

KoGIs - Kompetenzzentrum für die Gestaltung der Informationssysteme  
Bremen.online GmbH

## Umsetzung der bremenweiten Suche (bremen.de und KoGIs-Instanzen)

Präambel

Die umfassend vorhandenen Online-Informationen der Bremischen Verwaltung sind auf zahlreiche Internetauftritte verteilt, vor allem in [www.bremen.de](http://www.bremen.de) sowie den Auftritten der Ressorts und Dienststellen der bremischen Verwaltung und der bremischen Gesellschaften. Durch die zunehmende Umstellung der Auftritte auf SixCMS unter Anwendung der Mustertemplates für die Bremische Verwaltung bietet sich die Möglichkeit, durch den Einsatz von standardisierten Diensten eine globale Suche über alle Inhalte anzubieten.

**Inhalt**

---

<i>Einleitung</i> .....	3
Ausgangssituation .....	3
Ziel .....	3
bremen.de .....	3
KoGIs .....	3
<i>Allgemeine Anforderungen</i> .....	4
Sucheingabe .....	4
Performance/System .....	4
Zentrale Anforderungen: .....	4
Anforderungen an Antwortzeit: .....	5
Relevanz/ Gewichtung: .....	5
Anforderungen an die Ergebnisdarstellung .....	5
<i>Systemspezifische Anforderungen</i> .....	6
KoGIs .....	6
<b>Vorschlag für die Umsetzung der Suche in grafischer Form:</b> .....	6
bremen.de .....	8

## Einleitung

---

Die umfassend vorhandenen Online-Informationen der Bremischen Verwaltung sind auf zahlreiche Internetauftritte verteilt, vor allem in [www.bremen.de](http://www.bremen.de) sowie den Auftritten der Ressorts und Dienststellen der bremischen Verwaltung und der bremischen Gesellschaften. Durch diese dezentrale Verteilung der Informationen wächst der Bedarf nach einer zentralen Suchmöglichkeit für alle Inhalte.

Dieses Papier stellt die Anforderungen der unterschiedlichen Systeme ([bremen.de](http://www.bremen.de) und KoGIs) zusammen, um darauf aufbauend verschiedene Umsetzungsmöglichkeiten zu prüfen und ein Konzept zu erstellen.

## Ausgangssituation

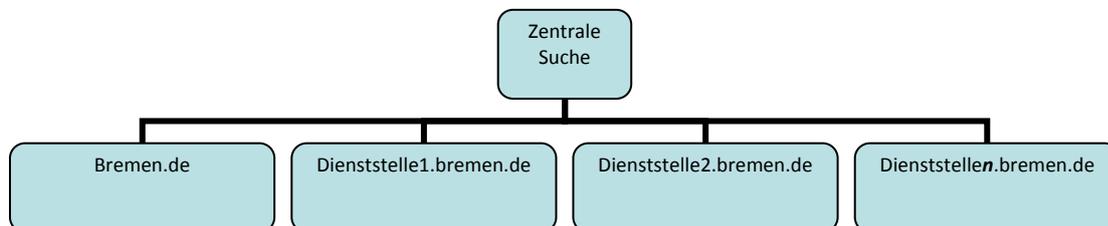
---

Die Internetauftritte der bremischen Verwaltung sind mit Hilfe der auf SixCMS basierenden Basismodule umgesetzt und –mit 2 Ausnahmen- zentral auf einem Server verortet. Alle Auftritte sind geschlossene Systeme (Instanzen) mit eigenen Datenbanken (aktuell ca. 90 Instanzen). Das Stadtportal [bremen.de](http://www.bremen.de) läuft ebenfalls auf SixCMS (nicht mit Hilfe der Basismodule), aber als geschlossenes System in einer gesonderten Serverfarm.

## Ziel

---

Die Inhalte der KoGIs-Instanzen und [bremen.de](http://www.bremen.de) sollen über eine Suche erschließ- und auffindbar sein.



## [bremen.de](http://www.bremen.de)

---

[Bremen.de](http://www.bremen.de) soll dann über einen zentralen Suchindex (Inhalte aus [bremen.de](http://www.bremen.de) und den KoGIs-Instanzen) verfügen, in dem jede gestartete Suche läuft. Somit gibt es auf [bremen.de](http://www.bremen.de) nur eine Suchmöglichkeit (,rotes Suchfeld'), die Quellen der Ergebnisse werden nur in der Ergebnisdarstellung deutlich gemacht.

## KoGIs

---

Die lokale Suche auf den KoGIs-Instanzen wird weiterhin erhalten bleiben. Die Instanzen sollen also über eine zweistufige Suche verfügen (lokal + ,bremenweit'). Grundsätzlich wird davon ausgegangen, dass diese immer zeitgleich ausgeführt wird (d.h. bei einer Sucheingabe wird sowohl lokal als auch ,bremenweit' gesucht und die Ergebnisse zweistufig dargestellt). Nur bei zu erwartenden Performanceproblemen sollte darüber nachgedacht werden, zunächst nur lokale Suche über das rote Suchfeld auszulösen und bei der Ergebnisdarstellung einen Hinweis auf die ,bremenweite Suche' einzubringen, die dann manuell ausgelöst werden kann..

## Allgemeine Anforderungen

---

Die nachfolgenden Anforderungen gelten grundsätzlich für die ‚bremenweite Suche‘ sowohl für die Inhalte der KoGIs-Instanzen wie auch für bremen.de.

## Sucheingabe

---

- Es soll keine weitere Suchmaske für die Sucheingabe geben, die Suche erfolgt über die lokalen Suchfelder.
- Die Suchmaschine soll nicht nach Boolescher Logik erfolgen, sondern nach Termhäufigkeit und –gewichtung die relevantesten Treffer am Anfang platzieren. Dazu gehört auch, dass automatisch die Treffer bei Mehrworteingaben oben stehen, die ALLE Suchbegriffe enthalten (entspr. UND- oder NEAR-Operator) und danach die Treffer erscheinen in denen nur eines der Suchwörter vorkommt (entspr. ODER-Operator).
- Es soll keine Suche nur nach ganzen Wörtern geben, sondern immer automatisch Teilwort-Suche.
- Suche nach Phrasen (Wortfolgen) soll möglich sein (in Anführungszeichen) bzw. in der Suchergebnisliste automatisch bevorzugt angezeigt werden, auch wenn nicht explizit nach Phrasen gesucht wurde (vgl. ODER / UND).
- Ein Suchbegriff soll mindestens 2 Zeichen lang sein.
- Die URL soll nicht manipulierbar sein, d.h. direkte Eingabe der Searchvalues unterbinden. Die URL soll zum Bookmarks einer Suche aber geeignet sein.
- Auch nationale Sonderzeichen (Buchstaben mit sog. diakritischen Zeichen) werden für die Suche in entsprechende „Basis“-Buchstaben umgewandelt (z.B. â -> a).
- Das Euro-Zeichen wird auch mit der Suche „EUR“ oder „EURO“ gefunden und umgekehrt.
- Sinnvoll wäre eine (auf jeden Fall für bremen.de, wenn nicht generell interessant) Rechtschreibfehler-toleranz, am besten wäre ein Feature wie 'Meinten Sie?'

## Performance/System

---

- Die Suchmaschine ist ein eigener Server.
- Die Suche wird in Echtzeit abgefragt.
- Die Kommunikation zwischen den lokalen Instanzen und der globalen Instanz sollte über http oder darauf aufbauende Protokolle erfolgen (ggf. überflüssig?).
- Es muss eine möglichst zeitnahe Indizierung der Artikel erfolgen, wobei keine Offline-Artikel indiziert werden dürfen. Die Löschung und Änderung der Artikel muss ebenso zeitnah erfolgen.
- Für die Einbindung externer Quellen ist eine Schnittstelle zu entwickeln, die auch zu dokumentieren ist.

## Zentrale Anforderungen:

---

- Es sollen alle Dokument-Typen durchsucht und die gesamten Inhalte indiziert werden.
- Es sollte eine Stoppwortliste benutzt werden, um „inhaltslose“ Wörter wie „oder“, „weil“, „aber“ usw. von der Suche auszuschließen.
- Es soll keine erweiterte Suche geben.
- Es sollen auch Inhalte indiziert werden, die nicht dargestellt werden (bspw. Metadaten des IFG).

### Anforderungen an Antwortzeit:

---

- Die Suche erfolgt über die Artikel, nicht über Seiten.
- Ausgangswert: ca. 5000 Artikel (Seiten, Personen etc.) bei ca. 90 Instanzen aus der Verwaltung.
- Ausgangswert bremen.de: ca. 50.000 Artikel (inkl. Visitenkarten, Straßenregister, Veranstaltungen).
- Antwortzeit : Suchergebnis < 1s .

### Relevanz/ Gewichtung:

---

- Die Relevanz eines Artikels ist das Kriterium für die Sortierung bei der Ausgabe der Treffer. Zur Ermittlung der Relevanz werden die unten aufgeführten Kriterien bewertet. Da uns einige Kriterien wichtiger als andere sind, werden sie gewichtet. Die Gewichtung erhöht die Relevanz, wodurch der Artikel in der Trefferliste weiter oben angezeigt wird. Der Faktor zur Gewichtung der Kriterien muss später angepasst werden können.
  - Titel
  - Schlüssel- bzw. Schlagworte
  - Überschrift 1
  - Überschrift 2
  - Überschrift 3
  - Überschrift 4
  - Überschrift 5
  - Überschrift 6
  - Text (auch Links, Infoboxen, etc. in Verbindung mit den Seiten)
- Bevorzugt werden bei Mehrworteingaben die Treffer in folgender Rangfolge ausgegeben:
  - Phrasen
  - UND-Verknüpfung
  - ODER-Verknüpfung

### Anforderungen an die Ergebnisdarstellung

---

- Das Online-/Change-Datum eines Artikels wird in der Liste der Ergebnisse nicht angezeigt.
- Die unterschiedlichen Artikeltypen (News, Personendaten, Veranstaltungen, Links, Downloads) sollten gekennzeichnet werden, ebenso die Dokumente (.doc, .pdf etc).
- Kennzeichnung von Ergebnissen der ‚bremenweiten Suche‘
  - Die Quelle (Dienststelle) des Suchergebnisses muss **vor** dem Titel des Suchergebnisses aufgeführt werden, jedoch etwas kleiner (analog zu Google) und wegen der Länge eine Zeile drüber

[PDF] [Ausbildung & Beruf](#)

Dateiformat: PDF/Adobe Acrobat

Impressum. Herausgeber. Bundesministerium. für Bildung und Forschung (BMBF). Referat

Öffentlichkeitsarbeit. 11055 Berlin. Bestellungen ...

[www.bmbf.de/pub/ausbildung\\_und\\_beruf.pdf](http://www.bmbf.de/pub/ausbildung_und_beruf.pdf) - [Ähnliche Seiten](#)

- Kennzeichnung externer Links muss durch das entsprechende Symbol und einen Tooltip erfolgen (analog zu der bisherigen inhaltlichen Verwendung von externen Links).

## Systemspezifische Anforderungen

---

### KoGIs

---

- Die Funktionalitäten der lokalen Suche müssen übernommen werden.
- Nach der Suchabfrage werden zunächst nur die lokalen Ergebnisse aufgeführt.
- Geschützte Seiten und Dokumente sind grundsätzlich von der Suche ausgenommen.
- Ein Link-Text am Seitenende der lokalen Suche muss eindeutig kommunizieren, dass es Ergebnisse auf der Ebene der ‚bremenweiten Suche‘ gibt: Anzeige eines Links am Ende der Trefferseiten „Auf anderen offiziellen Internet-Seiten der Freien Hansestadt Bremen wurden XXX weitere Treffer gefunden. Ergebnisse anzeigen“.
  - Wenn man bei der lokalen Suche auf den Link für weitere Suchergebnisse in der ‚bremenweiten Suche‘ geklickt hat,
    - a) werden die Ergebnisse der ‚bremenweiten Suche‘ in der gleichen lokalen Instanz angezeigt, aus der heraus der Link zur ‚bremenweite Suche‘ angeklickt worden ist
    - b) wird im Titel (title), der Kapitelüberschrift und im Text vor den Suchergebnissen deutlich auf die Erweiterung des Suchindexes (‚bremenweiten Suche‘) hingewiesen.
  - Der Nutzungspfad – wie er bei der lokalen Suche vorhanden ist – soll in der ‚bremenweiten Suche‘ nicht erscheinen.
  - Kennzeichnung von Dokumenten wie PDF ebenfalls zusätzlich mit dem Symbol und einem Tooltip, analog zu den Inhalten. Die Steuerung sollte mit den Funktionen (Mimetypes unter Weitere Container) übereinstimmen. Das bedeutet, es sollen alle bereits „bekannten“ Mimetypes übernommen werden.
  - Gibt es keine Treffer, wird der Hinweis unterdrückt.
- Es gibt keine erweiterte Suche nach der Ergebnisliste der ‚bremenweite Suche‘. Die erweiterte Suche bezieht sich ausschließlich auf die lokale Suche.
- Im Falle der Nichterreichbarkeit etc. der ‚bremenweiten Suche‘ muss sichergestellt sein, dass
  - die Anzeige der lokalen Ergebnisse nicht beeinträchtigt ist,
  - der Nutzer einen Hinweis erhält,
  - die Suchabfragen abgebrochen werden können (Timeout),
  - das Backend der dezentralen Instanzen nicht beeinträchtigt ist.
- Zu klären ist, ob tatsächlich alle Inhalte von bremen.de über die Instanzen gefunden werden sollten oder ob z.B. auf den Stadtplan verzichtet werden könnte. Außerdem sollte überlegt werden, die Ergebnisse der ‚bremenweiten Suche‘ für die KoGIs-Instanzen zu filtern und nur Ergebnisse aus der Verwaltung (d.h. von anderen KoGIs-Instanzen und dem Verwaltungsbereich von bremen.de) anzuzeigen. Wichtig ist dabei, dass die Steuerung möglich ist, d.h. den Algorithmus bei Bedarf anpassen werden kann.
- Da es vorkommen kann, dass in verschiedenen Quellsystemen dieselben Inhalte gefunden werden können (bspw. Inhalte der Visitenkarten, die in KoGIs-Instanzen gepflegt werden und deren Inhalte dann auch auf bremen.de vorhanden sind), d.h. die Suchmaschine sollte die Dopplungen so weit wie möglich über einen Algorithmus abfangen.

### Vorschlag für die Umsetzung der Suche in grafischer Form:

---

Das Suchergebnis, erweitert um den Link zur ‚bremenweite Suche‘ (hier gibt es verschiedene Möglichkeiten der Quellenangabe):

## > Suchergebnis

Gesucht wurde nach: **schicktanz**

Suchergebnis: 1-2 von 2

**Isabella Schicktanz**

Telefon: + 49 421 361 59962

Telefax: + 49 421 496 5337

E-Mail: [Isabella.Schicktanz@afz.bremen.de](mailto:Isabella.Schicktanz@afz.bremen.de)

### Organisation und Aufgaben

Leiter des Aus- und Fortbildungszentrums ist Holger Wendel (AFZ-L). Das AFZ ist untergliedert in vier Referate: Organisationsrechtlich als nicht-rechtsfähige Anstalt eigenständig, ...  
(unter: Kurzvorstellung | [Organisation und Aufgaben](#))

[zur erweiterten Suche](#)

### Weitere Suchergebnisse in der bremischen Verwaltung

Auf anderen Internet-Seiten der bremischen Verwaltung wurden XXX weitere Treffer gefunden.

[ERGEBNISSE AUS DER VERWALTUNG ANZEIGEN](#)

## > Suchergebnis in der brem. Verwaltung

Gesucht wurde auf den Seiten der brem. Verwaltung nach: **schicktanz**

Suchergebnis: 1-5 von 5

[Quelle: Senator für Finanzen]

### Organisation und Aufgaben [E→](#)

Projektleiterin des Teilprojektes Internet ist Isabella **Schicktanz**. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. ...

### Organisation und Aufgaben [E→](#)

[Quelle: Senator für Finanzen]

Projektleiterin des Teilprojektes Internet ist Isabella **Schicktanz**. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. ...

### Organisation und Aufgaben [E→](#)

[Quelle: Senator für Finanzen]

Projektleiterin des Teilprojektes Internet ist Isabella **Schicktanz**. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. ...

### Organisation und Aufgaben [E→](#)

Projektleiterin des Teilprojektes Internet ist Isabella **Schicktanz**. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. ...

[Quelle: Senator für Finanzen]

### Organisation und Aufgaben [E→](#)

Projektleiterin des Teilprojektes Internet ist Isabella **Schicktanz**. Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed diam nonummy nibh euismod tincidunt ut laoreet dolore magna aliquam erat volutpat. ...

[Quelle: Senator für Finanzen]

## [bremen.de](http://bremen.de)

---

- Eine Filterung nach allgemein verständlichen Kategorien und unterschiedliche Sortiermöglichkeiten der Treffer sind möglich (s. berlin.de), etwa nach
  - Behördeninformationen
  - Veranstaltungen
  - Adressen
  - Straßennamen / Stadtplanausschnitten
  - Artikel
  - Pressemitteilungen
- Es müssen bei Bedarf neue Kategorien von Dokumenten in die Suche integriert werden können, z.B. Schlagwörter oder Kleinanzeigen.
- Die Ansicht der Trefferlisten muss durch uns anpassbar sein, z.B. möchte b.o. ggf. bei Straßennamen auch einen Stadtplanausschnitt mit Google Maps in die Ergebnisliste einbauen können.
- Die Länge der Ergebnisanzeige soll variabel eingestellt werden können.
- Redaktionelle Steuerung der Relevanzsortierung
- Individuelles Highlighten / Boosten von Treffern
- Key-matching-Funktionalität, d.h. speziell ausgewählte URLs stehen am Anfang der Trefferliste
- Bei Bedarf müssen PremiumPlus Visitenkarten besonders hervorgehoben werden können.
- Textboxen mit Tipps oder redaktionell aufbereitete Trefferseiten können eingebaut werden.
- Bei 0-Treffer-Meldung und ev. auch auf Ergebnisseiten können individuell programmierte Tagclouds eingebaut werden. (können wir gesondert programmieren, muss auf der Ergebnisseite integriert werden)
- Am Anfang aller Suchergebnisse werden bezahlte Textlinks zu den Suchbegriffen angezeigt (wird gesondert programmiert, muss auf der Ergebnisseite integriert werden)
- Grundsätzlich muss die Indexierung steuerbar sein, um auch die Einbindung des Thesaurus (Gewichtung der Schlagwörter/Synonyme) zu gewährleisten.