

Data Science – (k)eine Teenagerliebe?

Thilo Stadelmann, Swiss ICT Symposium, 05. November 2011



Teenagerliebe?



Dan Ariely

6. Januar 2013 · 



Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...

 Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

 890 mal geteilt

 Vorherige Kommentare anzeigen

49 von 89

Quelle: <https://www.facebook.com/dan.ariely/posts/904383595868>

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 · 🌐

Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...

👍 Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

↪ 890 mal geteilt

💬 Vorherige Kommentare anzeigen

49 von 89

Quelle: <https://www.facebook.com/dan.ariely/posts/904383595868>

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 · 🌐

Big data is like teenage sex: **everyone talks** about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...

👍 Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

↪ 890 mal geteilt

💬 Vorherige Kommentare anzeigen

49 von 89

Quelle: <https://www.facebook.com/dan.ariely/posts/904383595868>

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 · 🌐



Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it...

👍 Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

↪ 890 mal geteilt

💬 Vorherige Kommentare anzeigen

49 von 89

Quelle: <https://www.facebook.com/dan.ariely/posts/904383595868>

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 · 🌐

Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it... → **Keiner macht es!?**

👍 Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

↪ 890 mal geteilt

💬 Vorherige Kommentare anzeigen

49 von 89

Quelle: <https://www.facebook.com/dan.ariely/posts/904383595868>

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 ·

Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it... → **Keiner macht es!?**

Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

Agenda

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 ·

Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it... → **Keiner macht es!?**

Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

Agenda

- Data Science in der öffentlichen Wahrnehmung

Teenagerliebe?

Data Science \neq Big Data !
Aber: Hohe Korrelation in
der aktuellen öffentlichen
Wahrnehmung.



Dan Ariely

6. Januar 2013 ·

Big data is like teenage sex: everyone talks about it, nobody really knows how to do it, everyone thinks everyone else is doing it, so everyone claims they are doing it... → **Keiner macht es!?**

Walter Erquinigo, Jorge Vergara Roa, Guilherme Felitti und 1.901 anderen gefällt das.

Agenda

- Data Science in der öffentlichen Wahrnehmung
 - Spannende Anwendungsfälle, die vielleicht überraschen
- Enorm viele Branchen können profitieren

Beliebte Data Science Beispiele

Beliebte Data Science Beispiele

- Kauf-Empfehlungen: Bspw. «Netflix Prize» (2009)
Quelle: <http://www.netflixprize.com/>





Beliebte Data Science Beispiele

- Kauf-Empfehlungen: Bspw. «Netflix Prize» (2009)
Quelle: <http://www.netflixprize.com/>
- Abwanderungs-Prognosen: Telkos, Versicherungen, ...
Quelle: <http://engineering.zhaw.ch/de/engineering/forschung/publikationen-t.html?pi=206434&gu=96884>





Beliebte Data Science Beispiele

- Kauf-Empfehlungen: Bspw. «Netflix Prize» (2009)
Quelle: <http://www.netflixprize.com/>
- Abwanderungs-Prognosen: Telkos, Versicherungen, ...
Quelle: <http://engineering.zhaw.ch/de/engineering/forschung/publikationen-t.html?pi=206434&gu=96884>
- Warenkorb-Analyse: Bspw. Windeln & Bier, Bücher, ...
Quelle: <http://web.onetel.net.uk/~hibou/Beer%20and%20Nappies.html>





Beliebte Data Science Beispiele

- Kauf-Empfehlungen: Bspw. «Netflix Price» (2009)
Quelle: <http://www.netflixprize.com/>
- Abwanderungs-Prognosen: Telkos, Versicherungen, ...
Quelle: <http://engineering.zhaw.ch/de/engineering/forschung/publikationen-t.html?pi=206434&gu=96884>
- Warenkorb-Analyse: Bspw. Windeln & Bier, Bücher, ...
Quelle: <http://web.onetel.net.uk/~hibou/Beer%20and%20Nappies.html>
- Risiko und Missbrauch: Lebensversicherungen, Kreditkarten-nutzung
Quelle: Eric Siegel, «Predictive Analytics», John Wiley & Sons, 2013



...



Beliebte Data Science Beispiele

- Kauf-Empfehlungen: Bspw. «Netflix Price» (2009)
Quelle: <http://www.netflixprize.com/>
- Abwanderungs-Prognosen: Telkos, Versicherungen, ...
Quelle: <http://engineering.zhaw.ch/de/engineering/forschung/publikationen-t.html?pi=206434&gu=96884>
- Warenkorb-Analyse: Bspw. Windeln & Bier, Bücher, ...
Quelle: <http://web.onetel.net.uk/~hibou/Beer%20and%20Nappies.html>
- Risiko und Missbrauch: Lebensversicherungen, Kreditkarten-nutzung
Quelle: Eric Siegel, «Predictive Analytics», John Wiley & Sons, 2013



...

→ Benutzerprofile zur Optimierung des Konsums

Datengetriebenes arbeiten und personenbezogene Daten

Datengetrieben arbeiten heisst:

- Den gesunden Menschenverstand beiseite lassen
 - Korrelation suchen, nicht direkt Kausalzusammenhänge
- ➔ Unbewusste Vorgänge werden normativ
Vergessen ist nicht vorgesehen
- ➔ Nur das gewöhnliche wird gelernt

Zum Weiterlesen: Byung-Chul Han, «Psychopolitik», S. Fischer, 2014



Datengetriebenes arbeiten und personenbezogene Daten

Datengetrieben arbeiten heisst:

- Den gesunden Menschenverstand beiseite lassen
 - Korrelation suchen, nicht direkt Kausalzusammenhänge
- ➔ Unbewusste Vorgänge werden normativ
Vergessen ist nicht vorgesehen
- ➔ Nur das gewöhnliche wird gelernt

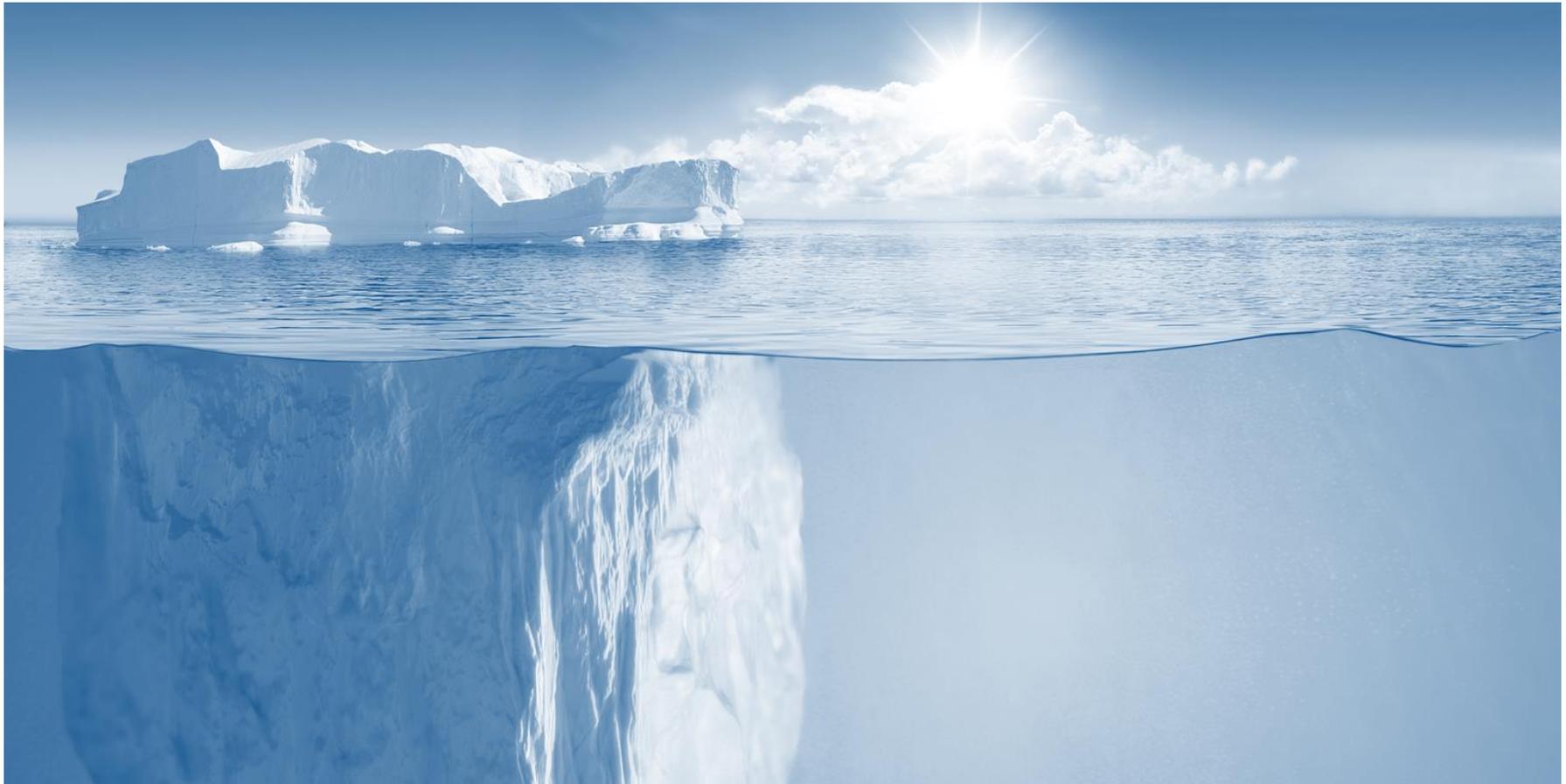
Zum Weiterlesen: Byung-Chul Han, «Psychopolitik», S. Fischer, 2014

Whatever you are doing with computers, you are changing our world. Is it for the better?

Quelle: Michael L. Brodie, http://www.michaelbrodie.com/michael_brodie_statement.asp



Data Science abseits der Hauptbühne



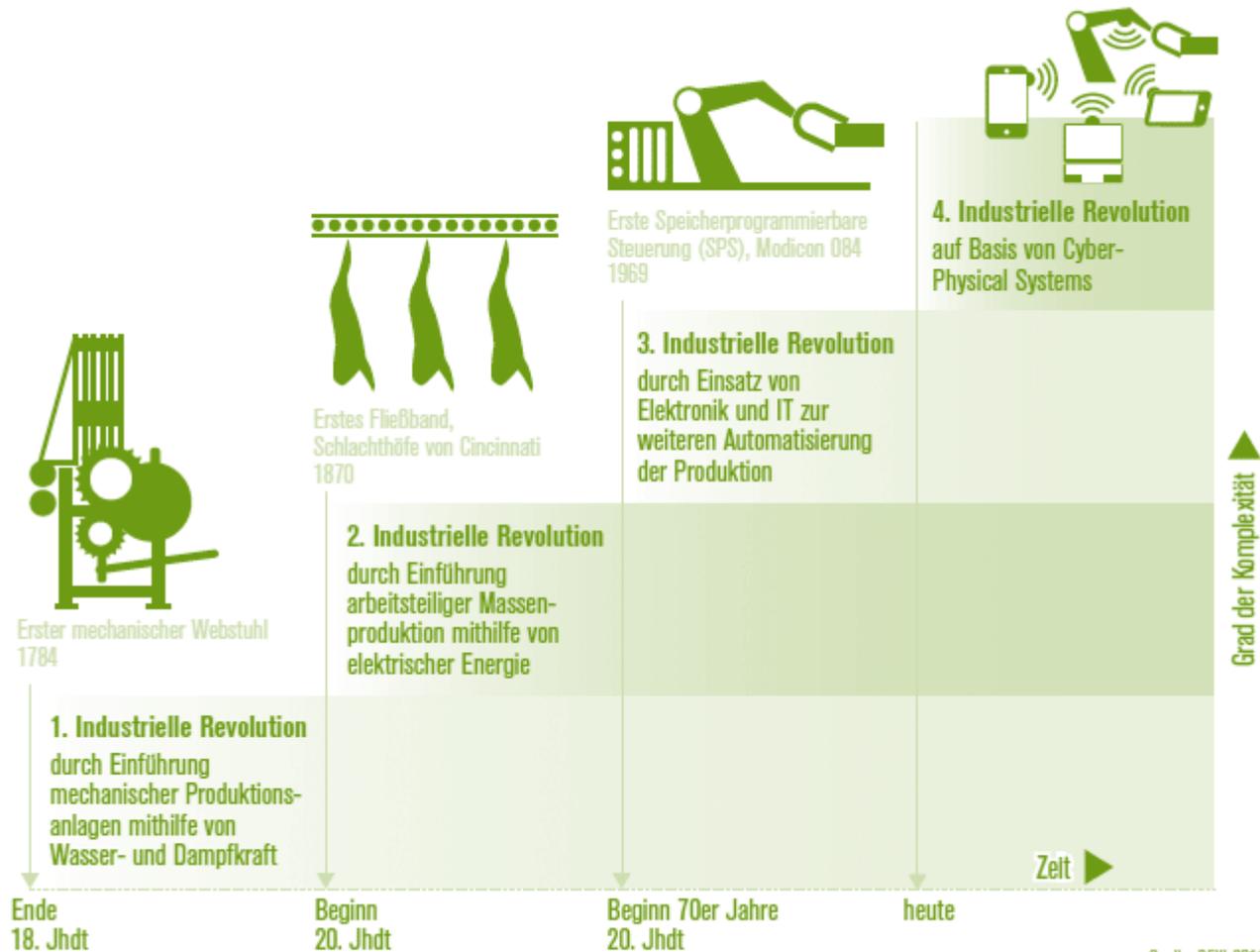
Beispiel automatische Vorhersage

Was?	Wer?
Grippe	Google Flu Trends: Sagt anhand Suchtrends Zunahme von Grippefällen 7-10 Tage vor der zuständigen staatlichen Seuchenstelle voraus
Systemfehler	Argonne National Laboratory: Vorhersage von Rissen im Kühlsystem von AKWs BNSF Railway: Vorhersage defekter Gleise 85% → Verhindert schwere Zugunglücke TTX: Vorhersage von Fehlern in hunderttausenden von Eisenbahnrädern zur Abschätzung des jährlichen Wartungsaufwands (98.5% Genauigkeit)
Ölfördermenge	National Iranian South Oil Company: Vorhersage der Tagesproduktion an Öl
Flugverspätung	Continental Airlines: Vorhersage der Verspätung von Flügen via Radardaten (Einsparung mehrerer 10 Millionen Dollar)
Fahrer- ablenkung	Ford Motor Company, TWT GmbH: Grad der Fahrerablenkung für Autolenker (86% Genauigkeit für Ford-System)

Quellen:

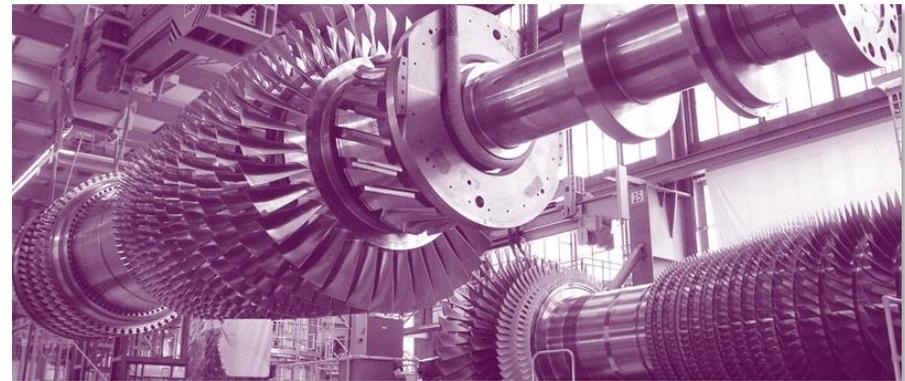
- Eric Siegel, «Predictive Analytics», John Wiley & Sons, 2013
- Thilo Stadelmann et al., «FABELHAFT - Fahrerblenkung: Entwicklung eines Meta-Fahrerassistenzsystems durch Echtzeit-Audioklassifikation», VDI Wissensforum, 2012

Beispiel Industrie 4.0



Beispiel vorausschauende Wartung

- Ausgangslage: Wartung grosser Maschinen ist teuer, Ausfall noch teurer
Ziel: Wartung, wenn Defekt kurz bevorsteht anstatt präventiv
Herausforderung: Genereller Ansatz - Tuning auf jede individuelle Maschine vermeiden



Lösung:

- Sensoren überwachen Betriebsparameter kontinuierlich
- Datenanalyse z.B. "in der Cloud"
- Selbstadaptive Signalverarbeitungsansätze
- Verfolgen von Änderungen über die Zeit

Beispiel Informationsbeschaffung / Suche

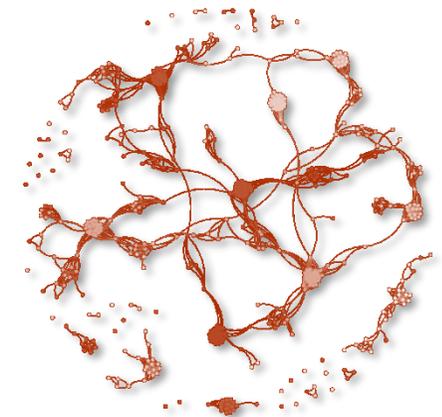
- Ausgangslage: Ca. 7'500 Förderstiftungen in der Schweiz
- Ziel: Visuelle Suche anhand eines Förderantrags soll
geeignetsten Sponsor zu Tage fördern
- Herausforderung: Quantifizieren & Darstellen der inhaltlichen
Ähnlichkeit von Stiftungszwecken

stiftungschweiz.ch



Lösung:

- Entwicklung eines mehrsprachigen Suchsystems
- Extrem kleiner Datensatz für aussagekräftige Suchverfahren
(7'500 Stiftungszweck-Dokumente)
 - ➔ Extrem recall-orientiert (Suchleistung vs. Präzision der Ergebnisse)
- Sehr grosser Datensatz für intuitive Visualisierungen
 - ➔ “Forced Directed Layout” auf Ähnlichkeitsmatrix der Wörter sowie Themenmodellierung



Beispiel Open Data Exploration («SODES»)

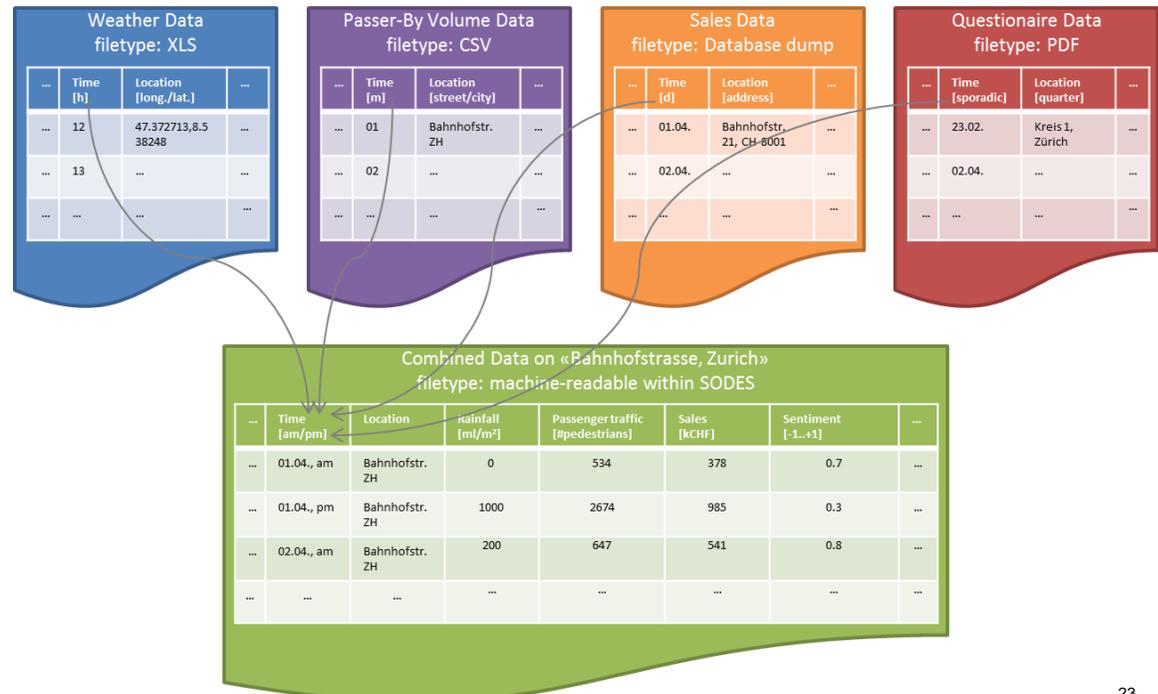
Ausgangslage: Open Data könnte sich als Goldgrube erweisen – wenn es nicht so zeitaufwändig zu integrieren wäre

Ziel: Eine Plattform für einfachen und intuitiven Zugang, Integration und Exploration verschiedener Datenquellen

Herausforderung: Betrieb der Plattform

Lösung:

- Konzept und Methodik vorhanden
- Betriebspartner gesucht



Fazit

- Fertigungsautomation
- Vorausschauende Wartung
- Suche und Informationsbeschaffung
- Intelligente Vorverarbeitung von Daten
- ...

➔ Viele lohnende Anwendungsfälle für Advanced Analytics jenseits von «BI»



Mehr Data Science?

- Weiterbildung: dlab.zhaw.ch/teaching
- Konferenz: dlab.zhaw.ch/sds2015
- Projekte: datalab@zhaw.ch