DER SPIEGEL

DER SPIEGEL

Digas and Oracle Text

DER SPIEGEL

Digas and SPIEGEL Digital – an overview

- Digas?
- Document Base, Hardware Architecture, Queries
- Software: Service Oriented Architecture
- New Clients: SPIEGEL Digital, SPIEGEL Mobil
- Digas at Axel Springer Verlag
- Issues for Improvement
- Plans for 2007

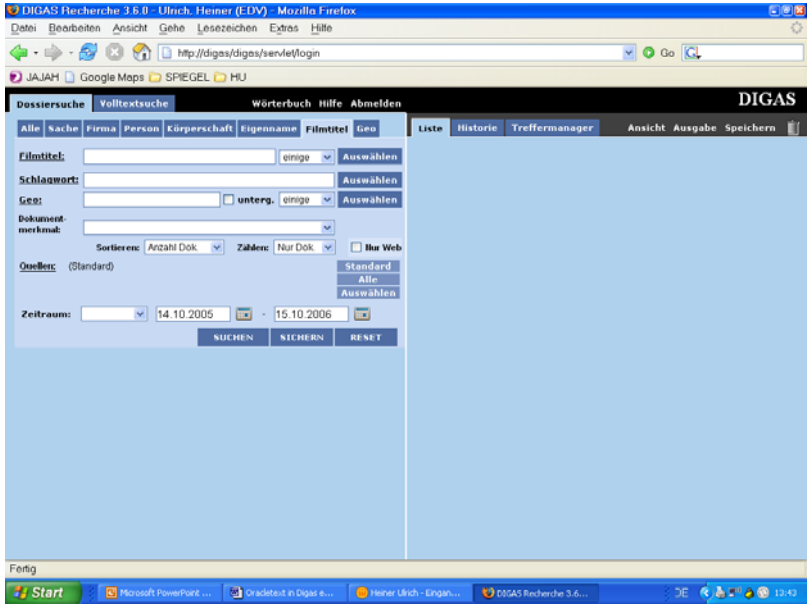
DER SPIEGEL

Digas?

- Digas is the Digital Archiving System developed at DER SPIEGEL, starting in the early 1990ies. Since 1996 Digas runs on a Oracle database, since 2001 we use Oracle Text as the full text index.
- Digas is not only a large full text database. What makes Digas different are the dossiers. A dossier is a number of articles regarding a certain topic, person or company.
- Since Jan 2005 Axel Springer Verlag started to use the Digas Archiving System. Axel Springer Verlag is our first partner, our goal is to stay ahead in the development of our archiving system and find more partners in german publishing.

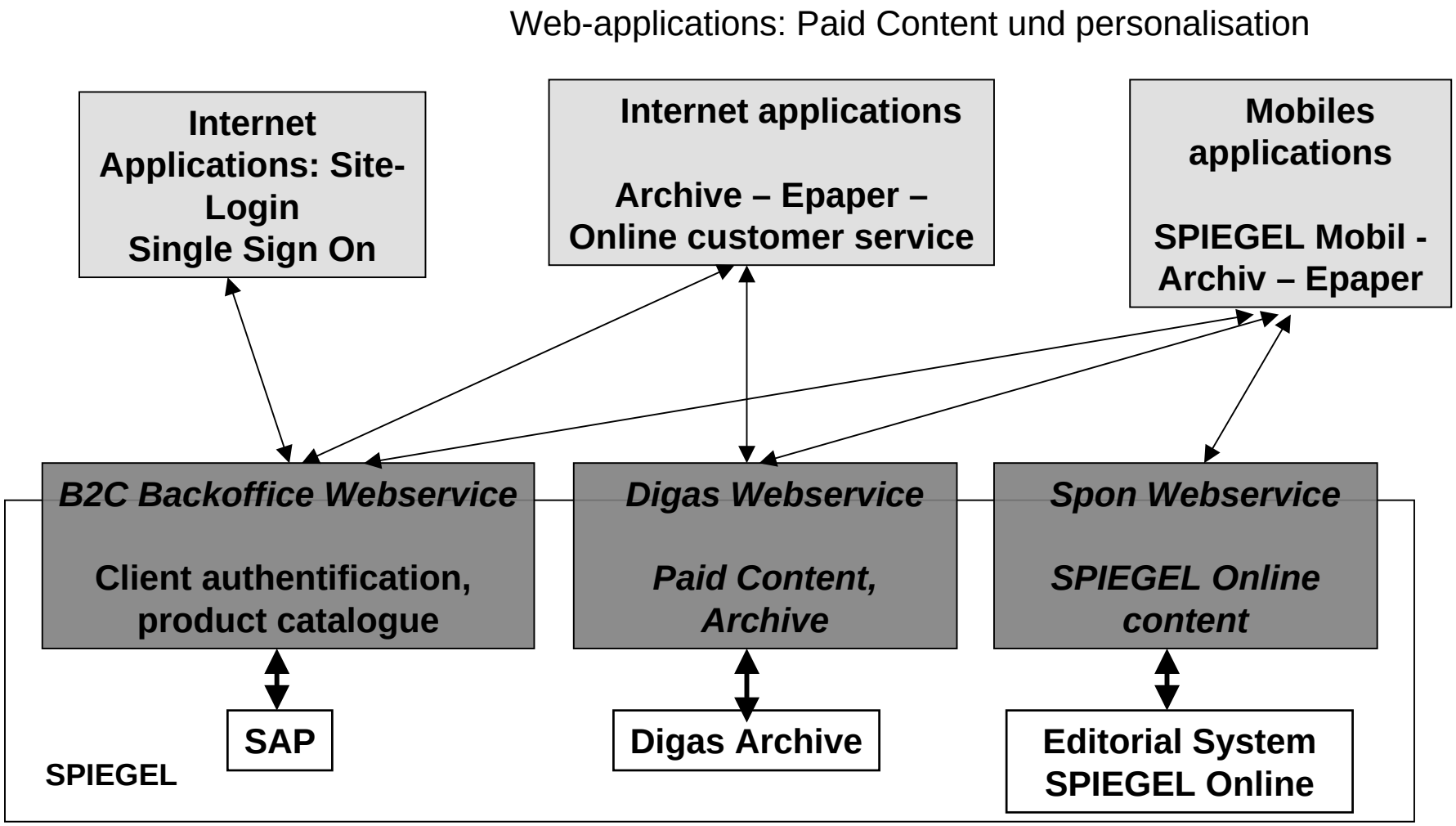
Digas Clients

- The client we rolled out in 2002 and which is still used by most researchers
- Live presentation via VPN



DER SPIEGEL

Digas Web Service – Part of the Service Oriented Architecture at DER SPIEGEL



DER SPIEGEL

Digas at SPIEGEL Online

- SPIEGEL Online is the largest German language news-site – around 400 Mio page impressions monthly.
- Based on the Service Oriented Architecture we developed many different „Digas“-client-applications like the site search which presents relevant Web-Dossiers (paid content) in the result-set, the E-Paper application and the new Tag Cloud („Themen des Tages“ – daily topics). Also based on these Web Services the SPIEGEL mobile site was developed by a company which specialises in mobile applications.
- These applications form the SPIEGEL Digital family of products which integrate paid content in the editorial content of SPIEGEL Online.
- There are up to 50 000 site searches at SPIEGEL Online, right now about 30 000 customers are using the paid content E-Paper and archive application either on top of their print subscription or only digitally.

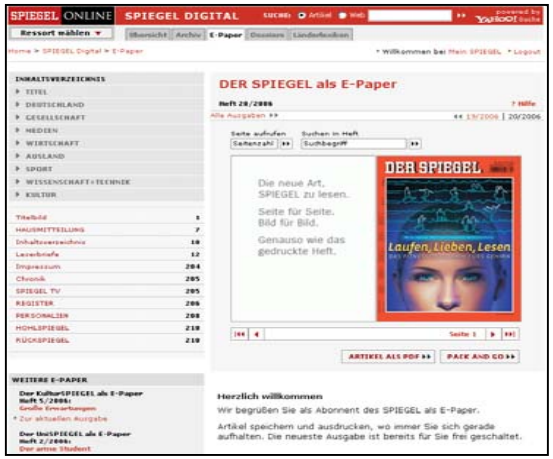
SPIEGEL Digital: Archive, Epaper, Web-Dossiers

- Archive

http://service.spiegel.de/digas/archiv

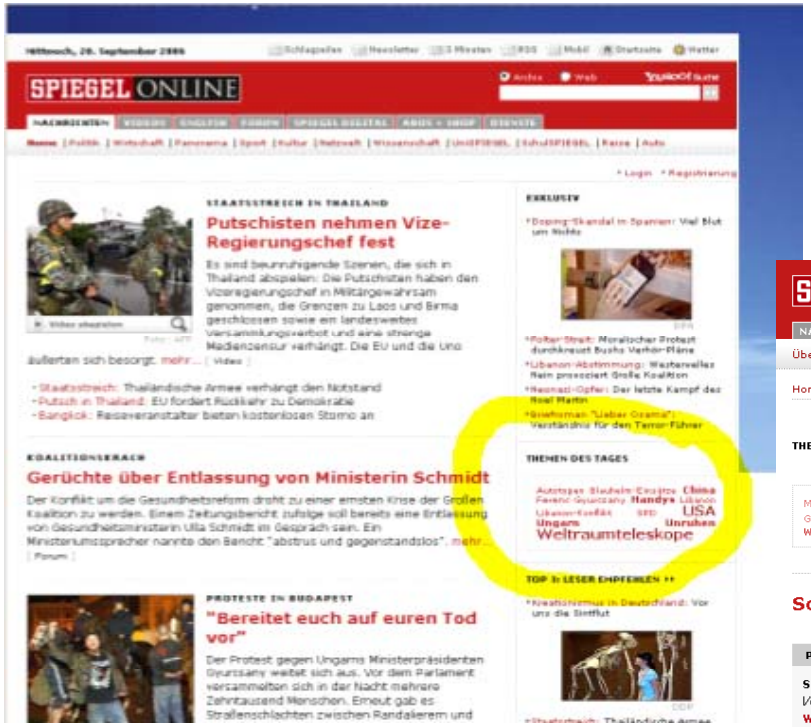
- E-Paper

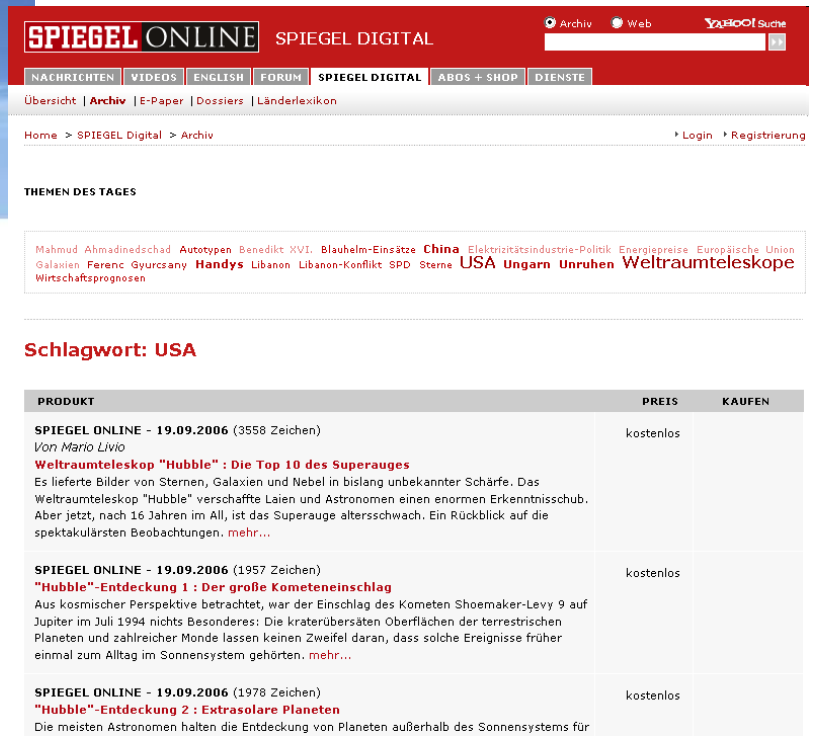
http://service.spiegel.de/digas/servlet/epaper?Q=SP



DER SPIEGEL

Site Search, archive and „Themen des Tages“ (Tag Cloud, daily topics) at SPIEGEL Online





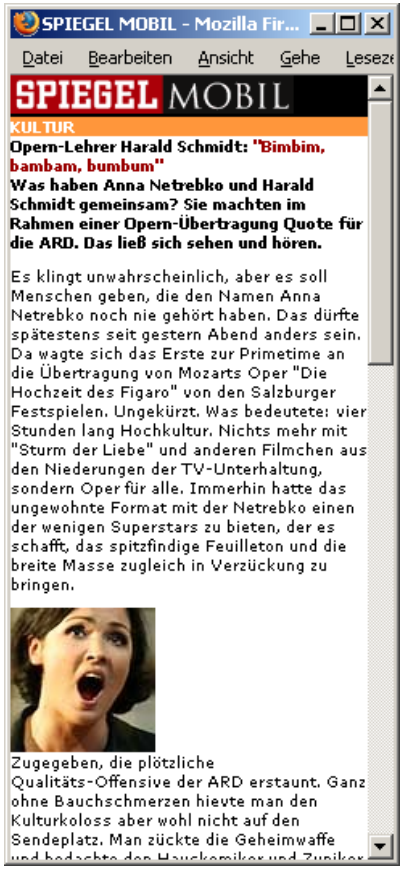
SPIEGEL mobil

- SPIEGEL Online news, optimized for different mobile devices

DER SPIEGEL

SPIEGEL mobil

- Articles and picture galleries



DER SPIEGEL

SPIEGEL mobil

- Video content for mobiles with broadband connection

DER SPIEGEL

SPIEGEL mobil

- SPIEGEL Archive and E-Paper mobile (for PDAs and Blackberries)

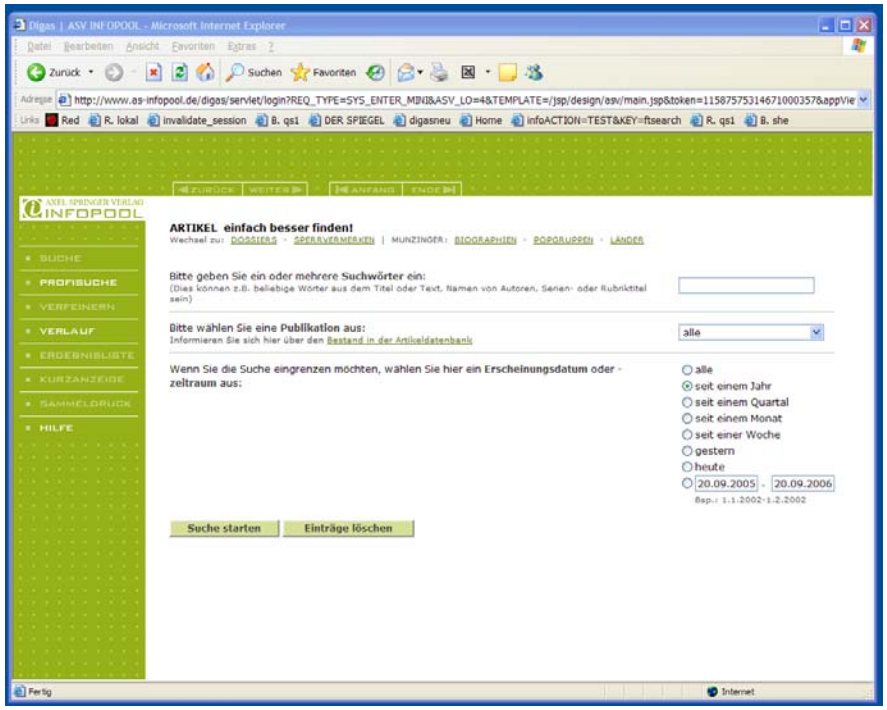
DER SPIEGEL

Digas Clients at Axel Springer Verlag

- Since the beginning of 2005 Axel Springer Verlag is using Digas as their archiving and research application. AS has their own client application for their editorial staff with a different look and feel but the same basic search functionality. The Digas Web-Service is used in the new editorial system built by Unisys for fully integrated research functionality in the editorial worklfow.

DER SPIEGEL

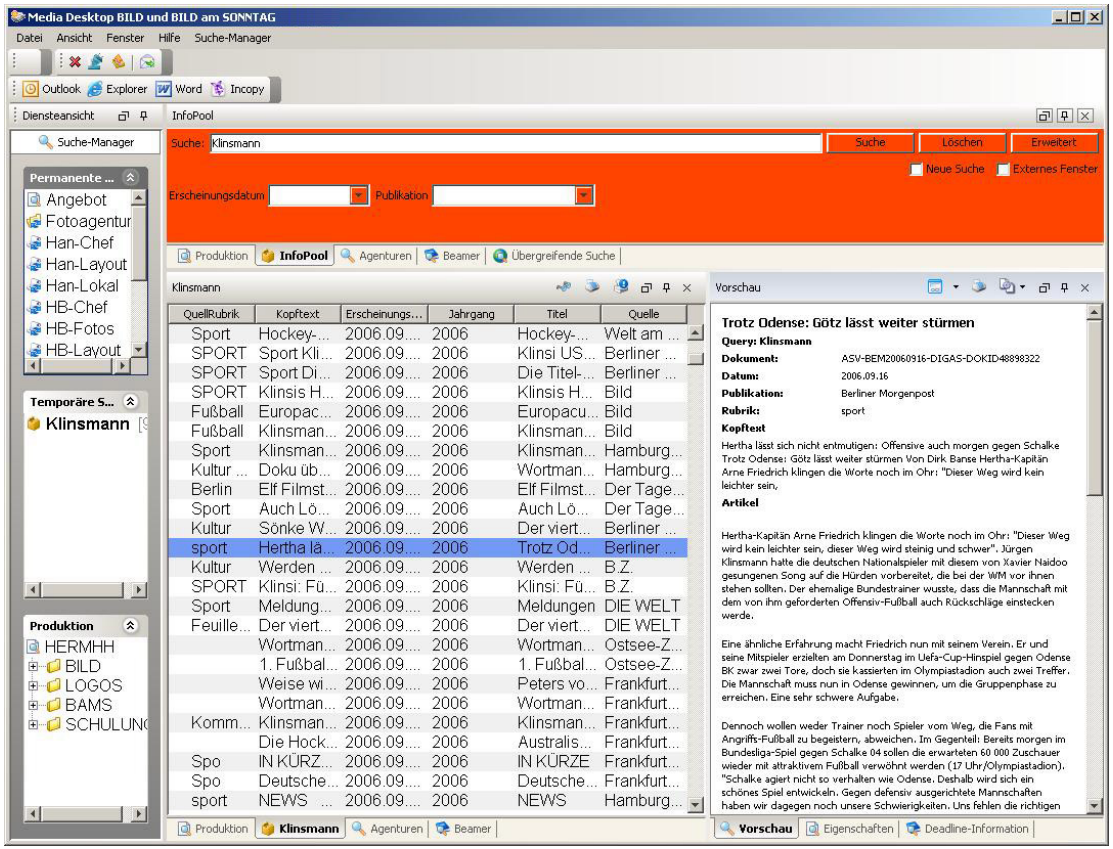
Digas-Archive at Axel Springer Verlag (AS)



- Different look and feel, same basic functionality

DER SPIEGEL

Digas Webservice: „Media Desktop“, part of the editorial system developed for the Bild-Zeitung



- The editorial system of the Axel Springer Verrlag (developed by Unisys) connects via Digas Webservice to the SPIEGEL Archive.

DER SPIEGEL

Document base, Hardware, activity

- Until 2 Years ago: HP Superdome, 48 GB memory, 16 CPU, comparex 7500 disk array.
- 2001 we had 10 million documents, around 1000 queries per day.
- Today: P690, 64 GB memory (20 GB buffer cache), 8 CPU, Hitachi 9585 disk array
- 2006 our archive consists of 32 million documents, our users make up to 10.000 queries per day, up to 2500 queries per hour.
- We are committed to an average query performance <3 seconds (and 70% of the queries must be <3 seconds, 80% <6 seconds, 90% <12s, 98% <30, 99% <60s).

DER SPIEGEL

Tools

- „Old“ Digas:
  Tomcat 4.x, CORBA (JDK1.4)

- Today in production:
  JBoss 4.0.3, EJB 2.1, Axis 1.4 for Webservices, Java 5

- Applications released in 2007:
  JBoss 4.0.4, EJB 3.0 (JSR181 WebServices), Java 5

DER SPIEGEL

Oracle text version

- In Production: 10.1, will be updated to 10.2 shortly
- 11g beta: waiting for a release which runs on red hat 4

DER SPIEGEL

Projects in 2007

- Improvement of the categorizing-process. One partner in doing this would be picturesafe, are there others?
- RAC with Oracle Text? Licensing?
- 11g: Index Management: Online-Indexing on commit, manage fragmentation via shadow index
- 11g: Composite Index, how can we improve the performance of dossier searches
- 11g: User defined scoring

DER SPIEGEL

Contact:

- Heiner Ulrich
Head of IT, Editorial Systems
DER SPIEGEL
heiner_ulrich@spiegel.de