From data-driven startup to large company in a decade

Dr. Andreas Both
Head of Research and Development
Unister GmbH
Germany

European Data Forum 2012
Copenhagen, June 6-7, 2012
... challenges of Big Data and the emerging Data Economy and to develop suitable action plans for addressing these challenges...
Claim

... challenges of Big Data and the emerging Data Economy and to develop suitable action plans for addressing these challenges ...
Claim

... challenges of Big Data and the emerging Data Economy and to develop suitable action plans for addressing these challenges ...
... challenges of Big Data and the emerging Data Economy and to develop suitable action plans for addressing these challenges ...
Personal Retrospective . . .

E-commerce is challenging. The complexity of the business model is very high. Everything has to work smoothly, otherwise the punishment comes quickly.

Some lessons:
- Improve your business model. Care about our customers.
- Rapid impact of decisions.
- Increasing data complexity.
Personal Retrospective . . .

. . . joining Unister, April 2010

- eCommerce is challenging
- complexity of business
- everything has to work smoothly
- *punishment* comes quickly
Personal Retrospective . . .

...joining Unister, April 2010

- eCommerce is challenging
- complexity of business
- everything has to work smoothly
- *punishment* comes quickly

some lessons

- improve your business model
- care about our customers
Personal Retrospective . . .

joining Unister, April 2010
  - eCommerce is challenging
  - complexity of business
  - everything has to work smoothly
  - *punishment* comes quickly

some lessons
  - improve your business model
  - care about our customers
  - rapide impact of decisions
  - increasing data complexity
It’s about the business activities, stupid.

- Should we really talk about data?
- Why should we talk about data processes?
- Why should we talk about the data analytics?
It’s about the business activities, stupid.

• Should we really talk about data?
• Why should we talk about data processes?
• Why should we talk about the data analytics?

➢ Is it just about the business?
➢ Is data a core value?
IT companies vs. traditional business

**classic business**
need natural resource
to establish processes
depending on location

**e-business**
need data
to establish processes
independent from location

Dr. Andreas Both, Head of R&D, Unister
IT companies vs. traditional business

**classic business**
- need natural resource to establish processes depending on location

**e-business**
- need data to establish processes independent from location

SME will find many niches within the market.
IT companies vs. traditional business

**classic business**
- need natural resource to establish processes depending on location

**e-business**
- need data to establish processes independent from location

SME will find many niches within the market.

...establish large companies too!
Unister’s Success Story

- Internet startup mainly located in Leipzig, Germany
- private limited company (German GmbH)
- founded in 2002, 5 founders
- managing director: Thomas Wagner
Unister’s Success Story

- Internet startup mainly located in Leipzig, Germany
- private limited company (German GmbH)
- founded in 2002, 5 founders
- managing director: Thomas Wagner
- eCommerce company, B2C
  - national and international activities
  - > 40 web-portals and services
  - > 13.22 mio. unique user / month in Germany

\(^a\)AGOF e.V. / internet facts 2012-01
Unister’s Success Story

- Internet startup mainly located in Leipzig, Germany
- private limited company (German GmbH)
- founded in 2002, 5 founders
- managing director: Thomas Wagner
- eCommerce company, B2C
- national and international activities
  - > 40 web-portals and services
  - > 13.22 mio. unique user / month in Germany a

IT-driven approach

aAGOF e.V. / internet facts 2012-01
Unister’s Success Story
Unister’s Success Story

Number of Employees

- 2003: 1 employee
- 2004: 7 employees
- 2005: 38 employees
- 2006: 106 employees
- 2007: 185 employees
- 2008: 372 employees
- 2009: 701 employees
- 2010: 1157 employees
- 2011: 1530 employees

Dr. Andreas Both, Head of R&D, Unister
IT companies

Data is the foundation

- getting data is tough
- having data is not enough
- managing data is challenging
IT companies

Data is the foundation

- getting data is tough
- having data is not enough
- managing data is challenging

Next parts of the talk

- Data Access
- Data Integration
- (Big) Data Analyses
IT companies

Data is the foundation

- getting data is tough
- having data is not enough
- managing data is challenging

Next parts of the talk

- Data Access
- Data Integration
- (Big) Data Analyses

← Steps a start-up has to tackle!
IT companies

Data is the foundation

- getting data is tough
- having data is not enough
- managing data is challenging

Next parts of the talk

- Data Access
- Data Integration
- (Big) Data Analyses

← Steps a start-up has to tackle!
Steps Unister had tackled.
IT companies

Data is the foundation

- getting data is tough
- having data is not enough
- managing data is challenging

Next parts of the talk

- Data Access
- Data Integration
- (Big) Data Analyses

← Steps a start-up has to tackle!
Steps Unister had tackled.
Steps Unister has to tackle.
Data Access
Data Access

Observations

- business models need access of data
  - support, description, enrichment, …
Data Access

Observations
- business models need access of data
  ▶ support, description, enrichment, ...

Unister’s Success (step 1)
- was capable of integrating many data sets
- user-focussed data
Data Access

Observations

- business models need access of data
  - support, description, enrichment, ...

Unister’s Success (step 1)

- was capable of integrating many data sets
- user-focussed data

eCommerce demand

- local information
- link to local events
- (legal) constraints
  → such data not available
Data Access

Observations

• business models need access of data
  ▶ support, description, enrichment, . . .

Challenges

• *Open Data* should be established
• Standards have to be defined and followed

Unister’s Success (step 1)

• was capable of integrating many data sets
• user-focused data

eCommerce demand

• local information
• link to local events
• (legal) constraints
  → such data not available
Data Access: Needed Actions

- *Open Data* initiatives need support
- enhanced tool support
- political commitment
Data Access: Needed Actions

- *Open Data* initiatives need support
- enhanced tool support
- political commitment

- will establish (local) data economics
- locally connected companies could grow
Data Integration
Data Integration

Observations

• bread and butter of data-driven companies
• needs much effort
Data Integration

Observations

• bread and butter of data-driven companies
• needs much effort

Unister’s Success (step 2)

• fusion of different data sets leads to good user experience
Data Integration

Observations
- bread and butter of data-driven companies
- needs much effort

Unister’s Success (step 2)
- fusion of different data sets leads to good user experience

eCommerce demand
- matching processes
- integration tools
Data Integration

Observations

- bread and butter of data-driven companies
- needs much effort

Challenges

- establish distributed knowledge base
  - Linked Data paradigm
- NOSQL + SQL

Unister’s Success (step 2)

- fusion of different data sets leads to good user experience

eCommerce demand

- matching processes
- integration tools
Data Integration: Needed Actions

- support Linked Data Cloud
- companies need sound and solid data sets
Data Integration: Needed Actions

- support Linked Data Cloud
- companies need sound and solid data sets
- research on scalable data integration processes
- Cloud Computing → Big Data challenge
Data Analyses and Big Data
Data Analyses and Big Data

Observations

- disseminate, understand and ultimately benefit from increasing volumes of data
- example: social networks
Data Analyses and Big Data

Observations

- disseminate, understand and ultimately benefit from increasing volumes of data
- example: social networks

Unister’s Success (step 3)

- defining data analysis processes with impact
- pareto-optimal processes lead to good coverage
- analysis came to a limit because of many segments
Data Analyses and Big Data

Observations

- disseminate, understand and ultimately benefit from increasing volumes of data
- example: social networks

Unister’s Success (step 3)

- defining data analysis processes with impact
- pareto-optimal processes lead to good coverage
- analysis came to a limit because of many segments
demand

- descriptive analyses processes
- higher-level process interfaces
- good developers
Data Analyses and Big Data

Observations
- disseminate, understand and ultimately benefit from increasing volumes of data
- example: social networks

Challenges
- ...

Unister’s Success (step 3)
- defining data analysis processes with impact
- pareto-optimal processes lead to good coverage
- analysis came to a limit because of many segments

eCommerce demand
- descriptive analyses processes
- higher-level process interfaces
- good developers
Data Analyses and Big Data: Global Movement

DATA GROWTH

COST OF STORAGE/GB (DECREASE)

2011 2012 2013 2014 2015

source: ucsd.edu
Data Analyses and Big Data: Challenges

Source: hadapt.com

Dr. Andreas Both, Head of R&D, Unister
Data Analyses and Big Data: Challenges

The 3 V

- Volume
- Variety
- Velocity
Data Analyses and Big Data: Challenges

The 3 V
- Volume
- Variety
- Velocity

The +2 V
- Virality
- Viscosity

© 2011 R Wang & Insider Associates, LLC. All rights reserved.
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

- need systems for big data analyses

hadoop
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

- need systems for big data analyses
  - good support: Cloud Computing, . . .
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

- need systems for big data analyses
  - good support: Cloud Computing, . . .
- need people to operate on the systems
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

• need systems for big data analyses
  ▶ good support: Cloud Computing, . . .
• need people to operate on the systems
  → ok: some experience available
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

• need systems for big data analyses
  ▶  *good support*: Cloud Computing, . . .

• need people to operate on the systems
  →  *ok*: some experience available

• need people to develop applications
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge ...

levels of challenge

- need systems for big data analyses
  - *good support*: Cloud Computing, ...
- need people to operate on the systems
  - *ok*: some experience available
- need people to develop applications
  - *bad*: rarely taught
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

- need systems for big data analyses
  - *good support*: Cloud Computing, . . .
- need people to operate on the systems
  - *ok*: some experience available
- need people to develop applications
  - *bad*: rarely taught
- need people to think about using the network effect
Data Analyses and Big Data: Needed Actions

handle this is all about knowledge . . .

levels of challenge

- need systems for big data analyses
  - good support: Cloud Computing, . . .
- need people to operate on the systems
  - ok: some experience available
- need people to develop applications
  - bad: rarely taught
- need people to think about using the network effect
  - very bad: Talent Gap
Data Analyses and Big Data: Potential

Big data – The next frontier for innovation, competition, and productivity
(McKinsey May 2011)
Data Analyses and Big Data: Potential

*Big data – The next frontier for innovation, competition, and productivity*  
(McKinsey May 2011)

- Plenty of possibilities!
Summary: Most important activities

- Open Data will give a push
- well-developed tools are crucial for SME
- talent gap has to be tackled
Conclusion

Data Access, Data Integration, Data Analyses
Conclusion

Data Access, Data Integration, Data Analyses

*Big Data*

Cost of Storage/GB (Decrease)

2011 2012 2013 2014 2015

Dr. Andreas Both, Head of R&D, Unister
Conclusion

Data Access, Data Integration, Data Analyses

Big Data

successful companies

COST OF STORAGE/GB (DECREASE)

2011 2012 2013 2014 2015

Number of Employees

2003 2004 2005 2006 2007 2008 2009 2010 2011

Slide 23 Dr. Andreas Both, Head of R&D, Unister
From data-driven startup to large company in a decade

Dr. Andreas Both
Head of R&D
Unister GmbH
andreas.both@unister.de
+49 341 65050 24496
http://www.unister.de