td>=&gt; Commitment-Gaps</td>
</tr>
</tbody>
</table>

Life-cycle of Data
Gap Analysis Methodology: Type of Gaps

- Type of Gaps
- Awareness
  - Awareness Indicator 1
  - Awareness Indicator 2
  - Awareness Indicator 3
- Long-term Knowledge
- Implementation
- Commitment

Survey (Insight Report)
- Item A
- Item B
- Item C
Gap Analysis Methodology: Transforming Survey into Tree Structure

- 2 expert workshops with focus groups

Survey Design

Tree Structure for the Tool

Domain specific knowledge
Gap Analysis Methodology: Load Survey-Data into Tool

- Survey-Data is loaded into tool
- Performed by tool Administrator
Gap Analysis Methodology: Calculate and Analyze Gaps

Categories:
- Awareness
  - 0.4
- Knowledge
  - 0.2
- Implementation
  - 0.6
- Commitment
  - 0.3

Sub-categories:

Survey Items

Data Selection

Company type:
- Publishers

Employees:
- > 100

Storage:
- Please select
  - 0-1 GB
  - 1-100 GB
  - 100-1000 GB
  - > 1000 GB

Gap level:
- 0.6-1.0
- 0.3-0.6
- 0.1-0.3

Total
- Gap: 0.38
Gap Analysis Methodology: Reporting and Documentation

- Reports are generated for communication

- Data, Views and Settings are stored for documentation and reproduction of results

- Performed by tool Administrator
Agenda

I. Introduction

II. Gap Analysis Framework & Methodology

III. Visual Gap Analysis IT-Tool (Live Demo)

IV. Example: Publishers Survey analysed
Tool for Gap-Analysis
Agenda

I. Introduction

II. Gap Analysis Framework & Methodology

III. Visual Gap Analysis IT-Tool (Live Demo)

IV. Example: Publishers Survey analysed
Analysis of Gaps: Publisher Community
All Respondents (185)

- The knowledge about long term preservation in the entire sample seems to be high
- Awareness and implementation is on a moderate level
- The biggest gap lies in the commitment to actually “live” and promote preservation. Particularly the willingness to share data seems to be low among the publishers.
"Can authors submit their underlying digital research data with their publication to you?"

**"No"**

(44)

- Long Term Preservation 0.33
  - Awareness 0.31
  - Knowledge 0.49
  - Implementation 0.27
  - Commitment 0.27

- Offers Data Upload and Access 0.16
- Verifies Data 0.23
- Sees Data Storage Need 0.34
- Preserves Data 0.50
- General Knowledge 0.89
- Technology Knowledge 0.08
- Offers Upload, Validation and Collaboration 0.05
- Preserving Strategies and Operations 0.10
- Policies 0.42
- Willingness to Pay 0.33
- Willing to share data 0.05

**"Yes"**

(106)

- Long Term Preservation 0.41
  - Awareness 0.43
  - Knowledge 0.51
  - Implementation 0.40
  - Commitment 0.32

- Offers Data Upload and Access 0.40
- Verifies Data 0.35
- Sees Data Storage Need 0.42
- Preserves Data 0.55
- General Knowledge 0.89
- Technology Knowledge 0.12
- Offers Upload, Validation and Collaboration 0.05
- Preserving Strategies and Operations 0.10
- Policies 0.48
- Willingness to Pay 0.38
- Willing to share data 0.08
“Are you planning to develop a validation process for data submission?”

<table>
<thead>
<tr>
<th>“In more than 5 years” (3)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Long Term Preservation 0.26</td>
<td></td>
</tr>
<tr>
<td>- Awareness 0.20</td>
<td>5 + Offers Data Upload and Access 0.27</td>
</tr>
<tr>
<td></td>
<td>5 + Verifies Data 0.13</td>
</tr>
<tr>
<td></td>
<td>8 + Sees Data Storage Need 0.38</td>
</tr>
<tr>
<td></td>
<td>2 + Preserves Data 0.33</td>
</tr>
<tr>
<td></td>
<td>3 + General Knowledge 0.78</td>
</tr>
<tr>
<td></td>
<td>8 + Technology Knowledge 0.00</td>
</tr>
<tr>
<td>- Knowledge 0.39</td>
<td>10 + Offers Upload, Validation and Collaboration 0.33</td>
</tr>
<tr>
<td></td>
<td>9 + Preserving Strategies and Operations 0.13</td>
</tr>
<tr>
<td></td>
<td>5 + Policies 0.27</td>
</tr>
<tr>
<td></td>
<td>4 + Willingness to Pay 0.17</td>
</tr>
<tr>
<td></td>
<td>1 + Willing to share data 0.00</td>
</tr>
<tr>
<td>- Implementation 0.24</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>- Commitment 0.14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>“In less than 3 years” (6)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>- Long Term Preservation 0.43</td>
<td></td>
</tr>
<tr>
<td>- Awareness 0.44</td>
<td>5 + Offers Data Upload and Access 0.43</td>
</tr>
<tr>
<td></td>
<td>5 + Verifies Data 0.37</td>
</tr>
<tr>
<td></td>
<td>8 + Sees Data Storage Need 0.38</td>
</tr>
<tr>
<td></td>
<td>2 + Preserves Data 0.58</td>
</tr>
<tr>
<td></td>
<td>3 + General Knowledge 1.00</td>
</tr>
<tr>
<td></td>
<td>8 + Technology Knowledge 0.27</td>
</tr>
<tr>
<td>- Knowledge 0.64</td>
<td>10 + Offers Upload, Validation and Collaboration 0.43</td>
</tr>
<tr>
<td></td>
<td>9 + Preserving Strategies and Operations 0.37</td>
</tr>
<tr>
<td></td>
<td>5 + Policies 0.40</td>
</tr>
<tr>
<td></td>
<td>4 + Willingness to Pay 0.38</td>
</tr>
<tr>
<td></td>
<td>1 + Willing to share data 0.00</td>
</tr>
<tr>
<td>- Implementation 0.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>- Commitment 0.26</td>
<td></td>
</tr>
</tbody>
</table>
“Does your organisation have a policy for preservation of digital publications?”

<table>
<thead>
<tr>
<th>“No” or “Don’t know” (86)</th>
<th>“Yes” (99)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Long Term Preservation 0.32</th>
<th>Long Term Preservation 0.43</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness 0.29</td>
<td>Awareness 0.42</td>
</tr>
<tr>
<td>Knowledge 0.48</td>
<td>Knowledge 0.52</td>
</tr>
<tr>
<td>Implementation 0.30</td>
<td>Implementation 0.40</td>
</tr>
<tr>
<td>Commitment 0.19</td>
<td>Commitment 0.36</td>
</tr>
</tbody>
</table>

- Offers Data Upload and Access 0.27
- Verifies Data 0.23
- Sees Data Storage Need 0.32
- Preserves Data 0.35
- General Knowledge 0.90
- Technology Knowledge 0.06
- Offers Upload, Validation and Collaboration 0.33
- Preserving Strategies and Operations 0.26
- Policies 0.26
- Willingness to Pay 0.27
- Willing to share data 0.05

- Offers Data Upload and Access 0.32
- Verifies Data 0.33
- Sees Data Storage Need 0.41
- Preserves Data 0.63
- General Knowledge 0.89
- Technology Knowledge 0.14
- Offers Upload, Validation and Collaboration 0.41
- Preserving Strategies and Operations 0.39
- Policies 0.60
- Willingness to Pay 0.42
- Willing to share data 0.07
"We offer online collaboratories and we preserve them (for the long term)"
**Does your organisation have a disaster recovery policy for its digital content?**

<table>
<thead>
<tr>
<th>Response</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No</strong></td>
<td>51</td>
</tr>
<tr>
<td><strong>Yes</strong></td>
<td>79</td>
</tr>
</tbody>
</table>

### No (51)

- **Long Term Preservation 0.35**
  - **Awareness 0.38**
  - **Knowledge 0.51**
  - **Implementation 0.25**
  - **Commitment 0.26**

### Yes (79)

- **Long Term Preservation 0.44**
  - **Awareness 0.43**
  - **Knowledge 0.52**
  - **Implementation 0.41**
  - **Commitment 0.39**
"If your organisation has not yet systematically organised long-term preservation, when do you expect to do this?"

- "Not planned yet" (31)
  - Long Term Preservation 0.30
    - Awareness 0.22
    - Knowledge 0.43
    - Implementation 0.22
    - Commitment 0.23

- "< 1 year" (51)
  - Long Term Preservation 0.43
    - Awareness 0.42
    - Knowledge 0.45
    - Implementation 0.40
    - Commitment 0.45
“How important is the reason ‘digital data potentially has economic value’ for preservation?”

“very important” (112)

- Long Term Preservation 0.35
  - Awareness 0.31
    - Offers Data Upload and Access 0.28
    - Verifies Data 0.22
  - Knowledge 0.49
    - Sees Data Storage Need 0.32
    - Preserves Data 0.40
    - General Knowledge 0.89
    - Technology Knowledge 0.08
  - Implementation 0.34
    - Offers Upload, Validation and Collaboration 0.36
  - Commitment 0.25
    - Preserving Strategies and Operations 0.32
    - Policies 0.41
      - Willingness to Pay 0.29
      - Willing to share data 0.04

“Yes” (26)

- Long Term Preservation 0.44
  - Awareness 0.46
    - Offers Data Upload and Access 0.34
    - Verifies Data 0.37
  - Knowledge 0.52
    - Sees Data Storage Need 0.50
    - Preserves Data 0.65
    - General Knowledge 0.91
    - Technology Knowledge 0.13
  - Implementation 0.39
  - Commitment 0.39
    - Preserving Strategies and Operations 0.37
    - Policies 0.54
      - Willingness to Pay 0.59
      - Willing to share data 0.04
Summary:
Gaps within the publisher community

- The Gap Analysis Framework and IT-Tool provides insight into community-specific gaps and leverages the PARSE.Insight survey data from WP3.
- Domain-specific knowledge is necessary to use the tool and contextualize results.
- The tool allows for different perspectives on communities ("Drill-Down").
- The degree of adopting and implementing preservation strategies differs strongly within the community of publishers.
- Gaps could be identified in three areas:
  - Commitment (big gaps)
  - Implementation (medium – big gaps)
  - Awareness (medium gaps)
- The knowledge dimension is not sufficiently covered by the publisher survey data.
- First means to close these gaps will be identified by follow-up expert workshops (Tool results to stimulate discussion).
Thank you for your attention