



Der Marktplatz MIA ermöglicht Business Intelligence auf dem Datenbestand des .DE Webs, beispielsweise für die Segmente Telco, Retail, Gaming, Financial Services, Internet oder Risk Analytics.

- 11/2011 - 10/2014
- Ca. 6 Mio. Euro
- BMWi funding: 58%
- ~20 FTE für 3 Jahre
- Kontakt: alexander.loeser@tu-berlin.de



Bundesministerium
für Wirtschaft
und Technologie



www.mia-marktplatz.de

Messen/Vorhersagen des Wertes einer Marke, eines Produktes, ...

- Welche Top-10 Web-Sites beeinflussen das Kaufverhalten der Kunden für 1100 ccm Motorräder?
- Sollte Steffi Graf immer noch Nudeln bewerben?

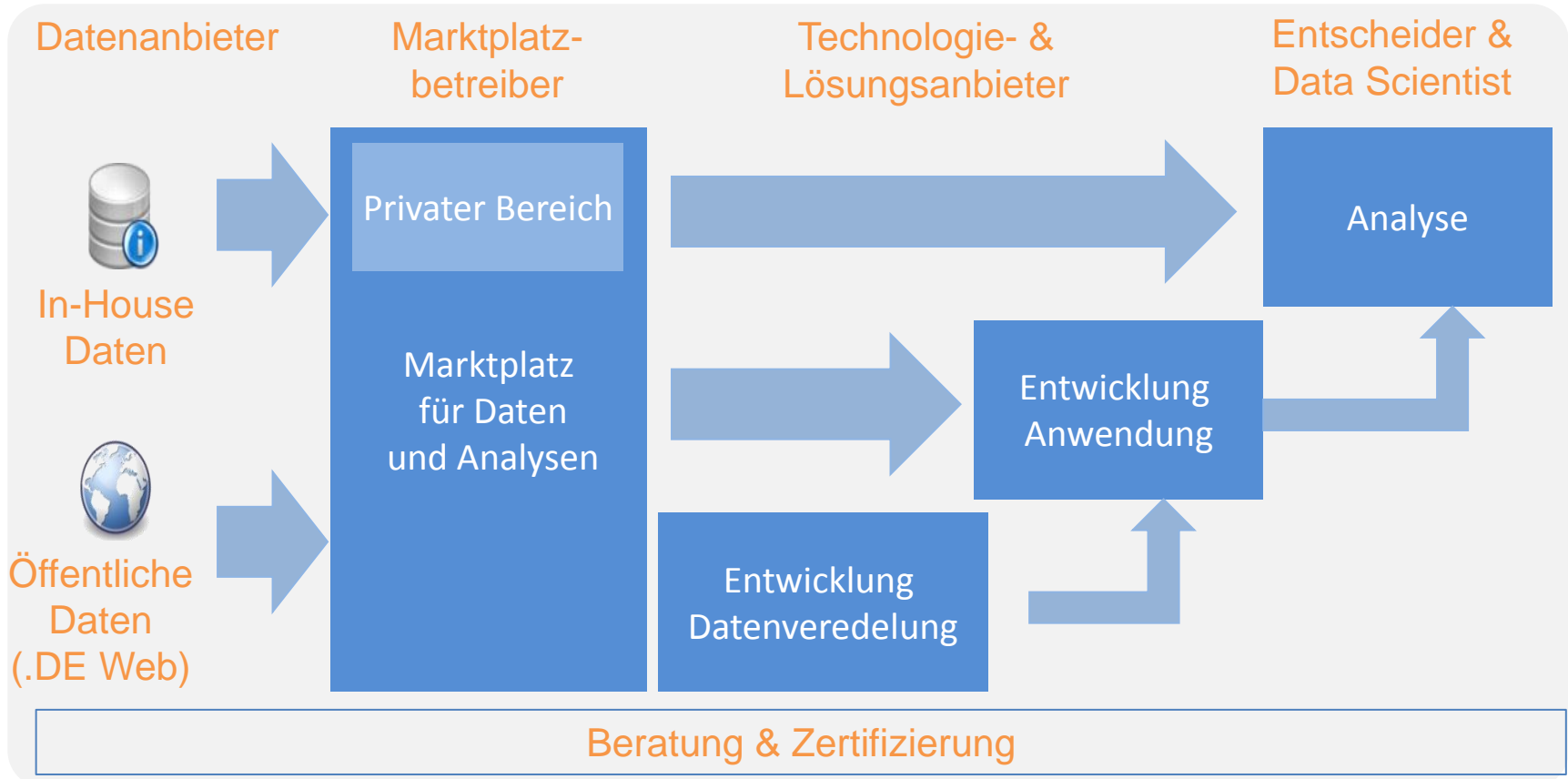
Identifikation von aufschlussreichen Korrelationen

- Welche nicht-alkoholischen Getränke trinken Fans bei einem Bundesligaspiel im Winter?
- Kaufen Besitzer eines teuren Smart-Phones auch häufiger ‚Business-Class‘ Flüge?

Wert = Verwertbarer Informationsvorsprung + Exklusivität der Information innerhalb meiner Peer-Gruppe (+ Umfangreiche Rechte für Weiterverkauf)

Quelle: Umfrage mit Experten und CEOs, u.a. von SAP, IBM, Factual.com, GfK Retail, Thomson Reuters. Löser, et.al.: Pricing Models for Data Markets. To appear at BIRTE/VLDB 2012 Springer LNBIP.

Sieben Anbietergruppen für Daten und datennahe Dienstleistungen ermöglichen Entscheidern den ausschlaggebenden Informationsvorsprung.



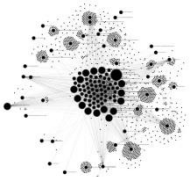
Quelle: Umfrage mit Experten und CEOs, u.a. von SAP, IBM, Factual.com, GfK Retail, Thomson Reuters. Löser, et.al.: Pricing Models for Data Markets. To appear at BIRTE/VLDB 2012 Springer LNBP.



> 500.000 Forum Postings / Tag



+ 85.000 Nachrichten / Tag von > 1.500 Quellen

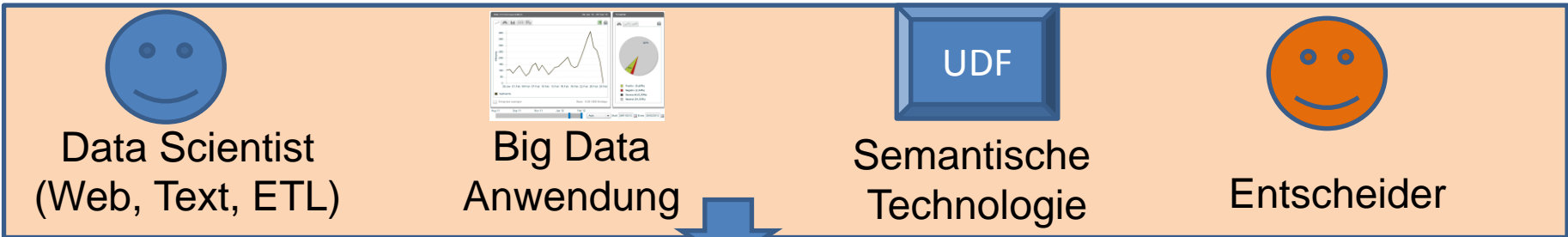


+ Kontinuierliches Crawling ausgesuchter Webseiten

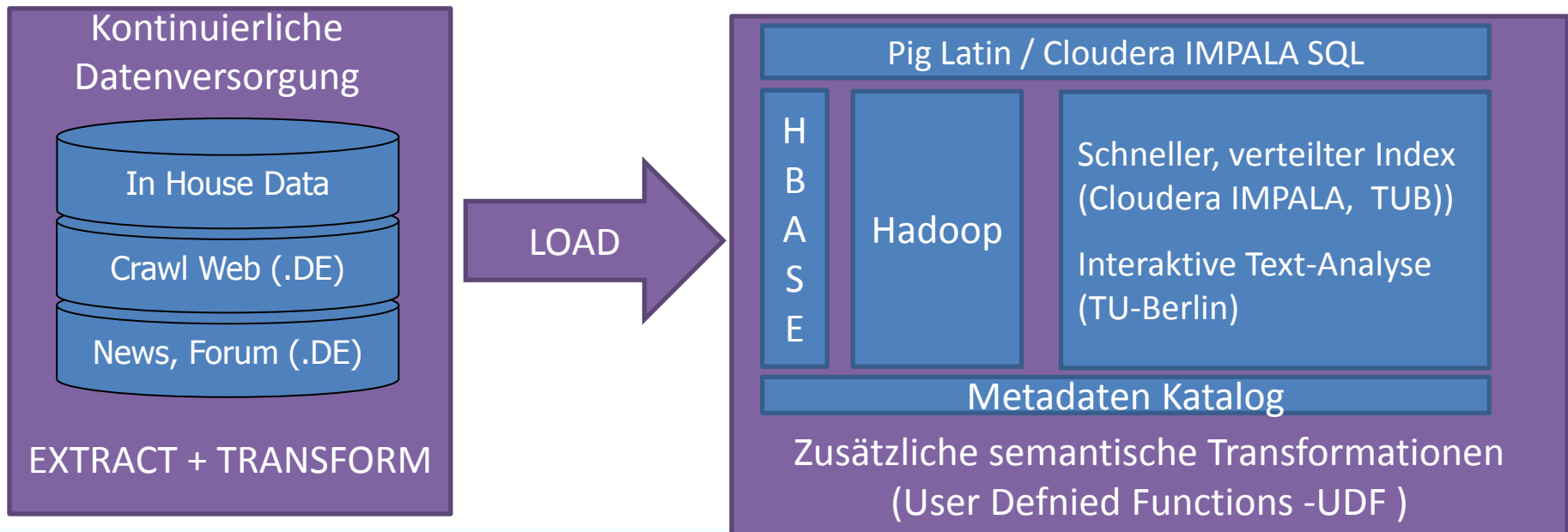
Aktueller Stand (auf MIA Cluster)

- > 80 Mill. Nachrichten (.DE)
- > 700 Mill. Crawl (.DE)
- > 300 Mill Blogs und Forenbeiträge (.DE)
- Für Universitäten: Clue Web09 Korpus (500 Mill. .COM, 50 Mill. DE)

Grundidee: Mehrere Kundengruppen teilen sich eine gemeinsame Infrastruktur und reichern einen gemeinsamen Datenbestand domänenspezifisch an.



MIA Anfragesprache (Neofonie)



- **Deutschlandradio 5/2012** „Im Meer der Webseiten: Der Online-Marktplatz MIA bietet Business-Intelligence für das deutschsprachige Web an“
<http://www.dradio.de/dlf/sendungen/computer/1755776/>
- **Deutschlandradio 03/2013** „Das Geschäft mit den Datenbergen: Big Data boomt in der IT-Branche“
<http://www.dradio.de/dlf/sendungen/computer/2035278/>
- **Heise iX 3/2013 Cloud Computing – Report.** Vertraute Wolke: BMWi gibt 50 Millionen Euro für Trusted Cloud aus.
<http://www.heise.de/ix/inhalt/2013/03/106/>
- **TU-Intern 4/2011** http://www.pressestelle.tu-berlin.de/fileadmin/a70100710/Fotos/TU_intern/2011/April/tui4_2011.pdf
- Digital Marktplatz für Business Intelligence. <http://www.digitalzeitschrift.de/mediathek.php?type=digitalk&id=329>
- **Splitter 03/2012** http://www.itdz-berlin.de/dokumente/splitter/splitter_2012_3.pdf

- Unsupervised Discovery of Relations and Discriminative Extraction Patterns Alan Akbik, Larysa Visengeriyeva, Priska Herger, Holmer Hemsén, und Alexander Löser 24th. International Conference on Computational Linguistics (COLING) 2012
- Pricing Approaches for Data Markets. Alexander Muschalle, Florian Stahl, Alexander Löser, und Gottfried Vossen. BIRTE 2012 workshop with 38th. International Conference on Very Large Data Bases (VLDB) 2012
- Beyond Search: Retrieving Complete Tuples from a Text-Database. Alexander Löser, Christoph Nagel, Stephan Pieper, und Christoph Boden Springer journal on Information System Frontiers (ISF) 2012 ('Journal Impact Factor' 1.4)
- KRAKEN, N-ary Facts in Open Information Extraction. Alan Akbik and Alexander Löser NAACL-HLT 2012 Joint Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction (AKBC-WEKEX).
- Fact-Aware Document Retrieval for Information Extraction. Christoph Boden, Alexander Löser, Christoph Nagel and Stephan Pieper Datenbank-Spektrum Volume 12, Number 2 (2012), 89-100, DOI: 10.1007/s13222-012-0088-4