Data Quality Management in Data Exchange Platforms – An Approach for the Industrial Data Space in Germany

CHRISTOPH QUIX

QDB WORKSHOP, SEPTEMBER 5, 2016
Products are becoming hybrid: Integration of classical & digital services
Role of Data is Changing

- Data as result of a process
- Data as process enabler
- Data as enabler for products
- Data as a product

Value

Time
Data Quality Definition

Product-oriented
- Based on features of the product

Application-oriented
- Fulfills requirements of users

Process-oriented
- Compliance of production process with specifications

Value-oriented
- Price-performance ratio
How to measure the value?

- What is the value of a product?
  - Costs of its production
  - Market value (negotiation, auction, ...)
  - Usefulness / value of benefit for a specific business process
    - Which additional costs would occur if you do not have that product?

- Can this also be applied to data?
  - Data is a digital product, it can be copied easily
## Data Items are Goods with a Value

<table>
<thead>
<tr>
<th>Company</th>
<th>Service</th>
<th>Ctry</th>
<th>Data Type</th>
<th>Valuation</th>
<th>Value per Data item</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kroger</td>
<td>Super Markets</td>
<td>US</td>
<td>Customer profiles incl. Buying profile</td>
<td>Market value</td>
<td>1.6 EUR(^1)</td>
</tr>
<tr>
<td>facebook</td>
<td>Social Network</td>
<td>US</td>
<td>User profiles</td>
<td>Market value</td>
<td>225 USD(^2,3)</td>
</tr>
<tr>
<td>Festo</td>
<td>Automation technology</td>
<td>DE</td>
<td>Parts master data</td>
<td>Production costs</td>
<td>500 to 5,000 EUR(^4)</td>
</tr>
<tr>
<td>Syngenta</td>
<td>Agro chemicals</td>
<td>CH</td>
<td>Parts master data</td>
<td>Value of benefit</td>
<td>184 CHF(^5)</td>
</tr>
</tbody>
</table>

1) [http://www.wsj.de/nachrichten/SB11446175161338053998704580212211843086060](http://www.wsj.de/nachrichten/SB11446175161338053998704580212211843086060)
2) [http://en.wikipedia.org/wiki/Facebook; 890 million daily active users.](http://en.wikipedia.org/wiki/Facebook; 890 million daily active users.)
3) [http://www.ft.com/cms/s/0/ecc0f050-37a3-11e4-bd0a-00144feabdc0.html#axzz3RH6POTH; Marktkapitalisierung von 200 Mrd. USD.](http://www.ft.com/cms/s/0/ecc0f050-37a3-11e4-bd0a-00144feabdc0.html#axzz3RH6POTH; Marktkapitalisierung von 200 Mrd. USD.)
Internal IDS Connector

Industrial Data Space

External IDS Connector

Upload / Download / Search

Internet

Third Party Cloud Provider

Industrial Data Space Broker

Index

Registry

IDS Connector

Industrial Data Space App Store

Vocabularies

Apps

Company A

Company B

Upload / Download

Download
Key Features of the IDS

- Decentralized P2P Architecture
- Sovereignty
- Data as Goods
- Security
- Data Exchange
- Governance
- Collaborative Rules
- Trust
- Certified Participants
- Open neutral and user driven
- Scalability
- Network Platform and Services
- Network
Technical Architecture of an IDS Connector

- Connectors provide a controlled environment for the basic functions and data exchange
- Control is guaranteed by a virtualization concept with specific security functions (application container technology, e.g., Docker)
- There are different levels of security

Diagram:
- Application-specific modules, system adapters, ...
- IDS Data Core
- Core IDS Container
- Application Container Management
- Core OS
Data Quality in the IDS

- Data is traded like a product
- Does it need quality control like for regular products?
  - Who will perform the quality checks?
    - Provider
    - Consumer
    - Some trusted third party

- Data quality control needs to be integrated into the data exchange process

- Quality of the process
  - Certification of software
    - Can only guarantee basic functionality, compliance with rules, but not the quality of the data management processes
  - Certification of participants
Conclusion

- Digital services are being integrated into classical products
  - Features of the classical product become less important

- Data is a product and has a value

- Industrial Data Space provides an infrastructure for data exchange in a secure environment with certified software and certified participants

- Challenges
  - Valuation of data
  - Data quality control
  - Value-oriented and quality-oriented data integration