Verwaltungsdiensteportal service-bw
Systemarchitektur und semantische Techniken

Workshop Intelligente Portale, Karlsruhe, 31.05.2011

T Systems

Verwaltungsdiensteportal service-bw

Systemarchitektur - Schichtenmodell.

service-bw und mein service-bw wurden als 3-Schicht Architektur umgesetzt.

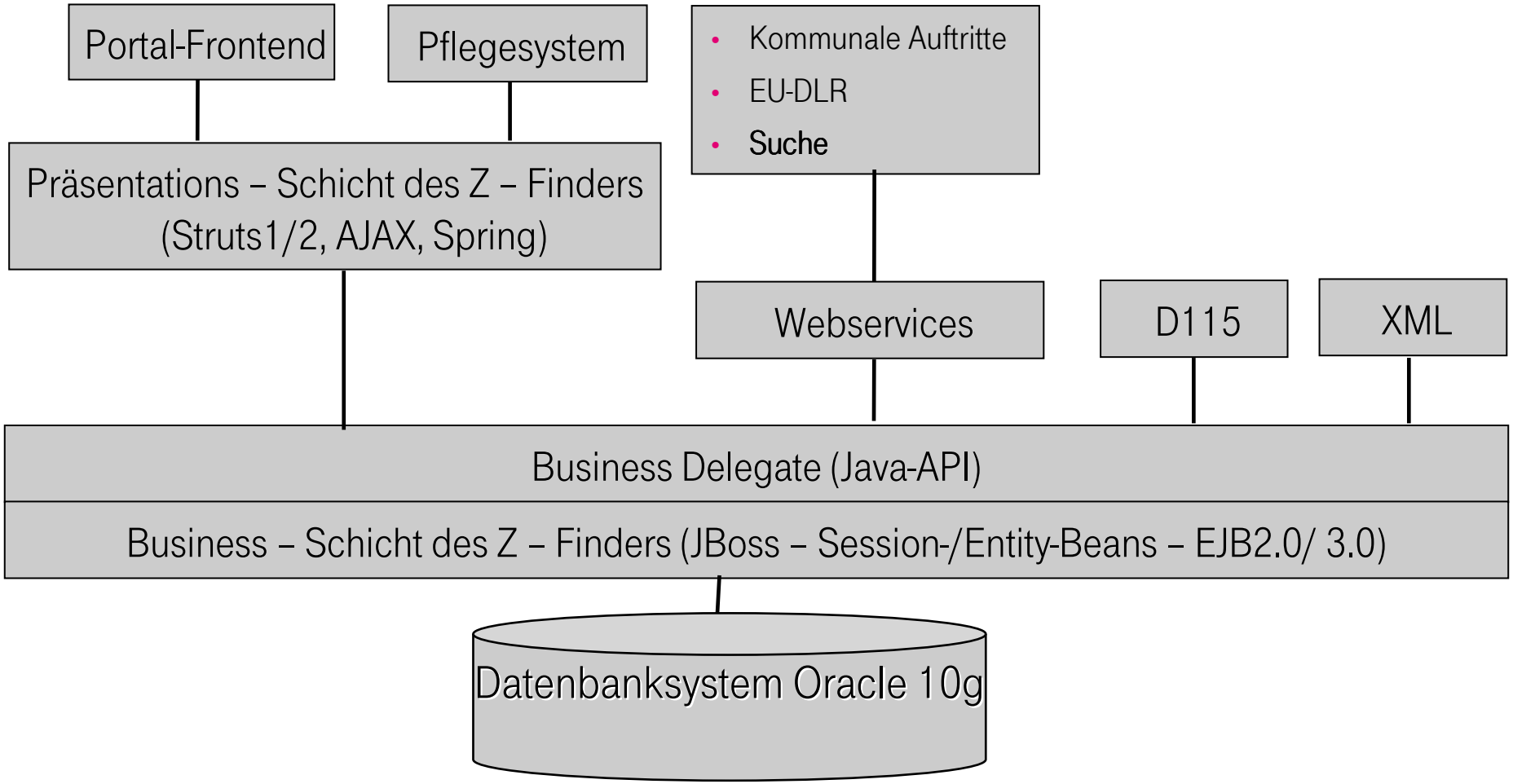
- Präsentationsschicht – Diese ist für die Repräsentation der Daten und den Empfang von Benutzereingaben verantwortlich und stellt die Benutzerschnittstelle dar.
- Logikschicht (Businessschicht) – Sie beinhaltet alle Verarbeitungsmechanismen. Hier ist die Anwendungslogik vereint.
- Datenhaltungsschicht – Sie enthält die Datenbank und ist verantwortlich für das Speichern und Laden von Daten.

Vorteile von n-tier Architekturen

- Geringe Komplexität der einzelnen Teile
- Verteilung der Implementierungsaufgaben
- Erleichterte Wartbarkeit (keine Client-Software, Austausch von Versionen)
- Skalierbarkeit, Sicherheit

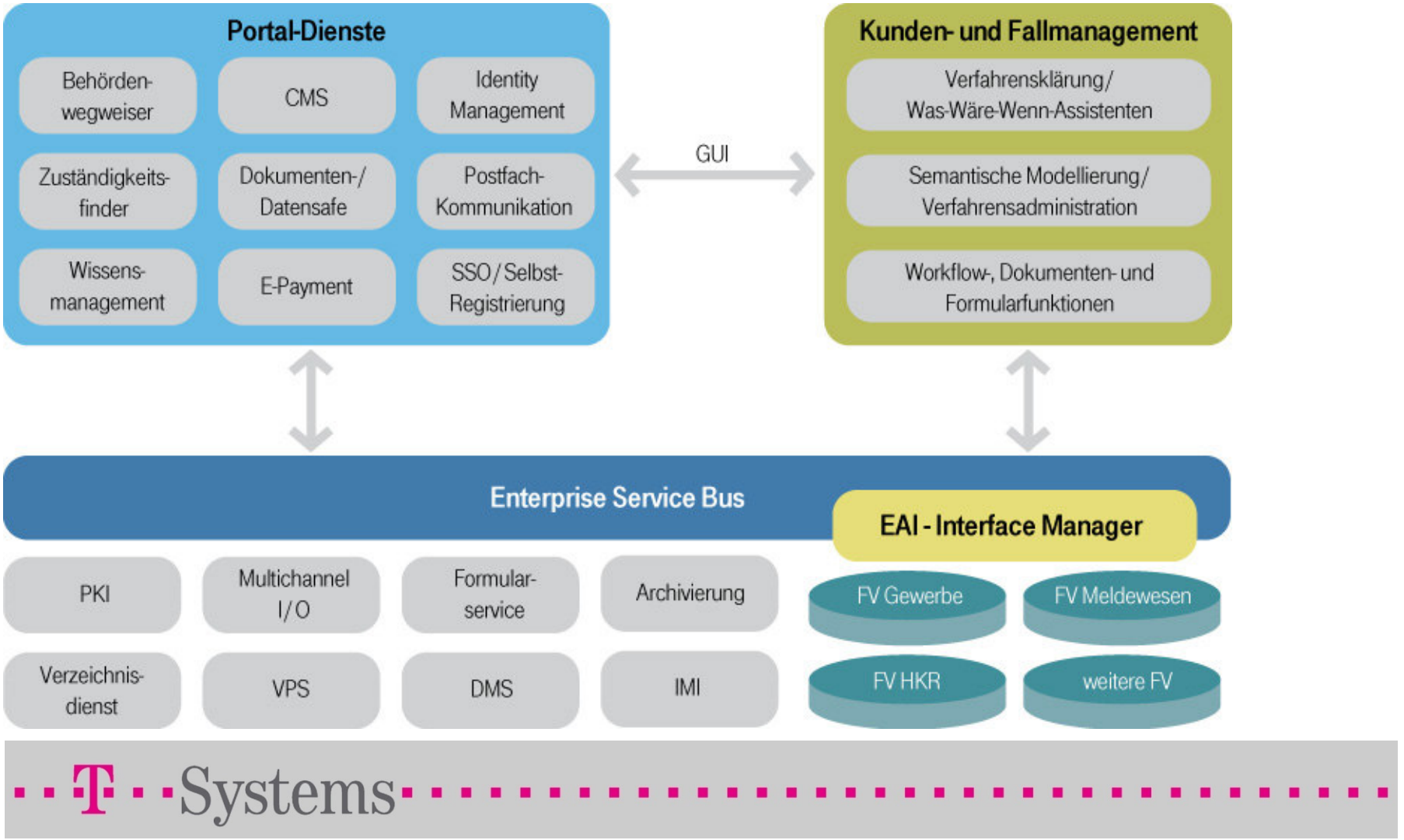
Verwaltungsdiensteportal service-bw.

Systemarchitektur – Übersicht



Verwaltungsdiensteportal service-bw

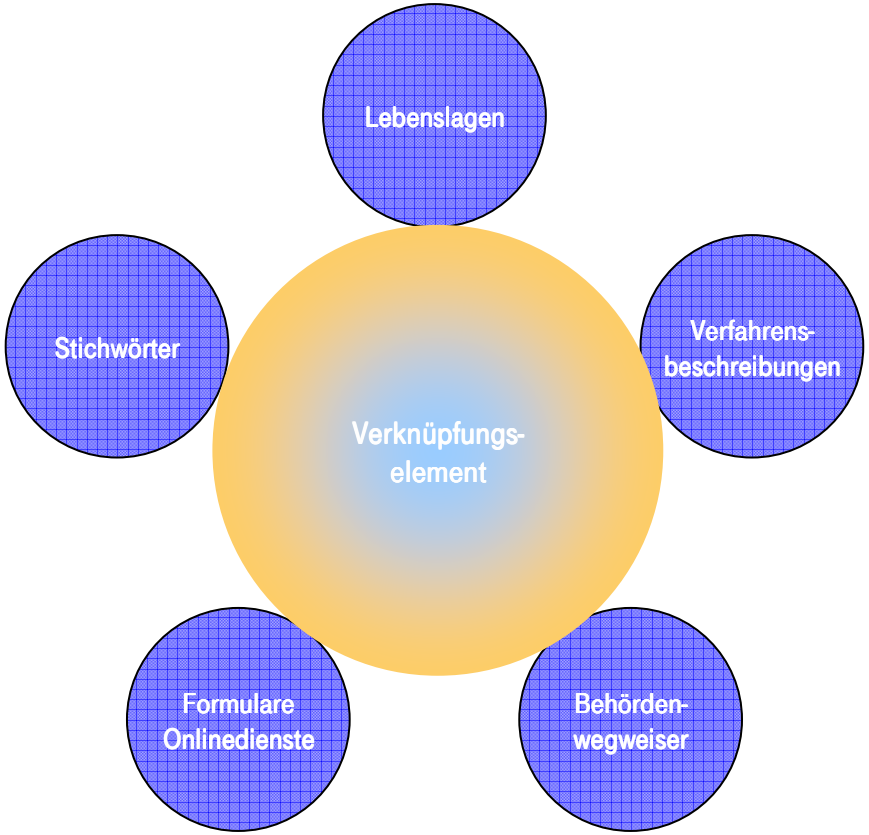
Umsetzung der EU-DLR - Komponenten



Verwaltungsdiensteportal service-bw

Systemarchitektur – Datenmodell.

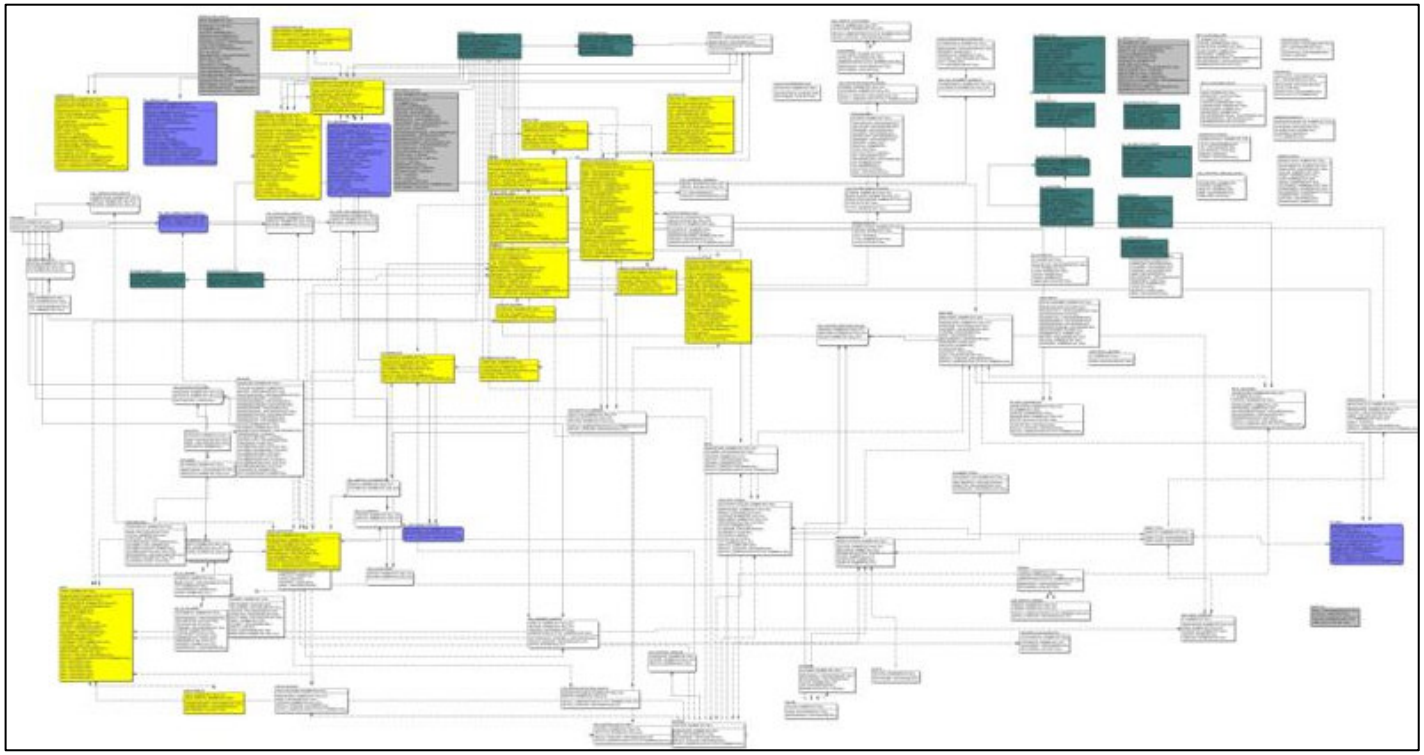
- service-bw stellt keine Sammlung statischer HTML-Seiten dar.
- Seiteninhalte werden zur Laufzeit, z.B. abhängig von der Regionalisierung des Nutzers, dynamisch generiert.
- Die Verbindung aller Inhaltstypen erfolgt durch ein technisches Datenbankelement („Verknüpfungselement“).



Verwaltungsdiensteportal service-bw

Systemarchitektur – Datenmodell.

- Die Vielzahl von Inhaltstypen und deren Beziehungen führt zu einem komplexen relationalen Datenmodell:



Verwaltungsdiensteportal service-bw

Semantische Techniken – Ausgangssituation.

Die Suche in service-bw erfolgte vor dem Relaunch (08/2009) mit Oracle-Text.

- Mit Hilfe von Oracle-Text wurde eine Volltextindexierung der Tabellen von service-bw durchgeführt. Die Suche erfolgte im Anschluss im Index und konnte sehr performant durchgeführt werden.

Problem: Da bei der Volltextsuche einfach ein Stringvergleich im Index durchgeführt wurde, wurden bei einer Suche nach „KFZ“ keine Treffer gefunden, die das Wort „Kraftfahrzeug“ enthielten und umgekehrt. Ähnlich verhielt es sich bei Tippfehlern bzw. unterschiedlicher Schreibweise von Suchbegriff und Begriff in service-bw (Bsp.: „Einkommensteuer“ und „Einkommenssteuer“).

Bei der Suche mit Oracle Text wurde die Bedeutung des Suchbegriffs nicht ausgewertet.

Verwaltungsdiensteportal service-bw

Semantische Techniken – Vorteile Ontologische Suche.

- Inhalte werden über die Semantik (Ontologie) auch dann gefunden, wenn diese den Suchbegriff nicht explizit beinhalten.
  Beispiel: Für die Suchanfrage „Umzug Stuttgart“ kann der Bezug zur Verfahrensbeschreibung „Autoanmeldung“ direkt gefunden werden, da diese in einem logischen Zusammenhang steht.
- Modell verbessert die Relevanz der gefundenen Inhalte, z.B. Anzeige an oberster Stelle oder das Ausschließen von Inhalten
- Phrasenerkennung, d.h. natürlich sprachliche Suchanfragen werden interpretiert und liefern direkt den relevanten Inhalt, wie z.B. eine bestimmte Lebenslage.
- Durch automatische Tippfehlererkennung werden relevante Treffer gefunden, selbst wenn der Suchende sich bei der Eingabe vertippt hat. Bei Nulltreffer-Suchergebnis Vorschlag alternativer Suchbegriffe „oder meinten Sie“?
- Automatische Vervollständigung der Suchbegriffe, die den Nutzer von vorneherein die passenden Begriffe vorschlägt und die Gefahr von Fehleingaben minimiert.

Verwaltungsdiensteportal service-bw

Semantische Techniken – Der Ontologiebegriff.

Definition:

„Eine Ontologie stellt [...] ein formalisiertes Modell der Welt oder eines Teils der Welt dar, über deren Begriffe und Zusammenhänge eine Gruppe von Experten / Nutzern Einigkeit erreicht hat.“ (H. Oppermann, H.-P. Schnurr, R. Studer in: „wissensmanagement“ 6/2001, S. 34)

Merkmale einer Ontologie:

- Trennung zwischen
  - *Konzept / Klasse* (= Idee)

  und
  - *Instanz* (=Objekt der realen Welt)
- Hierarchie der Ideen

Verwaltungsdiensteportal service-bw

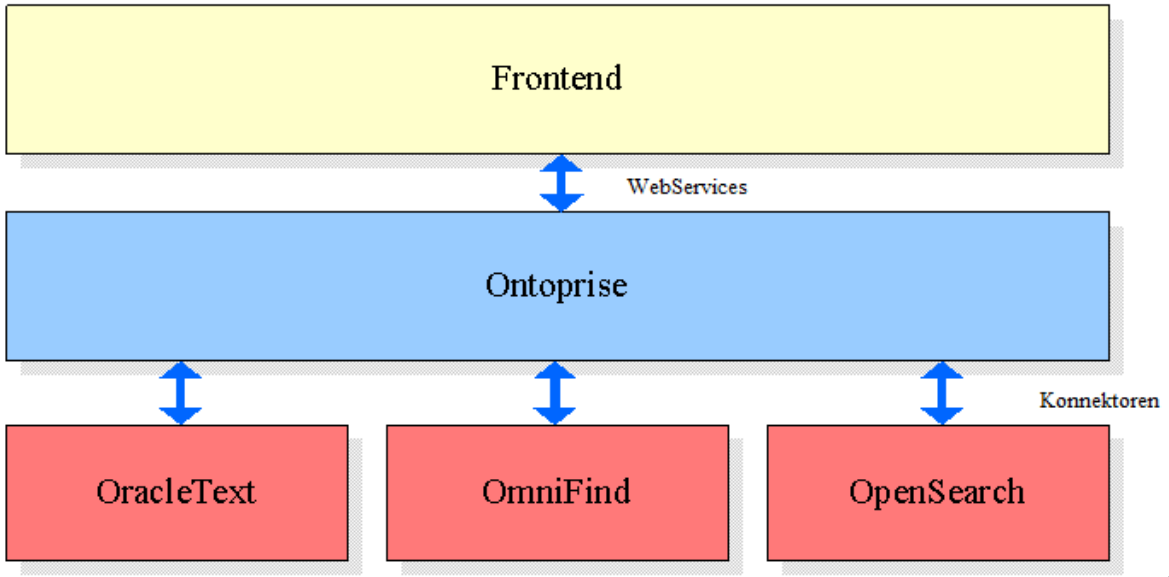
Semantische Techniken – Konzepte in service-bw.

- Für service-bw sind die Konzepte „Orte“, „Lebenslagen“, „Stichwörter“ und „Tätigkeiten“ in Ontoprise hinterlegt.
- Diese Konzepte werden regelmäßig einmal täglich mit den Daten aus service-bw aktualisiert. Eine manuelle Pflege der Konzepte ist somit nicht notwendig.
- Das Konzept „Orte“ dient der automatischen Regionalisierung
- Die Konzepte „Lebenslagen“ und „Stichwörter“ dienen u.a. als Wörterbuch für die Tippfehlerkorrektur sowie als Grundlage für die AJAX-Vorschlagsliste.
- Das Konzept „Stichwörter“ (inkl. Stichwortsynonyme) dient der sog. Query-Expansion. Gibt der Nutzer z.B. den Suchbegriff „heirat“ ein, dann generiert Ontoprise hieraus die Suchanfrage „(heirat ODER Ehegatten ODER Eheschließung ODER Hochzeit ODER Trauung ODER Vermählung)“

Verwaltungsdiensteportal service-bw

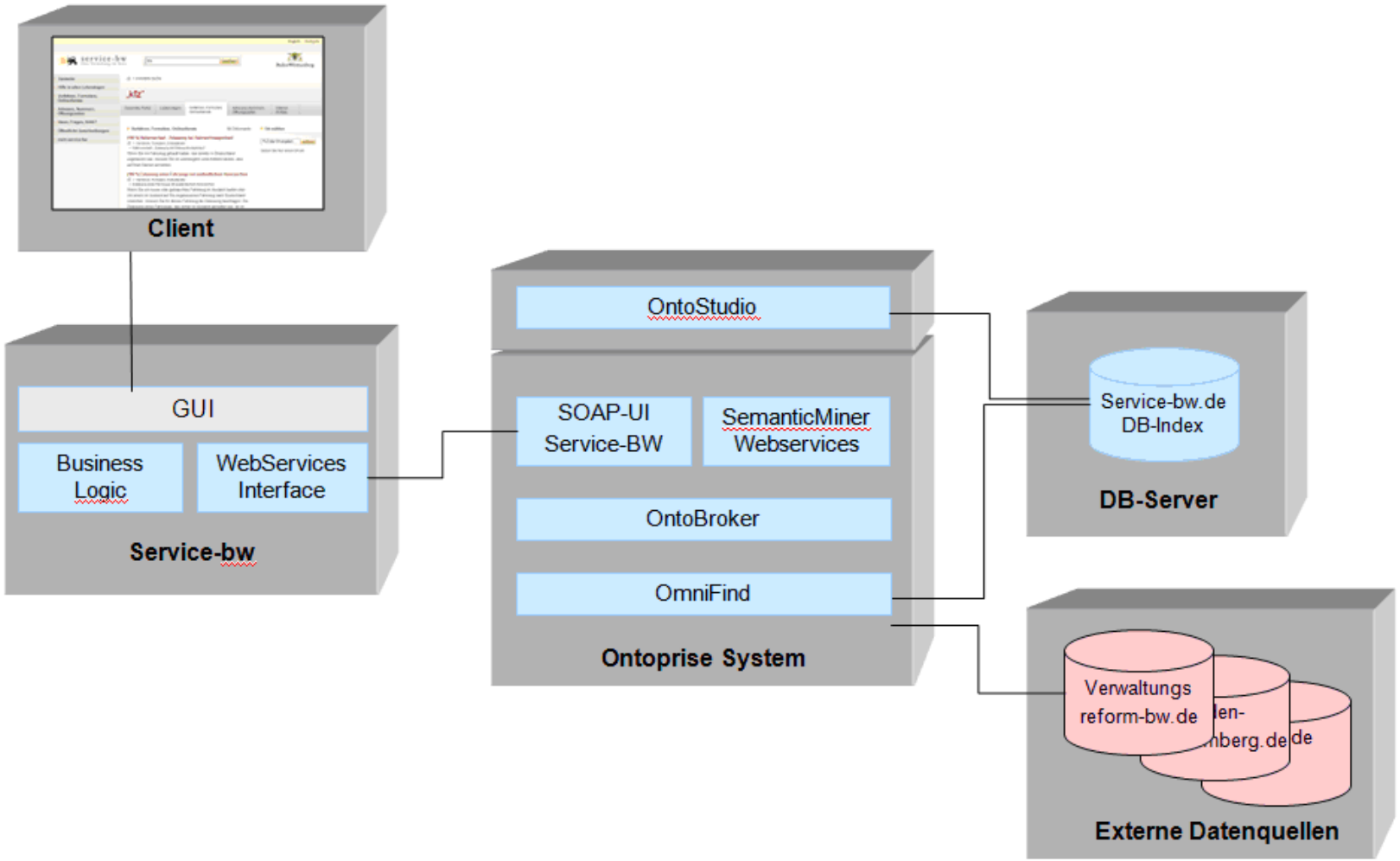
Semantische Techniken – Architektur.

- Ontoprise dient der Aufbereitung der Suchanfragen und Suchergebnisse
- Das Crawlen und Indexieren wird für die Inhalte von service-bw nach wie vor von OracleText übernommen, für die Fremdportalinhalte kommt IBM-OmniFind zum Einsatz. Zusätzlich wird zur Integration bestimmter Fremdportale (z.B. Umweltportal Baden-Württemberg) eine OpenSearch-Schnitt-stelle integriert.



Verwaltungsdiensteportal service-bw

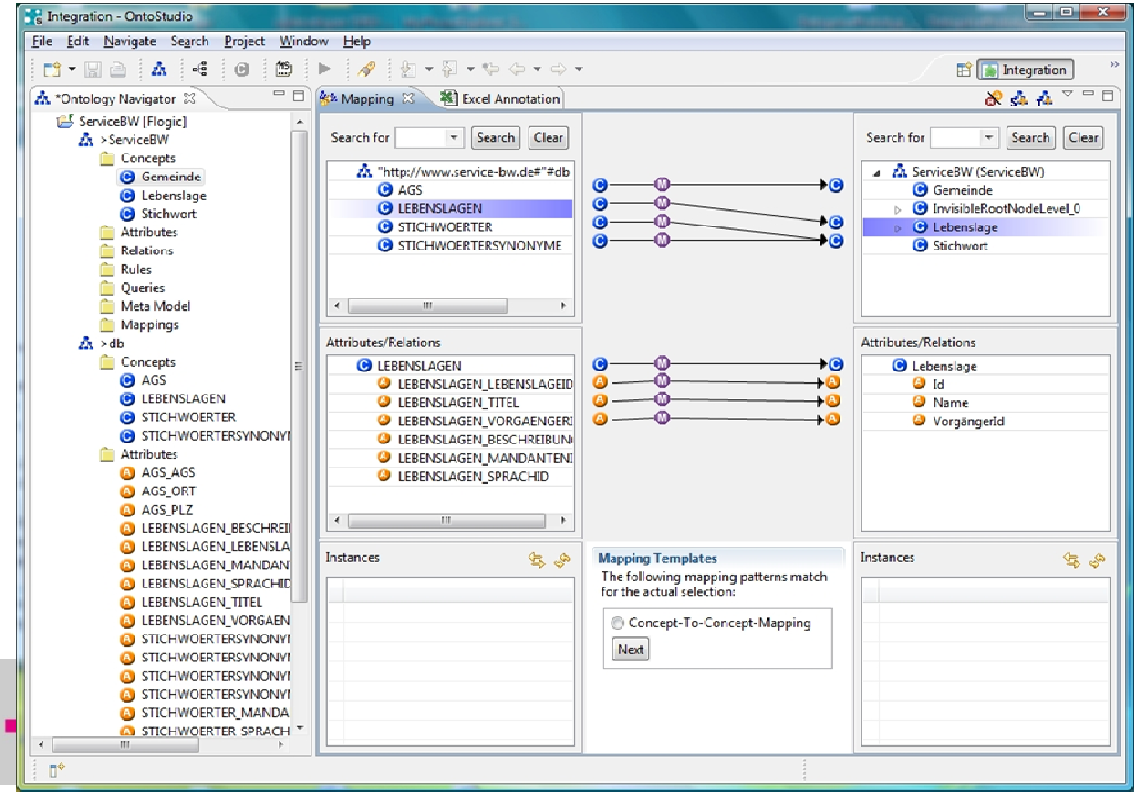
Semantische Techniken – Komponenten.



Verwaltungsdiensteportal service-bw

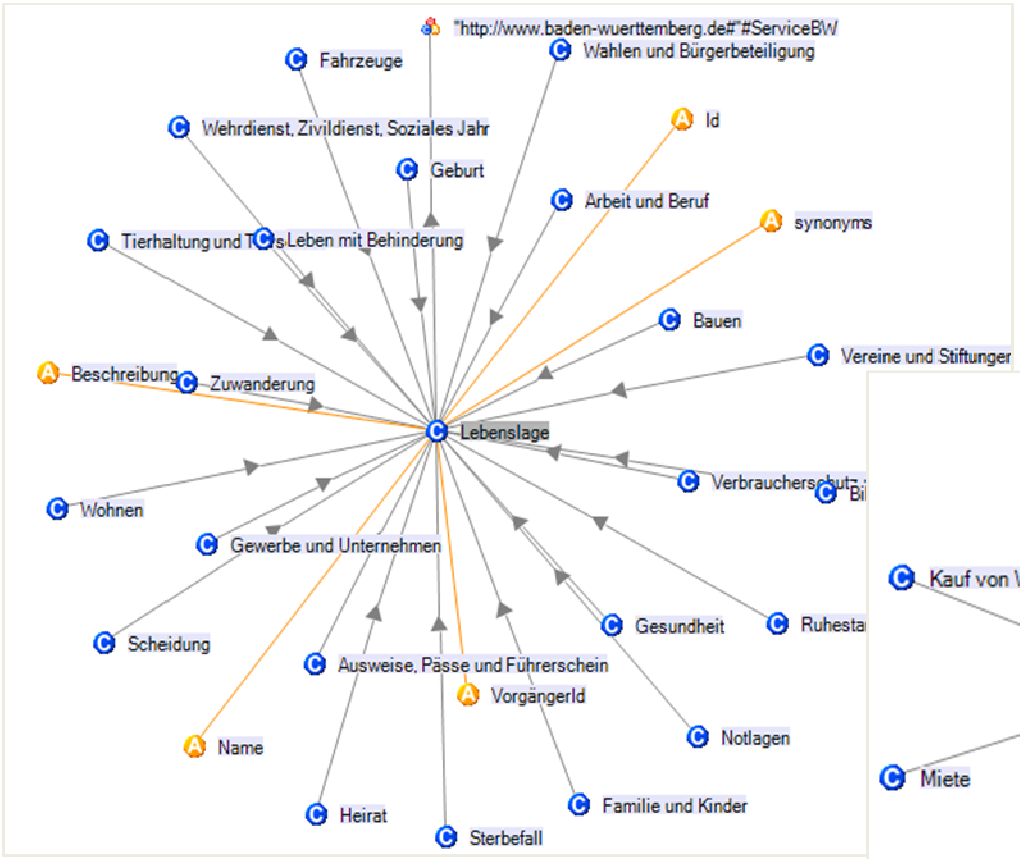
Semantische Techniken – Komponenten: Ontostudio

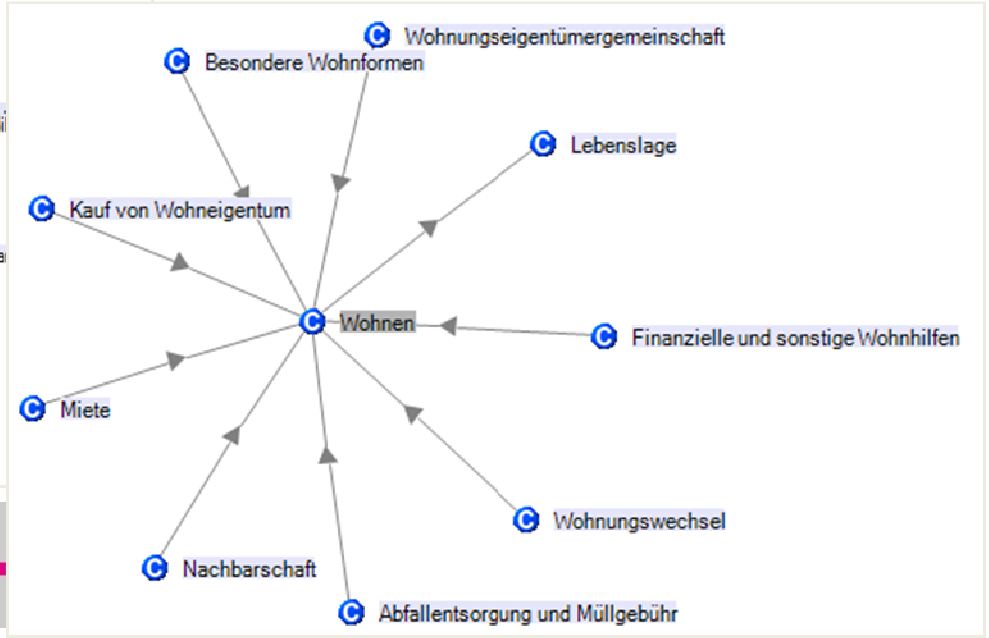
- OntoStudio ist eine Modellierungsumgebung zum Aufbau und zur Pflege von Ontologien. Mit Hilfe von OntoStudio können in vielfältiger Art und Weise Ontologien aufgebaut, angepasst und erweitert werden.
- Die im OntoStudio definierten Strukturen sind dann die Basis für den OntoBroker.



Verwaltungsdiensteportal service-bw

Semantische Techniken – Ausschnitte aus der Ontologie





Verwaltungsdiensteportal service-bw

Semantische Techniken – Komponenten: OntoBroker

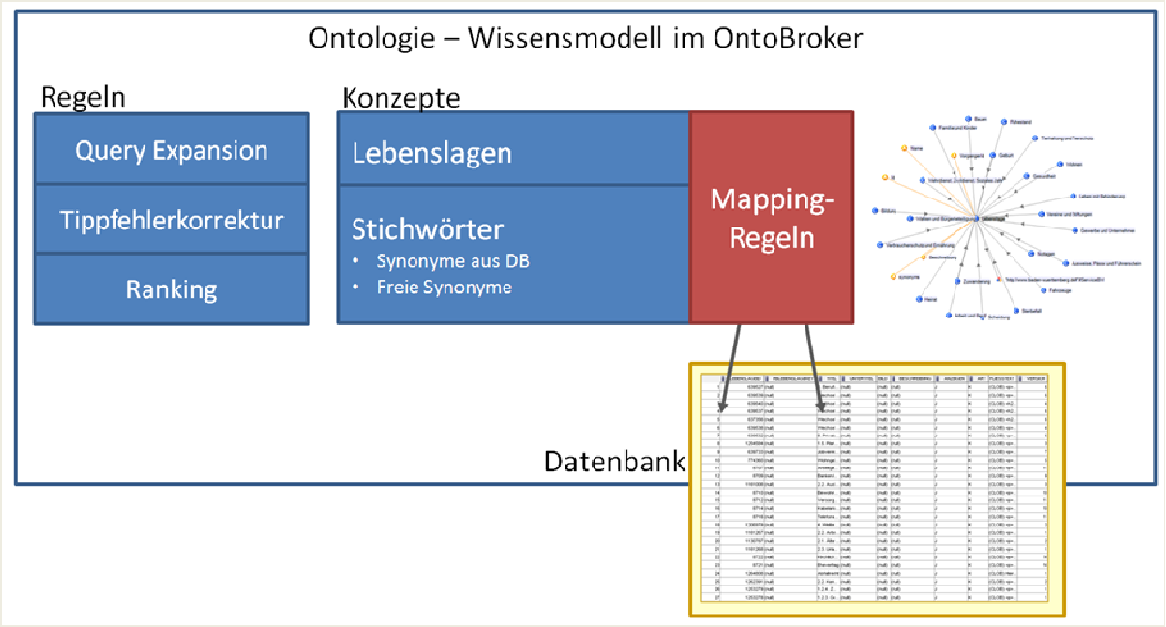
Die Komponente Ontobroker

- ist eine umfassende Semantic-Middleware.
- ist Schnittstelle zwischen service-bw und der Ontologie
- stellt Webservices zur Verfügung, die es ermöglichen, Abfragen an die Ontologie zu richten und aus den dort hinterlegten Informationen und Regeln Wissen zu generieren.
- ist für das physikalische Mapping zwischen der service-bw-Datenbank und der Ontologie zuständig. Darüber ist es möglich, mit einem automatischen, zyklischen Synchronisations-Prozess die Informationen aus der service-bw-Datenbank in die Ontologie zu überführen.

Verwaltungsdiensteportal service-bw

Semantische Techniken – Komponenten: OntoBroker(2)

- Bestandteile der Ontologie:



Verwaltungsdiensteportal service-bw

Semantische Techniken – Ablauf einer Suchanfrage

1. Analyse des Suchstrings („Parsing“):
   Suchstring (z.B. „Hochzet Heidelberg)“ wird an speziellen Webservice übergeben. Dieser führt mit Hilfe der in der Ontologie enthaltenen Konzepte und der übergebenen Sprache eine Tippfehlerkorrektur durch und liefert die passenden Konzepte mitsamt ihren Synonymen zurück. „Hochzet“ wird als „Hochzeit“ erkannt, die ein Synonym des Stichwortes „Heirat“ ist. Weitere Synonyme sind z. B. „Braut“, „Bräutigam“, „Ehegatten“ und „Trauung“.
2. Bereinigung der Ergebnisse:
   Zu jedem Suchbegriff wird nur der Begriff ausgewählt, der am Besten passt. Im Beispiel wird nur „Heirat“ und „Heidelberg“ zurückgeliefert.
3. Analyse des bereinigten Ergebnisses:
   Die übriggebliebenen Begriffe werden untersucht, ob es sich um Gemeinden handelt. Ist dies der Fall, wird der Begriff automatisch aus dem Suchstring entfernt und stattdessen für die Regionalisierung der Suchanfrage verwendet.

Verwaltungsdiensteportal service-bw

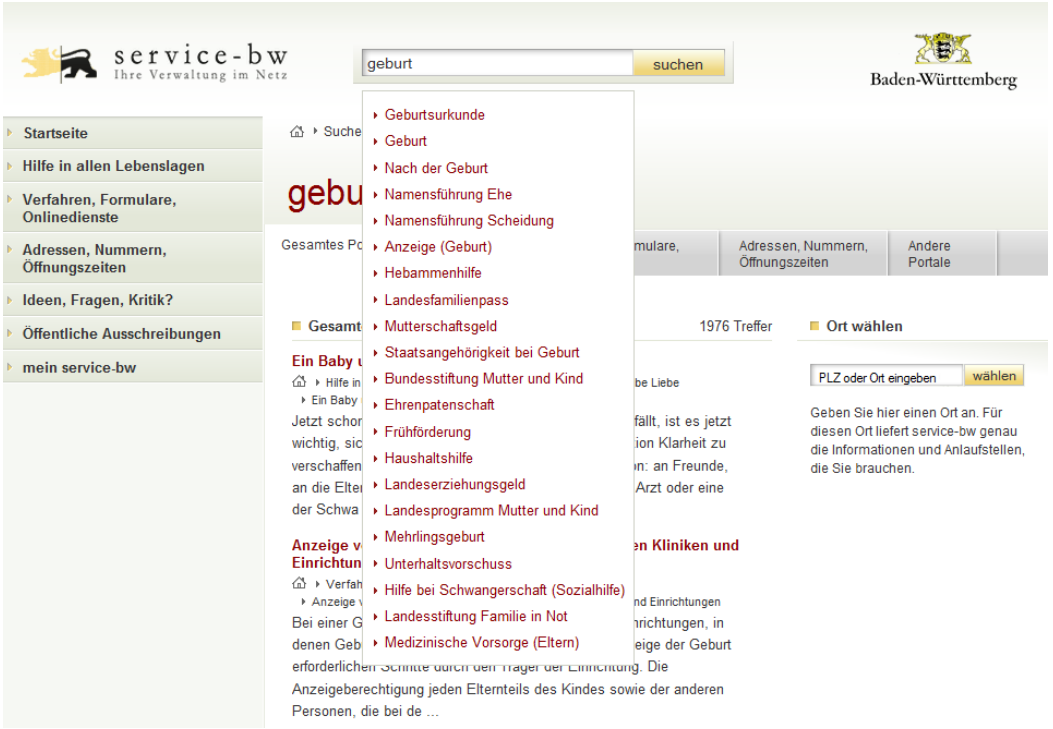
Semantische Techniken – Ablauf einer Suchanfrage(2)

4. Reduzierter Suchstring wird an Webservice übergeben, der daraus ein Anfragemodell erstellt, das er gemeinsam mit einer lesbaren Repräsentation zurückgibt. In unserem Beispiel lautet diese Repräsentation:
(Heirat ODER Braut ODER Bräutigam ODER Ehegatten ODER Eheschließung ODER Hochzeit ODER Trauung ODER Verlobter ODER Vermählung)
5. Dieses Anfragemodell wird dann, gemeinsam mit weiteren Parametern wie z. B. der Sprache, dem ermittelten Gemeindeschlüssels oder einer Mandanten-Id an die Suchwebservices übergeben. Diese können entweder über alle Objekttypen suchen, oder sich auf einen bestimmten Bereich beschränken. Bspw. kann über alles gesucht werden, oder nur nach Formularen und Verfahren, nach Lebenslagen oder nach Verwaltungseinheiten.
6. In den Suchwebservices wird das Anfragemodell in eine Datenbankabfrage übersetzt. Bei der Datenbankabfrage kann auf die einzelnen Felder der verschiedenen Objekte gewichtet zugegriffen werden (Boosting-Faktoren).

Verwaltungsdiensteportal service-bw

Semantische Techniken – Einbindung Ontoprise.

- Einbindung der ontologischen Suche Ontoprise
  - **AJAX-Vorschlagsliste**
  - **Semantische Suche, Suche nach Synonymen**
  - „Meinten Sie?“-Funktionalität
  - Automatische Tippfehlerkorrektur
  - Regionalisierung anhand Ortsnamen im Suchstring („Hochzeit Stuttgart“)
  - Suchergebnisse mit kurzem Anriss
  - Einbindung von Fremdportalen über Omnifind



Verwaltungsdiensteportal service-bw

Semantische Techniken – Einbindung Ontoprise(2).

- Einbindung der ontologischen Suche Ontoprise
  - AJAX-Vorschlagsliste
  - Semantische Suche, Suche nach Synonymen
  - **„Meinten Sie?“-Funktionalität**
  - Automatische Tippfehlerkorrektur
  - Regionalisierung anhand Ortsnamen im Suchstring („Hochzeit Stuttgart“)
  - Suchergebnisse mit kurzem Anriss
  - Einbindung von Fremdportalen über Omnifind

Verwaltungsdiensteportal service-bw

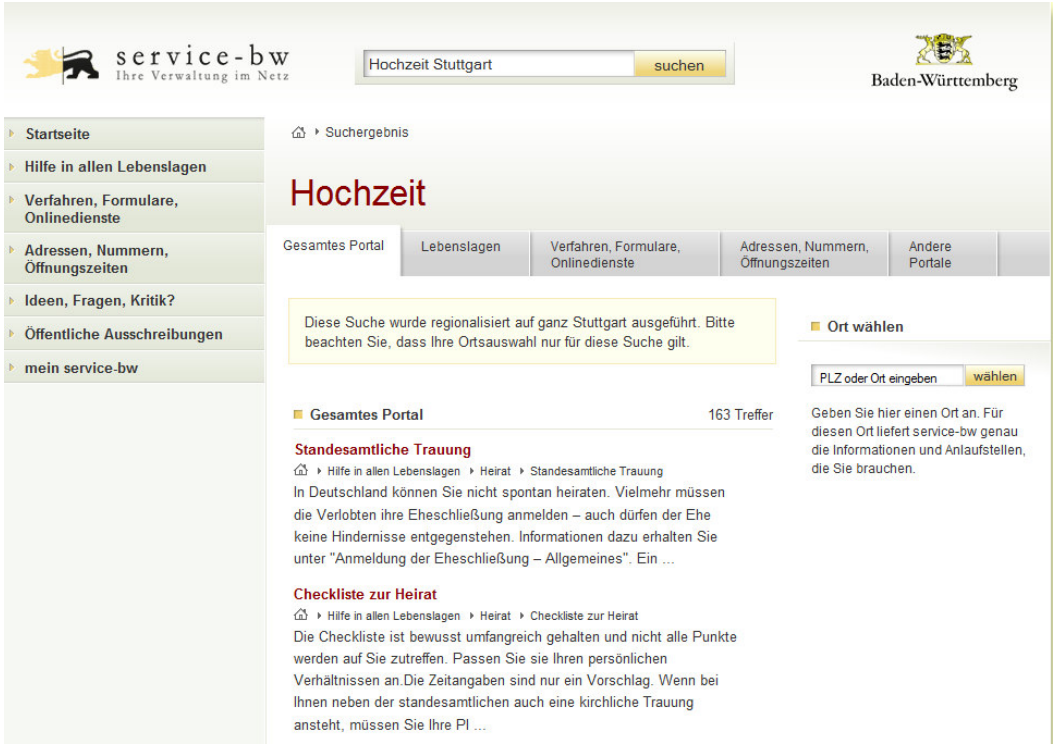
Semantische Techniken – Einbindung Ontoprise(3).

- Einbindung der ontologischen Suche Ontoprise
  - AJAX-Vorschlagsliste
  - Semantische Suche, Suche nach Synonymen
  - „Meinten Sie?“-Funktionalität
  - Automatische Tippfehlerkorrektur
  - Regionalisierung anhand Ortsnamen im Suchstring („Hochzeit Stuttgart“)
  - Suchergebnisse mit kurzem Anriss
  - Einbindung von Fremdportalen über Omnifind

Verwaltungsdiensteportal service-bw

Semantische Techniken – Einbindung Ontoprise(4).

- Einbindung der ontologischen Suche Ontoprise
  - AJAX-Vorschlagsliste
  - Semantische Suche, Suche nach Synonymen
  - „Meinten Sie?“-Funktionalität
  - Automatische Tippfehlerkorrektur
  - **Regionalisierung anhand Ortsnamen im Suchstring („Hochzeit Stuttgart“)**
  - **Suchergebnisse mit kurzem Anriss**
  - Einbindung von Fremdportalen über Omnifind



Verwaltungsdiensteportal service-bw

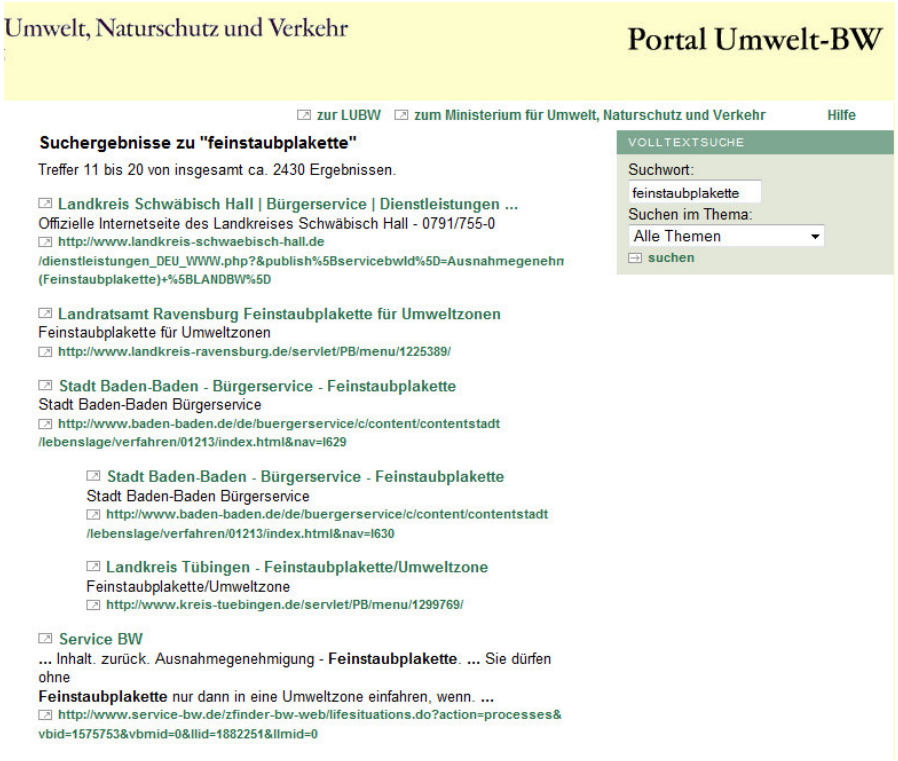
Semantische Techniken – Einbindung Ontoprise(5).

- Einbindung der ontologischen Suche Ontoprise
  - AJAX-Vorschlagsliste
  - Semantische Suche, Suche nach Synonymen
  - „Meinten Sie?“-Funktionalität
  - Automatische Tippfehlerkorrektur
  - Regionalisierung anhand Ortsnamen im Suchstring („Hochzeit Stuttgart“)
  - Suchergebnisse mit kurzem Anriss
  - Einbindung von Fremdportalen über Omnifind

Verwaltungsdiensteportal service-bw

Semantische Techniken – Schnittstellen.

- Eingehende Schnittstellen:
  - Webservice-Schnittstelle: Die Suche in service-bw kann über eine Web-Service-Schnittstelle angesprochen werden.
  - Im Rahmen des Projekts D115 wurde zusätzlich eine OpenSearch-Schnittstelle integriert, die es ermöglicht auf einfache Art und Weise eine regionalisierte Suche innerhalb der Verfahrensbeschreibungen von service-bw durchzuführen. Diese Schnittstelle wird inzwischen auch vom Umweltportal Baden-Württemberg genutzt, um bei einer Suche im Umweltportal auch Treffer von service-bw anzuzeigen.



Verwaltungsdiensteportal service-bw

Semantische Techniken – Schnittstellen(2).

- Ausgehende Schnittstellen:
  - In Kooperation mit dem Umweltportal Baden-Württemberg wurde von Ontoprise ein OpenSearch-Connector entwickelt, der es ermöglicht, die auf Google-Search-Appliance basierende Suchmaschine des Umweltportals anzusprechen und die Suchergebnisse in die Ergebnisliste „Andere Portale“ zu integrieren.

Verwaltungsdiensteportal service-bw

Semantische Techniken – Suche EU-DLR.

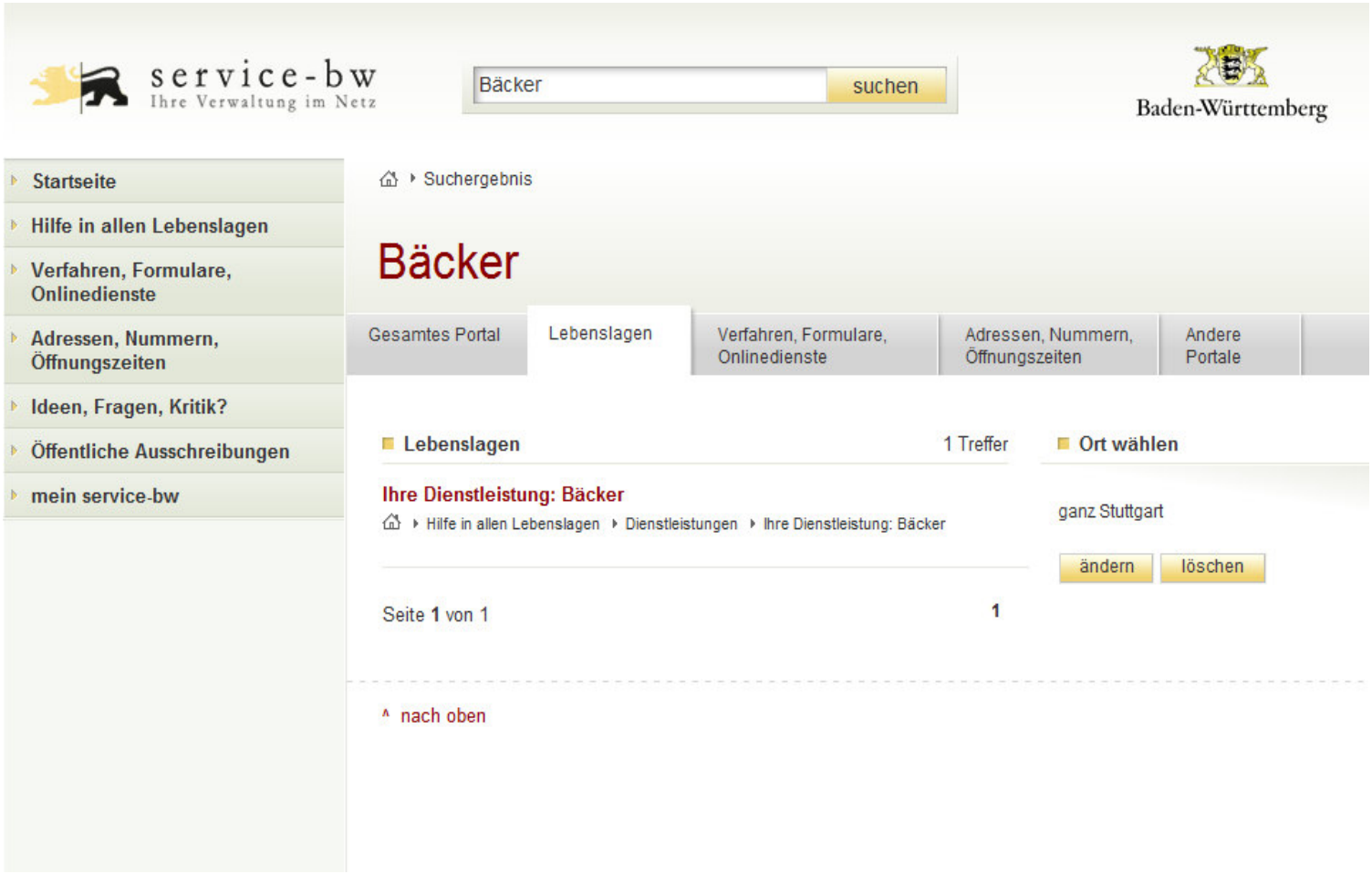
- EU-DLR
  - Suche nach Tätigkeiten
  - Automatische Übernahme von Tätigkeiten

Verwaltungsdiensteportal service-bw

Semantische Techniken – Suche EU-DLR(2).

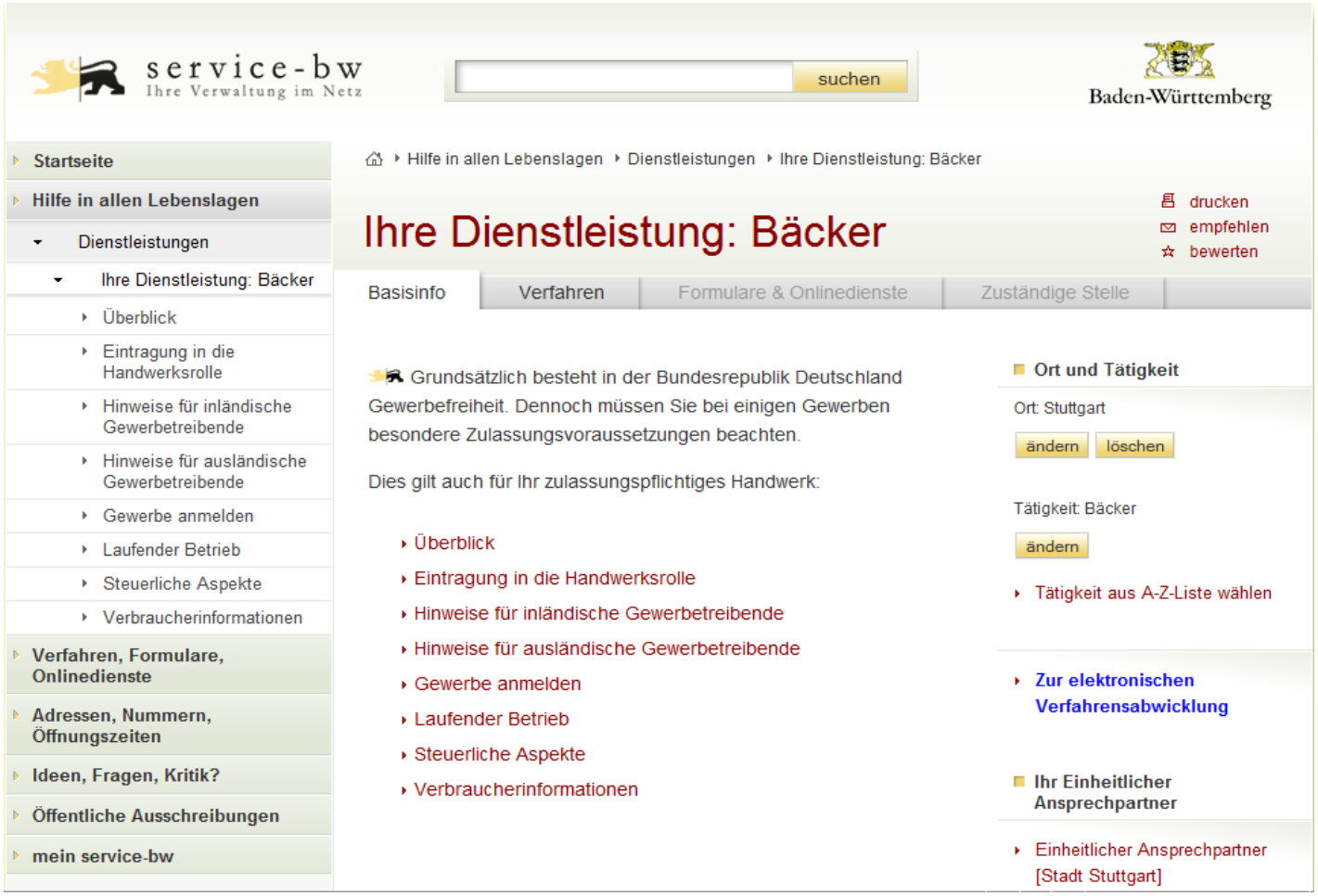
Verwaltungsdiensteportal service-bw

Semantische Techniken – Suche EU-DLR(3).



Verwaltungsdiensteportal service-bw

Semantische Techniken – Suche EU-DLR(4).



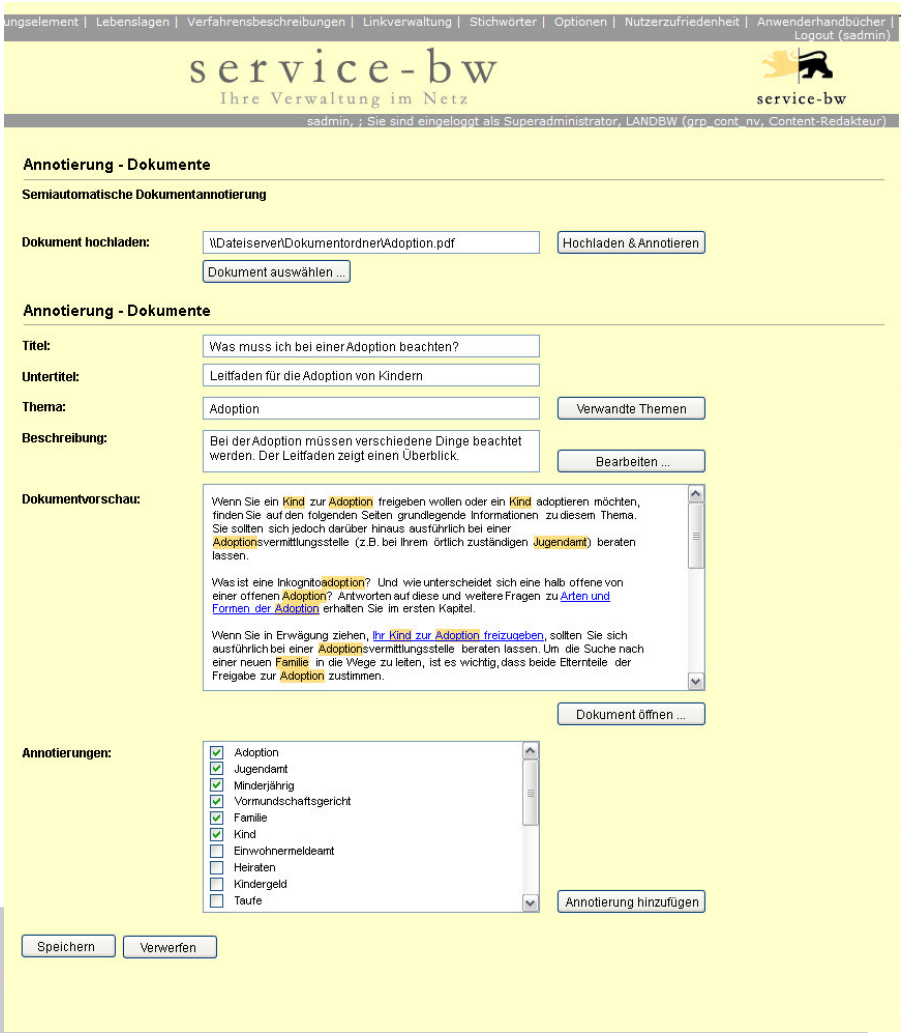
Verwaltungsdiensteportal service-bw

Semantische Techniken – Ausblick AdminCenter

Ontologiepflege AdminCenter

(Halb-)automatische Verschlagwortung

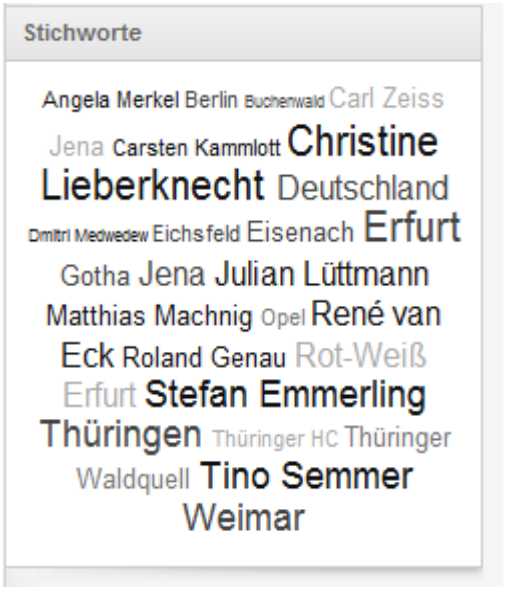
- Aktuell können im AdminCenter den einzelnen Informationsobjekten Verknüpfungselemente zusammen mit den zugeordneten Stichworten und Stichwortsynonymen zugewiesen werden. Deutlich einfacher wäre es, dem Redakteur beim Bearbeiten eines Informationsobjekts gleich eine Vorschlagsliste mit Schlagworten zu unterbreiten, die der Redakteur direkt bearbeiten und übernehmen kann.
- Die Schlagwortliste könnte dann auch unabhängig von der Suche dazu genutzt werden, den Nutzer auf ähnliche Dokumente aufmerksam zu machen.

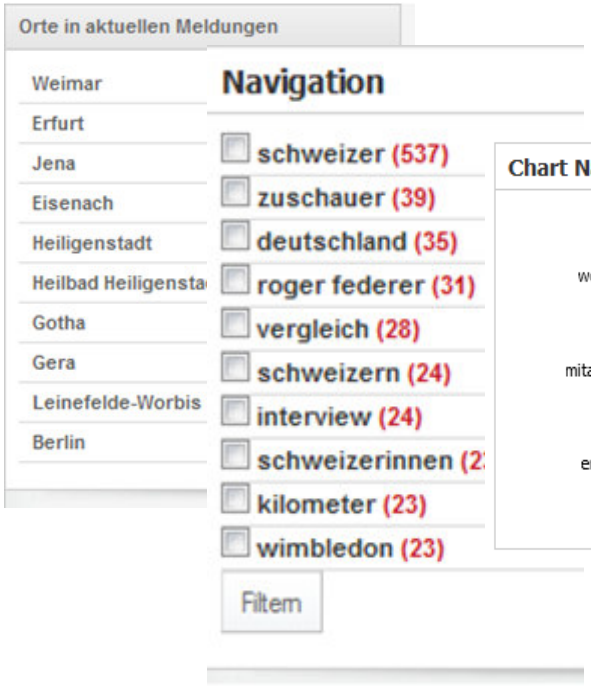


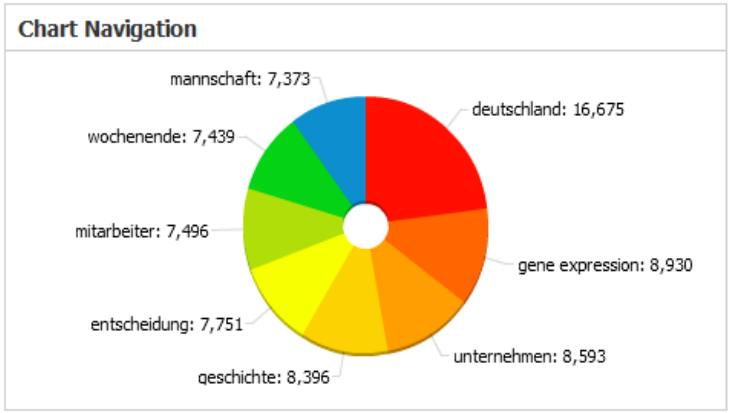
Verwaltungsdiensteportal service-bw

Semantische Techniken – Ausblick (Automatische Filter)

- TagCloud-Filter
  - Stichwörter bilden die TagCloud
  - Schriftgröße richtet sich nach Häufigkeit
- Einfacher Filter
  - Einschränkung der Suchergebnisse durch Filterung
- Chart Navigation-Filter
  - Aktualisierung des Charts nach Filterauswahl
  - animierte Auswahl und Aktualisierung







Verwaltungsdiensteportal service-bw

Semantische Techniken - Fazit der Ontoprise Integration

- Die Qualität der Suchergebnisse hat sich gegenüber der Oracle-Text-Suche stark verbessert.
- Durch kontinuierliche redaktionelle Anpassungen an der Ontologie (insbesondere Stichwortverzeichnis und Stichwortsynoynme) sowie Justierung verschiedener Schwellwerte kann die Qualität weiter verbessert werden.
- Die Synergieeffekte, die durch direkte Rückkopplung zwischen Pflegesystem und Ontoprise möglich sind, sollten genutzt werden.
- Die Suche in service-bw nutzt bisher nur einen Teil der Möglichkeiten, die sich durch den Einsatz von Ontoprise ergeben. Es sollte nach weiteren sinnvollen Erweiterungen der service-bw-Suche recherchiert werden.

Verwaltungsdiensteportal service-bw

Kontakt

Peter Dümig

T-Systems International GmbH

Segment Länder und Kommunen

Fasanenweg 5, 70771 Leinfelden-Echterdingen

+49 711 972 44259 (Tel.)

+49 951 91427206 (PC-Fax)

+49 160 58 25 467 (Mobil)

E-Mail: peter.duemig@t-systems.com

Internet: http://www.t-systems.com