

LUPO

Bereitstellung flexibel nutzbarer Dienste in Landesumweltportalen

*T. Schlachter; W. Geiger; R. Weidemann; G. Zilly
Karlsruher Institut für Technologie, Institut für Angewandte Informatik
Hermann-von-Helmholtz-Platz 1, 76344 Eggenstein-Leopoldshafen*

*R. Ebel; M. Tauber
LUBW Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg
Griesbachstr. 1, 76185 Karlsruhe*

*K. Zetzmann
Ministerium für Umwelt, Klima und Energiewirtschaft Baden-Württemberg
Kernerplatz 9, 70182 Stuttgart*

*M. Trumpler; T. Sattler
DECON-network
Bannwaldallee 24, 76185 Karlsruhe*

*K. Adelhard; M. Möhnle
Bayerisches Staatsministerium für Umwelt und Gesundheit
Rosenkavalierplatz 2, 81925 München*

*A. Müller
Ministerium für Umwelt, Landwirtschaft, Ernährung, Weinbau und Forsten Rheinland-Pfalz
Kaiser-Friedrich-Str. 1, 55116 Mainz*

*U. Keim
K2 & Partner Managementberatung
Wörthstr. 8, 65343 Eltville*

*V. Bachmann; B. Köther
Ministerium für Landwirtschaft und Umwelt des Landes Sachsen-Anhalt
Olvenstedter Str. 4, 39108 Magdeburg*

*D. Keil
Thüringer Ministerium für Landwirtschaft, Naturschutz und Umwelt
Beethovenstr. 3, 99096 Erfurt*

1. EINLEITUNG.....	11
2. DIENSTE IN DEN LANDESUMWELTPORTALEN	11
3. TAG-CLOUDS	12
4. UMWELTNAVIGATOR BAYERN.....	14
5. NUTZUNG DER LUPO-DIENSTE IN KOMMUNEN	15
6. LIVE-SUCHE IN LUFTMESSDATEN	16
6.1 LUFTQUALITÄT ALS SCHULNOTE.....	16
6.2 ABBILDUNGSVORSCHRIFT FÜR DEN GEMEINDEBEZUG	17
6.3 ABLAUF DER LIVE-SUCHE.....	18
6.4 DARSTELLUNG DER ERGEBNISSE	19
7. ZUSAMMENFASSUNG UND AUSBLICK.....	20
8. LITERATUR.....	20

1. Einleitung

Die Landesumweltportale (LUPO) bieten bereits seit Jahren einen einfachen und übersichtlichen Zugang zu Umweltinformationen. Dabei erschließen sie auch viele Datenquellen, die über herkömmliche Internet-Suchmaschinen nicht auffindbar sind. Eine regelmäßige redaktionelle Kontrolle sichert die Qualität der eingebundenen Inhalte, die sich entsprechend des Anbieterkonzeptes hauptsächlich auf Daten von Behörden beschränken.

Eine zentrale Rolle innerhalb der Landesumweltportale spielt deren Volltextsuche. Neben der „klassischen“ Suchfunktion, die Ergebnisse aus dem Index der Suchmaschine in Listenform liefert, wurden in den vergangenen Jahren zunehmend fortgeschrittene Technologien erarbeitet und verwendet, die dem Nutzer zusätzliche Daten, Suchergebnisse und Hinweise bereitstellen, um seine Anfrage möglichst passend zu beantworten /1/. So wurde ein Umweltthesaurus eingebunden, um auch verwandte und umgangssprachliche Begriffe auffinden zu können. „Keymatches“ liefern zu vielen Suchanfragen von der Redaktion ausgewählte, besonders passende Ergebnisse und „OneBoxen“ (s. auch Kapitel 6) können neben den eigenen Suchergebnissen Daten aus verschiedenen externen Systemen ermitteln und direkt darstellen.

2. Dienste in den Landesumweltportalen

Das grundsätzliche Verhalten von OneBoxen entspricht einem REST (Representational State Transfer)-Service, d.h. die Datenquelle bietet eine definierte Adressierbarkeit von Objekten (hier: Umweltdaten) über URLs und ein relativ simples XML-basiertes Antwortformat, das zur maschinellen Weiterverarbeitung (z.B. zum Mashup) der Daten geeignet ist. OneBox-Dienste stehen grundsätzlich auch zur Nutzung in anderen Systemen außerhalb der Landesumweltportale zur Verfügung.

Auch die Landesumweltportale sowie ihre Suchmaschine stellen zunehmend Schnittstellen zur Verfügung, die von anderen Systemen genutzt werden können. Die OpenSearch-Schnittstelle der Volltextsuche wurde bereits ausführlich in /2/ beschrieben. Sie wird seit Ende 2009 erfolgreich im Umweltportal Deutschland (PortalU) und seit dem Frühjahr 2011 auch im staatlich-kommunalen Verwaltungsdienstportal Baden-Württemberg (service-bw) zur Anzeige von Suchergebnissen aus dem Umweltportal Baden-Württemberg (Portal Umwelt-BW, <http://www.umwelt.baden-wuerttemberg.de/>) genutzt. Umgekehrt löst eine Suche im Portal Umwelt-BW auch eine Suche bei service-bw aus und zeigt im Erfolgsfall die passenden Treffer in einer Box neben den eigenen Suchergebnissen an (Abbildung 1).

In diesem Sinne entwickelt sich die Architektur der Landesumweltportale zunehmend in Richtung eines serviceorientierten Ansatzes. Dem tragen auch die Entwicklungen im Projekt „Semantische Suche nach Umweltinformationen (SUI)“ Rechnung /3/. Hier soll ein „Search Broker“ die Moderation zwischen den Portalsystemen und den Zielsystemen (Suchmaschine, Datenquellen, Dienste) übernehmen und auf Basis der vom Nutzer gegebenen Suchanfrage die angeschlossenen Systeme gezielt und passend abfragen. Das Portal übernimmt anschließend die Darstellung der ggf. vielfältigen Ergebnisse.



Abbildung 1: Suchergebnisse aus service-bw als OneBox im Portal Umwelt-BW

3. Tag-Clouds

Tag-Clouds dienen im Allgemeinen der Darstellung einer Liste von Schlagworten, bei der über eine alphabetische Sortierung und die Darstellungsgröße gleichzeitig zwei Sortierkriterien zur Anwendung kommen können. Viele Internet-Magazine und Blogs verwenden Tag-Clouds als einen möglichen Einstieg für Nutzer, indem sie Schlagworte der am häufigsten nachgefragten Artikel in einer als Schlagwortwolke (Abbildung 2) dargestellten Form anbieten, bei der die Schriftgröße einzelner Schlagworte relativ zu deren Häufigkeit gewählt wird. Häufig nachgefragte Begriffe werden dabei größer dargestellt als andere. Ein Klick auf eines der Schlagworte löst in der Regel eine entsprechende Suchanfrage aus. Die Ergebnisliste stellt die zum Schlagwort gehörigen Artikel dar, ggf. wiederum sortiert nach einem bestimmten Kriterium wie der Relevanz des Schlagwortes für einen Artikel oder der Aktualität.



Abbildung 2: Kleine Tag-Cloud im Portal Umwelt-BW

Auch für die LUPO-Portale bestand der Wunsch nach Tag-Clouds dieser Art. Einer der Auslöser hierfür war die Verfügbarkeit von Suchprotokollen, in denen die Häufigkeit von Suchan-

fragen ausgewertet werden kann. Die Idee dahinter ist recht simpel: Themen, die zuletzt für eine gewisse Zahl von Nutzern interessant waren, könnten auch für einen großen Teil der übrigen Nutzer relevant sein. Ähnliche Ideen gehen sogar noch weiter: So werden durch Internet-Suchmaschinenbetreiber Anfragen zu Krankheitssymptomen regional ausgewertet um z.B. die Ausbreitung von Grippe-Epidemien erkennen zu können /4, 5/.

Bei einer Analyse der Suchbegriffe in den Umweltportalen ergab sich jedoch, dass viele Suchbegriffe, insbesondere durch den Anschluss der LUPO-Portale an das PortalU und damit die Suchmaschine MetaGer, für eine automatische Übernahme in eine solche Tag-Cloud nicht geeignet sind.

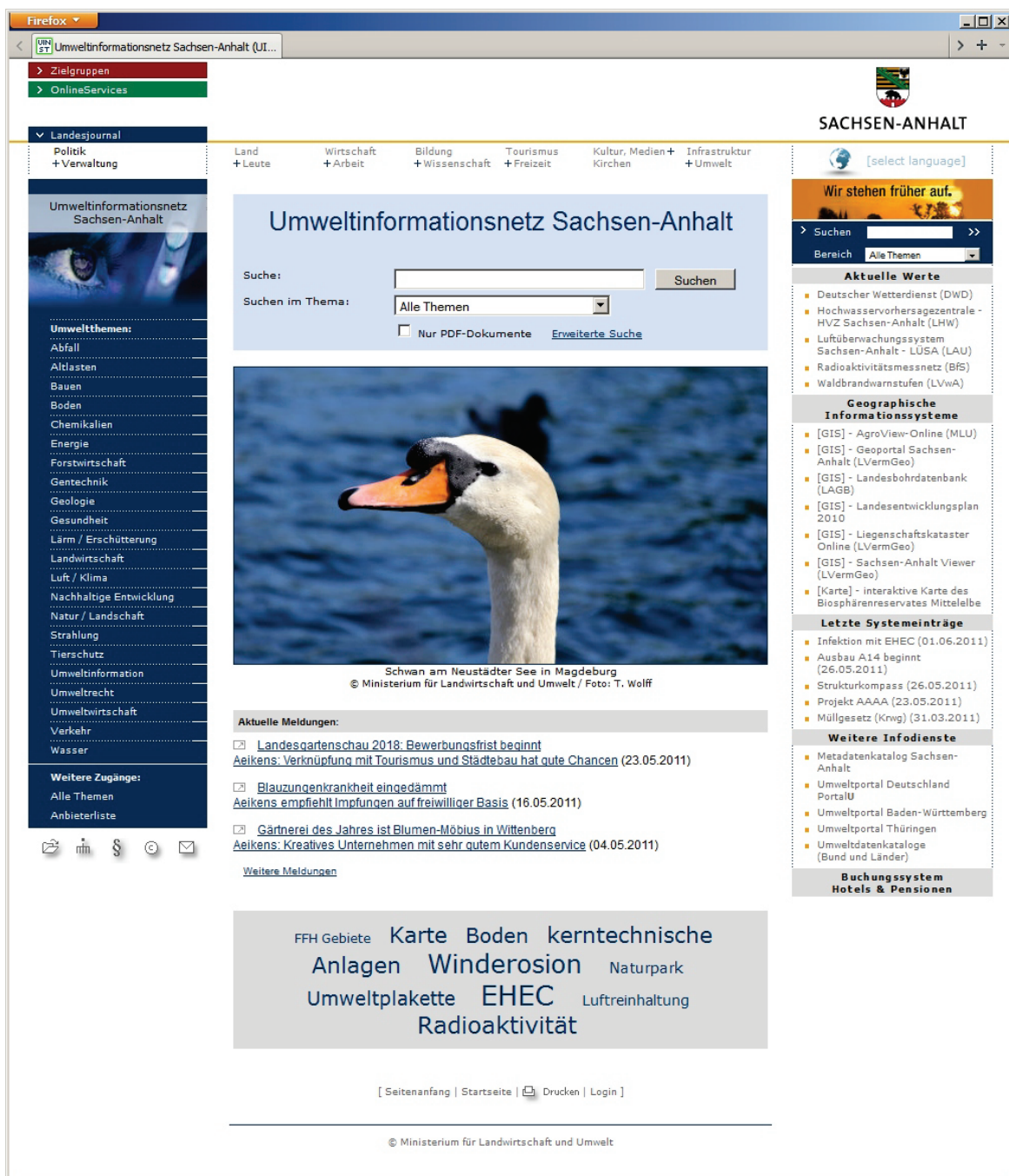


Abbildung 3: Tag-Cloud auf der Startseite des Umweltinformationsnetzes Sachsen-Anhalt. Die hellblaue Box stellt den überarbeiteten Einstieg in die Volltextsuche dar.

Deshalb wurde für die den LUPO-Portalen zugrunde liegende CMS-Plattform WebGenesis eine Tag-Cloud-Komponente entwickelt, welche die redaktionelle Pflege einer solchen Tag-Cloud ermöglicht. Die Suchbegriffe werden dabei über eine Schlagwortliste verwaltet. Dabei kann jedem Schlagwort als Zusatzinformation der zu verwendende Suchraum (Gesamt- oder Teilindex) der Suchmaschine zugeordnet werden. Die Reihenfolge der Schlagworte beschreibt dabei implizit ihre Relevanz.

Zur Präsentation der Schlagwortwolke stehen eine Reihe von Funktionen zur Verfügung, darunter die (zufällige) Auswahl einer bestimmten Anzahl von Schlagworten, die zufällige oder alphabetische Sortierung der Begriffe und die zufällige Auswahl von Schriftarten für einzelne Schlagworte. Diese Funktionen können innerhalb von WebGenesis zur Steuerung eines EJava-Templates verwendet werden, sodass die konkrete Darstellung einer Tag-Cloud in jedem LUPO-Portal individuell durch die Redakteure gestaltet werden kann. Die Tag-Clouds werden in den LUPO-Portalen von Baden-Württemberg, Thüringen und Sachsen-Anhalt (Abbildung 3) seit November 2010 produktiv genutzt. Die Tag-Cloud-Komponente für WebGenesis kann darüber hinaus auch unabhängig von den LUPO-Portalen eingesetzt werden.

4. Umweltnavigator Bayern

Am 24.11.2010 erfolgte der Beitritt des Freistaates Bayern in die LUPO-Kooperation. Im Rahmen eines Workshops wurden einerseits die Architektur und Konzepte der LUPO-Landesumweltportale präsentiert, andererseits auch Ideen und Anforderungen aus Bayern für ein bayerisches Landesumweltportal gesammelt.

Mit „Umweltnavigator Bayern“ steht der Name des Portals bereits fest. Ein erster Prototyp wurde entwickelt, der Suchindex für Bayern enthält bereits mehrere 100.000 Dokumente. Da ressortübergreifend gesucht werden soll, wird der Umweltnavigator als Angebot der bayerischen Staatsregierung positioniert. Hohe Priorität haben dabei auch die Einbindung von Videos und die Lenkung der Besucher zu den wichtigsten Web-Auftritten des Landes. Die Startseite wird die Umweltthemen dabei als innovative Auswahlliste in Art der Windrose eines Kompasses präsentieren. Auf den einzelnen Themenseiten sollen neben den wichtigsten Informationssystemen auch dynamisch die relevantesten Suchergebnisse angezeigt werden.



Abbildung 4: Erster Prototyp des Umweltnavigators Bayern

5. Nutzung der LUPO-Dienste in Kommunen

In Rheinland-Pfalz werden die LUPO-Technologien insbesondere für die Volltextsuche in der Wasserwirtschaft eingesetzt. Die Suchfunktionalität lässt sich dabei im Geoportale Wasser (<http://www.geoportale-wasser.rlp.de/>), mittlerweile aber auch in anderen Webangeboten, nutzen. Da im Suchindex auch wasserbezogene Inhalte vieler kommunaler Anbieter erfasst werden, soll die Suchfunktionalität künftig auch für Kommunen nutzbar gemacht, d.h. in den Web-Auftritten der Kommunen eingebunden werden können. Den Kommunen steht dafür der „DataScout“ zur Einbindung und Erstellung eigener Seiten zur Verfügung. Der DataScout enthält einen speziellen Template-Editor, mit dem die Anpassung der Recherche-Oberfläche an kommunale Anforderungen (insbesondere was das Layout betrifft) möglich ist. Ein Template für die Volltextsuche ist derzeit in Entwicklung. In Kombination mit einer entsprechenden Sammlung (Teilindex) innerhalb der Suchmaschine kann jede Gemeinde die für sie spezifischen Daten in einer scheinbar „eigenen“ Volltextsuche anbieten. Als Pilot für eine solche Einbindung dient der Bereich „Umwelt“ im Webangebot der Verbandsgemeinde Hagenbach

(<http://www.datascout.rlp.de/txppublic/39705/silver/index.html>). Der DataScout ist ein innovatives Beispiel dafür, wie LUPO-Dienste für Dritte nutzbar gemacht werden können.

6. Live-Suche in Luftmessdaten

Das Umweltportal Baden-Württemberg verwendet für die Suchfunktion eine Google Search Appliance (GSA) als Backend. Im Zuge der Weiterentwicklung des Portals wurde insbesondere die Suchfunktionalität optimiert und erweitert. Die integrierten OneBox-Module liefern zusätzlich zu den Ergebnissen aus der Volltextsuche weitere Ergebnisse, die im Umweltportal in der rechten Spalte in separaten Boxen angezeigt werden.

Eines dieser OneBox-Module integriert beispielsweise Inhalte des interaktiven Dienstes UDO (Umwelt-Daten und -Karten Online) der LUBW Landesanstalt für Umwelt, Messungen und Naturschutz Baden-Württemberg, die aufgrund des Einsatzes von JavaScript und der interaktiven Selektion mittels Formularen nicht von der Suchmaschine selbst gecrawlt werden können.

Nun sollten auch Inhalte aus dem Themenbereich Luft mit der OneBox-Technologie in die Suchergebnisse eingebunden werden. Im Vergleich zum UDO-Modul, welches als Ergebnis Links zu UDO-Schnellabfragen (parametrisierbare Workflows) liefert, sollten hier aktuelle Messdaten direkt in der OneBox dargestellt werden.

6.1 Luftqualität als Schulnote

Das Ziel war es, dem Besucher aktuelle und aussagekräftige Informationen aus dem Fachbereich Luft zu präsentieren. Bei der Suche nach geeigneten Messdaten fiel die Wahl schließlich auf den Kurzzeit-**Luftqualitätsindex** LuQx.

Im Themenportal Luft des LUBW-Internetauftritts bietet die Messnetzzentrale (MNZ) eine Interpretation der Luftqualität „als Schulnote“. Mit dem LuQx wird eine auch für den Laien interpretierbare Darstellung der aktuellen Luftmesswerte angeboten. Hierfür werden die Messwerte der fünf für die Luftbelastung maßgeblich relevanten Komponenten herangezogen. Diese sind:

- Stickstoffdioxid NO₂
- Feinstaub PM10
- Ozon O₃
- Schwefeldioxid SO₂
- Kohlenmonoxid CO

Die Messdaten werden hierbei in Indexklassen von 1 („sehr gut“) bis 6 („sehr schlecht“) eingeordnet, welche auf dem Schulnotenprinzip beruhen. Die schlechteste Note der Einzelkomponenten bestimmt schließlich den LuQx-Wert für die gesamte Messstelle. Dargeboten werden die aktuellen Werte auf den Webseiten der MNZ in Form einer Übersichtskarte. Die Darstellung der LuQx-Noten erfolgt dort in Form von farbigen Kreissymbolen, wobei durch die Farbgebung (von blau über gelb bis hin zu rot) die jeweilige Luftqualität zusätzlich zum Notenwert gekennzeichnet wird (Abbildung 5).

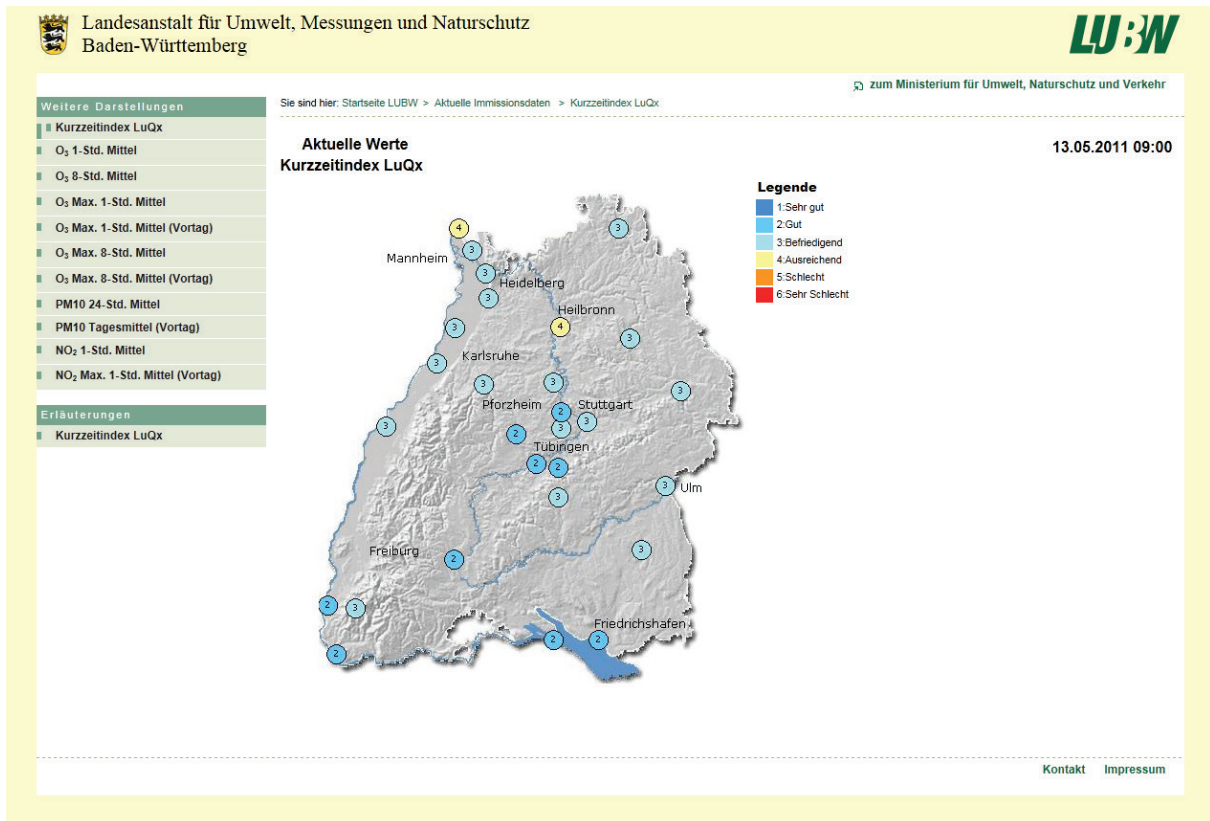


Abbildung 5: Übersichtskarte Kurzzeitindex LuQx im Themenportal Luft der LUBW

6.2 Abbildungsvorschrift für den Gemeindebezug

Die LuQx-OneBox will dem Benutzer vor allem bei der ortsbezogenen Suche passende Ergebnisse präsentieren. Der Anwender sucht nach lokalen Umweltinformationen, indem er einen Gemeinamen im Suchterm verwendet. Für diesen Zweck benötigt die OneBox eine Methode zur Abbildung von Gemeinden auf Luftmessstationen. Eine solche Abbildungsvorschrift wurde als Basis für die bisherige OneBox Luftmessstationen bereits mit der Fachseite entwickelt. Laut dieser Vorschrift wird die Gemeindegrenze um eine Pufferzone von 20 km erweitert (Abbildung 6). Die Luftmessstationen, die sich innerhalb dieser vergrößerten Geometrie befinden, werden für die Gemeinde als relevant betrachtet. Implementiert wurde diese Abbildung als Relationentabelle im Datenbankschema SMI, welches bereits die Daten für die Integration des UDO-Dienstes aufbereitet. Diese Tabelle wird asynchron, nämlich im Zuge der monatlichen Aktualisierung der Messstationsdaten, mit Hilfe einer rechenintensiven Geo-SQL-Prozedur gefüllt. Auf Basis dieser Zuweisungstabelle kann die OneBox schließlich zur Laufzeit eine performante Auswertung durchführen.

Ein grundsätzliches Problem stellte die Datenquelle für die aktuellen Werte des Luftqualitätsindex LuQx dar. Momentan ist die Datenbank, welche die Grundlage für die Webseiten mit den tagesaktuellen LuQx-Werten bildet, für die OneBox nicht erreichbar. Bis die technischen Voraussetzungen für die Anbindung der originalen Datenquelle erfüllt sind, werden die aktuellen LuQx-Werte aus der Webseite der Messnetzzentrale der LUBW ausgelesen. Sämtliche benötigten Informationen liegen als XML-Struktur in einer Imagemap vor, die für die Darstellung der Übersichtskarte „Aktuelle Werte Kurzzeitindex LuQx“ im Themenportal Luft des LUBW-Internetauftritts verwendet wird.

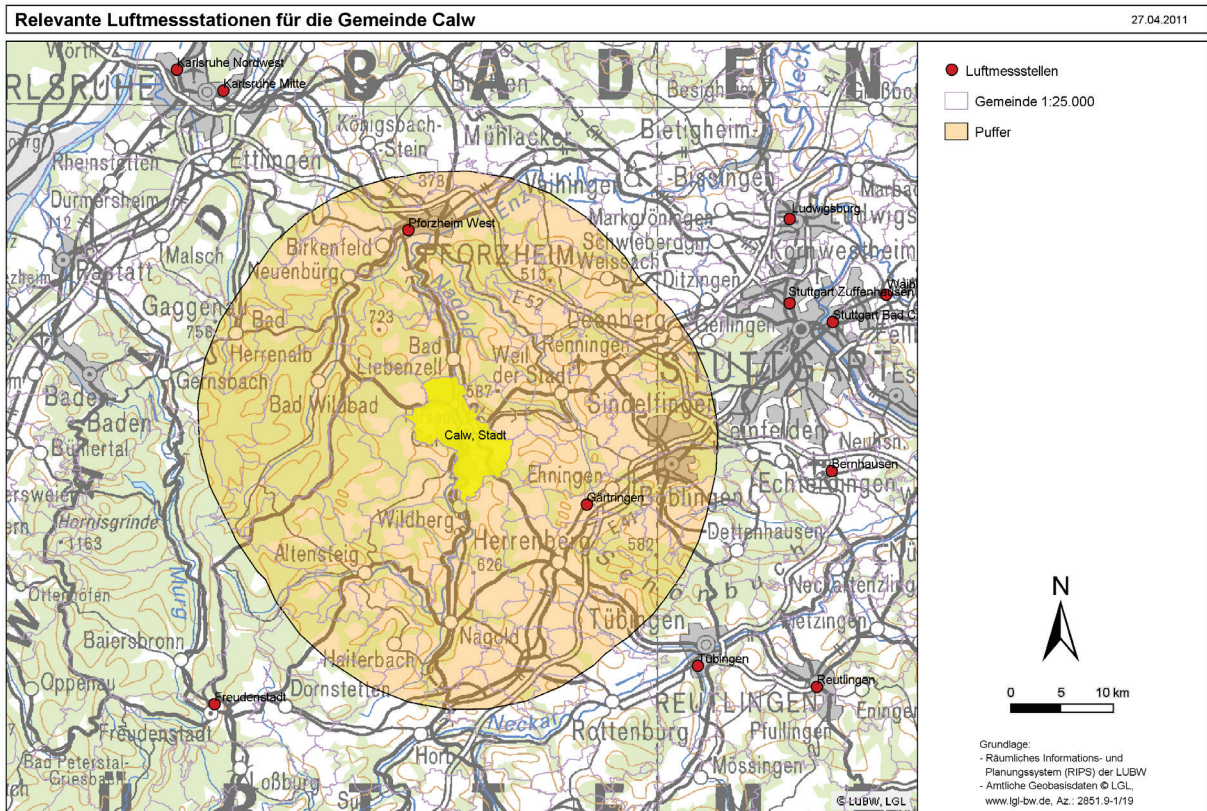


Abbildung 6: Ermittlung der für die Gemeinde Calw relevanten Luftmessstationen

6.3 Ablauf der Live-Suche

Sowohl die UDO-OneBox als auch der Vorgänger des LuQx-Moduls, die Luftmessstationen-OneBox, basieren auf der Volltextsuche der GSA. Die Grundlage bildet hierbei ein sogenannter Content-Feed, welcher Datenbankinhalte in die Suchmaschine lädt. Aus den aufbereiteten Daten wird ein internes Webdokument generiert und indexiert. Somit können dieselben Techniken zur Aufbereitung der Inhalte verwendet werden wie bei der redaktionellen Pflege der Umweltportalinhalte, nämlich die Integration sämtlicher suchrelevanter Informationen in das interne Webdokument. Dadurch gelang es beispielsweise auch, eine Suchmöglichkeit nach Teilgemeinden und Gemeindesynonymen generisch zu realisieren. Der Einsatz dieser Methodik war vor allem aufgrund des monatlichen Aktualisierungszyklus der relevanten Inhalte möglich.

Mit der LuQx-OneBox wollte man jedoch tagesaktuelle Ergebnisse direkt ausgeben. Hierzu war der GSA-Index nicht geeignet, vielmehr musste eine eigene Suchmethodik implementiert werden. Die Entwicklung einer Entsprechung der GSA-Volltextsuche mittels Generierung eines virtuellen Dokumentes für jede Messstation wäre zu aufwändig und ungeeignet für den Einsatz als Live-Suche. Bei dem für die LuQx-OneBox implementierten Suchalgorithmus wurde schließlich eine andere Strategie zugrunde gelegt. Anstatt sämtliche relevanten Informationen bei der Suche zu berücksichtigen, bietet die OneBox eine Ergänzung zur Volltextsuche, die nicht den Anspruch erhebt, hundertprozentig der Suchphrase zu entsprechen.

Der Suchalgorithmus der LuQx-OneBox besteht aus mehreren Phasen (Abbildung 7). Zunächst werden aus der Suchphrase die enthaltenen Gemeinde- und Stationsnamen extra-

hiert. Zu diesen Begriffen werden die passenden Messstationen aus einem Datenbankschema ausgelesen. Anschließend werden die aktuellen Werte für den Kurzzeitindex LuQx aus der in der Webseite hinterlegten Imagemap ausgelesen. Als Ergebnis liefert die OneBox schließlich die relevanten Luftmessstellen inklusive der aktuellen LuQx-Noten.

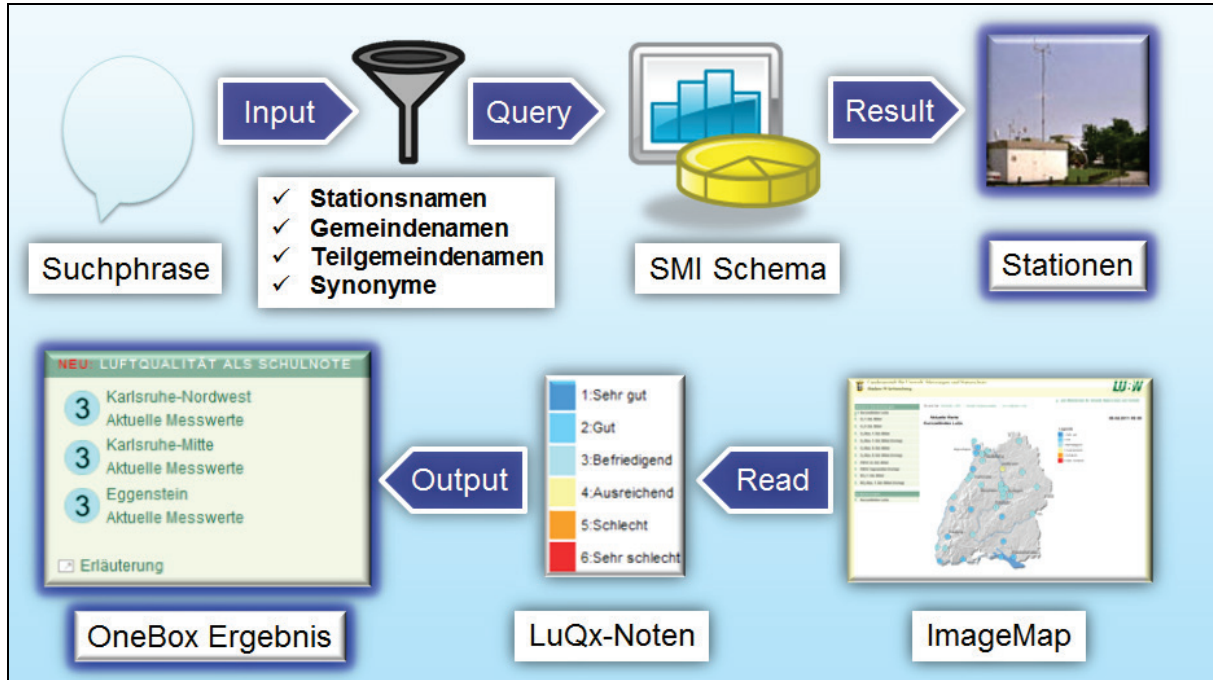


Abbildung 7: Suchvorgang der LuQx-OneBox

6.4 Darstellung der Ergebnisse

Die OneBox präsentiert den jeweiligen LuQx-Wert als farbiges Kreissymbol, wobei die Farbgebung der Präsentation auf der Webseite des Luftportals entspricht (Abbildung 8). Der jeweilige Stationsname verweist auf die Webseite mit Detailinformationen zu den LuQx-Werten. Darunter wird ein zusätzlicher Verweis „Aktuelle Messwerte“ auf die Webseite mit den für die LuQx-Klassifikation zugrundeliegenden Messwerten ausgegeben. Dadurch gelang eine Integration des bisherigen Luftmessstationen-Moduls in die LuQx-OneBox.



Abbildung 8: Die OneBox „Luftqualität als Schulnote“ für eine Suche nach „Stuttgart“

7. Zusammenfassung und Ausblick

An vielen in den vorigen Abschnitten dargestellten Beispielen wird die Entwicklung der Landesumweltportale hin zu einem diensteorientierten System deutlich. Dabei wurden und werden dort, wo es sinnvoll und wirtschaftlich ist, möglichst universell nutzbare Schnittstellen geschaffen, die die Einbindung von Umweltinformationen in andere Systeme und Kontexte ermöglichen. Dass sich diese Nutzung künftig nicht nur auf den Web-Bereich beschränken wird, zeigt das Projekt „LUPO mobil“, das die Nutzung von innerhalb und außerhalb des LUPO-Rahmens entwickelten Diensten auf mobilen Endgeräten ermöglichen soll /6/. Dabei wird insbesondere den technischen Möglichkeiten dieser Geräte (z.B. Standortbestimmung zur Ermittlung des Kontexts einer Anfrage) Rechnung getragen. Auch die im Projekt SUI gemachten Entwicklungen sollen in die Landesumweltportale einfließen. Neben einer erwarteten Verbesserung der Suchfunktionalität stehen hier die Verbesserung der Ergebnisdarstellung (Mashup) und die Möglichkeit zur Darstellung von Geoinformationen im Vordergrund. Außer Geoinformationen sollen aber auch weitere Medien eingebunden werden können, z.B. zunehmend Multimedia-Daten in Form von Bildern und Videos.

8. Literatur

- /1/ Schlachter, T. et al. (2009): LUPO – Fortgeschrittene Suchfunktionen in den Landesumweltportalen von Baden-Württemberg, Sachsen-Anhalt und Thüringen In: Mayer-Föll, R., Keitel, A., Geiger, W.; Hrsg.: Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt, Verkehr und benachbarte Bereiche in neuen Verwaltungsstrukturen, Phase IV 2008/09, Forschungszentrum Karlsruhe, Wissenschaftliche Berichte, FZKA 7500, S. 149-156.
- /2/ Schlachter, T. et al. (2010): LUPO – Ausbau der Suchfunktionalität der Landesumweltportale und Vernetzung mit PortalU. In: Mayer-Föll, R., Ebel, R., Geiger, W.; Hrsg.: Kooperative Entwicklung wirtschaftlicher Anwendungen für Umwelt, Verkehr und benachbarte Bereiche in neuen Verwaltungsstrukturen, Phase V 2009/10, KIT Scientific Reports 7544, S. 9-20.
- /3/ Bügel, U. et al. (2011): SUI für Umweltportale – Entwurf und prototypische Implementierung einer Architektur für die semantische Suche im Portal Umwelt-BW. In diesem Bericht.
- /4/ NZZ Blogs vom 2.12.2008: Epidemie-Warnungen von der Suchmaschine. http://www.nzz.ch/nachrichten/blogs/nzz_blogs/betablog/epidemie-warnungen_von_der_suchmaschine_1.1342919.html; besucht am 28.04.2011.
- /5/ Grippe-Trends weltweit verfolgen. <http://www.google.org/flutrends/>; besucht am 28.04.2011.
- /6/ Schlachter, T. et al. (2011): LUPO mobil – Ein Schichtenmodell zur Auswahl und Nutzung von Umweltdiensten auf mobilen Endgeräten. In diesem Bericht.