

# Nutzungsauswertung für Geowebdienste

Georg Ulbrich

Potsdam, 17.03.2016

## Einstieg



2 / 21

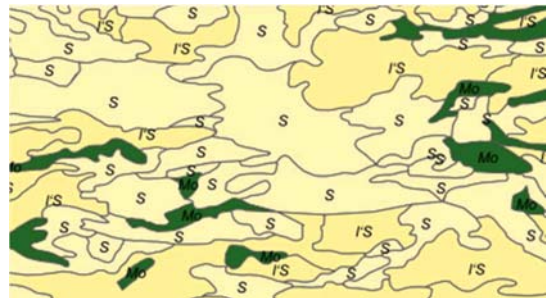
## Geowebdienste – standardisierte Operationen

- Standardisierung durch Open Geospatial Consortium (OGC)
- Verbreitete Dienstarten
  - Web Map Service (WMS): Darstellungsdienst
  - Web Map Tile Service (WMTS): Darstellungsdienst
  - Web Feature Service (WFS): Downloaddienst (Vektordaten)
  - Web Coverage Service (WCS): Downloaddienst (Rasterdaten)
  - Catalogue Service for the Web (CSW): Katalogdienst (Metadaten)
  - Web Processing Service (WPS): Prozessierungsdienst (z.B. Koordinatentransformation)
- Aufruf der Dienste erfolgt über URL des Dienstes ...  
z. B.: [http://inspire.brandenburg.de/services/boartsubstr\\_wms?](http://inspire.brandenburg.de/services/boartsubstr_wms?)
- ... mit Aufruf von **Parametern**

## Geowebdienste – standardisierte Operationen

- **SERVICE** WMS, WFS, WCS, ...
- **VERSION** 1.3.0, 2.0.0, ...
- **REQUEST** GetMap/GetFeature, GetCapabilities, ...
- **EXCEPTIONS** Format der Rückgabe bei Fehler

[http://inspire.brandenburg.de/services/boartsubstr\\_wms?  
SERVICE=WMS&  
VERSION=1.3.0&  
REQUEST=GetMap&  
LAYERS=vdlufta&  
FORMAT=image%2Fpng&  
EXCEPTIONS=XML&  
CRS=EPSG%3A25833&  
BBOX=419261,5871954,430563,5883256&  
WIDTH=400&HEIGHT=220](http://inspire.brandenburg.de/services/boartsubstr_wms?SERVICE=WMS&VERSION=1.3.0&REQUEST=GetMap&LAYERS=vdlufta&FORMAT=image%2Fpng&EXCEPTIONS=XML&CRS=EPSG%3A25833&BBOX=419261,5871954,430563,5883256&WIDTH=400&HEIGHT=220)



## Nutzungsauswertung für Geowebdiensten

- Geowebdienste sind *der* zentrale Bestandteil einer GDI
- Bislang eher wenig Kenntnisse über die Nutzung von Geowebdiensten
- Auswertung ist Grundlage für:
  - ... das Berichtswesen zum INSPIRE-Monitoring: Durchführungsbestimmung Monitoring & Reporting (2009/442/EG, Artikel 9)
  - ... eigene Erkenntnisse

## Nutzungsauswertung für Geowebdiensten

- Zentrale Fragestellungen:
  - Wie häufig wird (m)ein Geowebdienst genutzt?
  - Welche Arten der Nutzung (requests) gibt es anteilig?
  - Durch wen wird (m)ein Geowebdienst genutzt?
  - Verhältnis der Nutzung von Darstellungs- und Downloaddiensten?
  - Wo wird (m)ein Darstellungsdienst genutzt?
- Gesucht: Software / IT-Systeme zur Auswertung der Nutzung
- Analysesoftware für **Webseiten** sehr weit verbreitet:
  - Piwik, Google Analytics, AWStats, Webalizer, etracker, ...

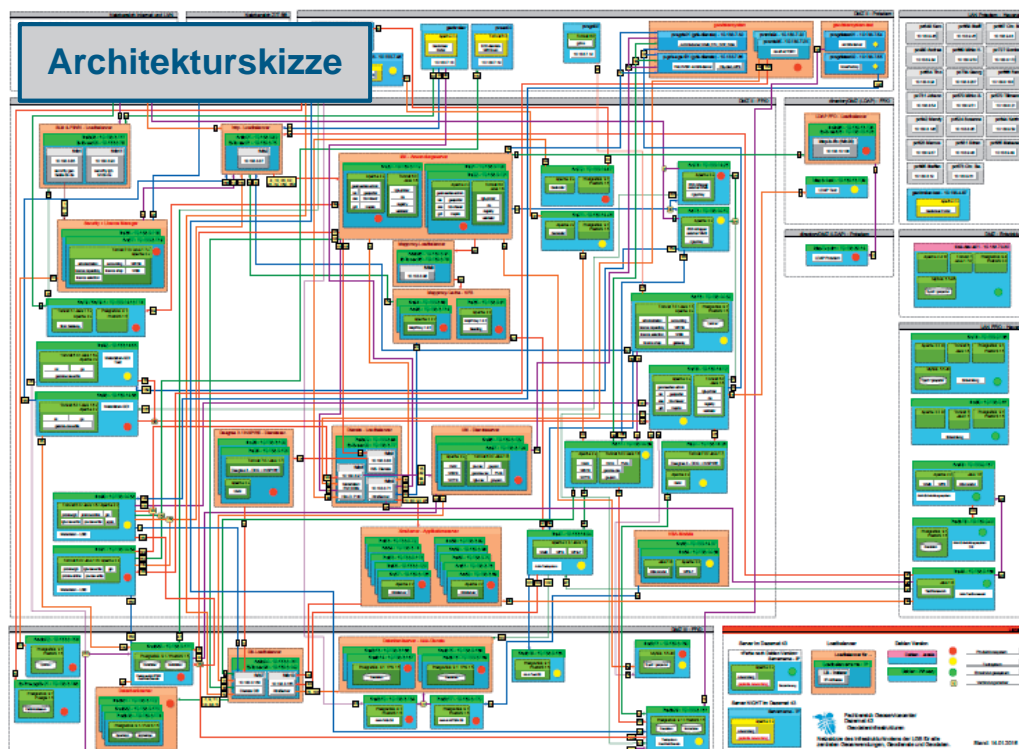
## Kenntnisse über die Nutzung von Geowebdiensten

- Verfahren:
  - Clientseitig (JavaScript, Cookies, Zählpixel)
  - Serverseitig (Auswertung von Server-Logfiles)
- Geowebdienste nur serverseitig auswertbar
- **Keine** Standard-Software für Auswertung von Geowebdiensten
- Umfangreiche Tests mit Piwik
  - Grundproblem: Viele verschiedene Dienste auf einem Serververbund, daher Filterung notwendig
- Bislang leider nicht erfolgreich ...
- Daher: Eine eigene Lösung muss (zunächst) her!

## Rahmenbedingungen

### Server-Architektur für Geowebdienste in der LGB

- Getrennte Systeme für Entwicklung, Test, Produktion
- Produktion: 44 Server mit ca. 275 Cores / 1 TB Ram / 4 TB Disk
- Geowebdienste der LGB (ISK) und für externe Kunden (INSPIRE)
- Je drei Server für ISK & INSPIRE hinter Loadbalancer
- 6 Apache Logfiles täglich (ca. 700 MB / 5,5 Mio. Zugriffe )
- >250 Geowebdienste (WMS, WFS, WCS, WMTS, CSW)



### Umsetzung des Tools



- Implementierung in Python
- Automatisierung über Cronjobs
- Überwachung per Check\_MK und Benachrichtigungen per E-Mail
- Nur ca. 700 Zeilen Code (31 KB)
- Verarbeitung der Logfiles eines Tages:
  - Prozessierungszeit: ca. 10 Minuten (Vergleich Piwik: ca. 8 Stunden)
  - Ram-Bedarf: ca. 50 MB, volle CPU-Auslastung

### Aufbau des Tools

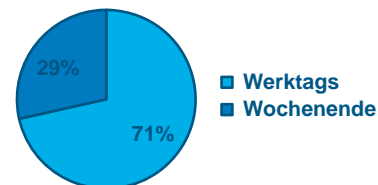
- Einlesen der Logfiles
- Zerlegung der Zeilen in Informationen (regulärer Ausdruck)
  - URL & Parameter, Method (post/get), Referrer, UserAgent, Status
  - Name Webservice, Type Webservice (WMS, WFS, ...), Request
- Filterung unerwünschter Einträge
  - Bots (SuMa, etc.), Checks (Check\_MK), Tests
  - Blacklists für: Referrer, UserAgent, Query
- Strukturierte Speicherung in einer Datenbank
  - Datum, Name Webservice, Typ Webservice, Request, Anzahl
  - Aggregiert für jeweils einen Tag

## Nutzungsstatistiken

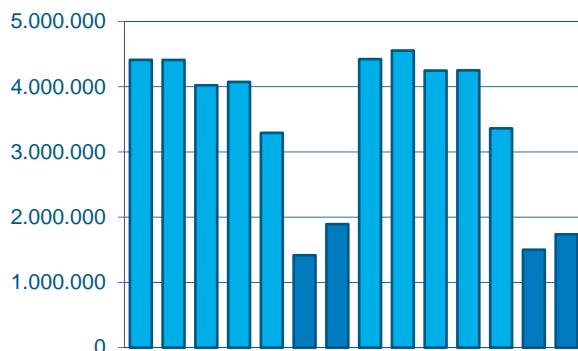
- Beispielhaft für die 6. + 7. KW (08.02. - 21.02.2016) – gefiltert
- Auswertung der Zugriff für:
  - alle Dienste auf isk.geobasis-bb.de & inspire.brandenburg.de
  - einen ausgewählten INSPIRE-Dienst (WMS+WFS)

## Anzahl der Zugriffe, gesamt

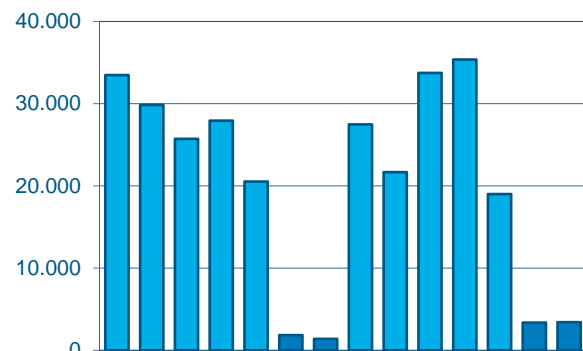
nach Wochentag (alle Dienste)



gesamt (alle Dienste)

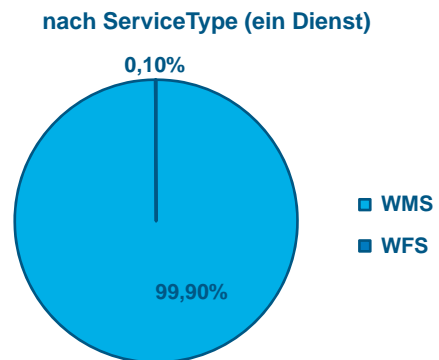
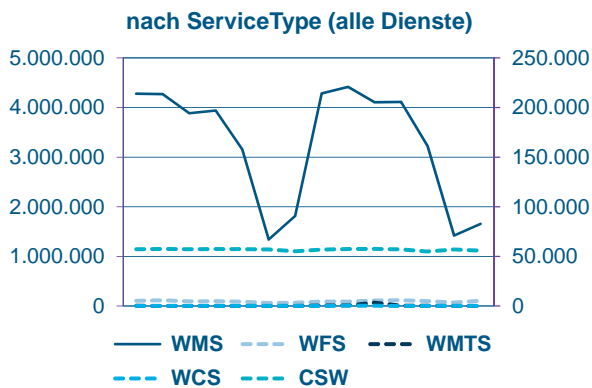


gesamt (ein Dienst)



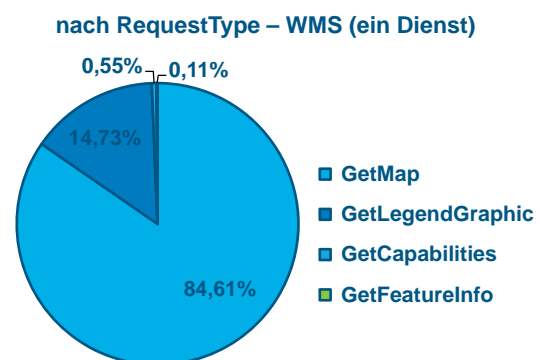
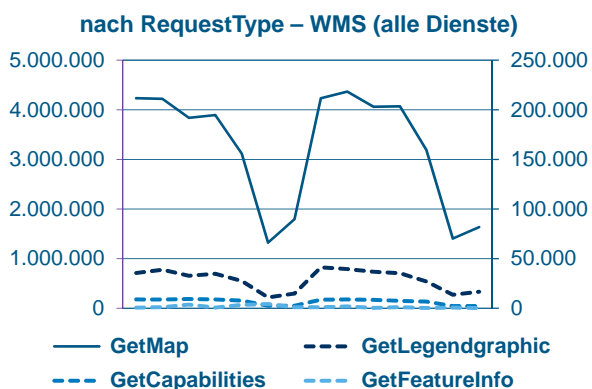
- Deutlicher Unterschied zwischen Werktagen und Wochenenden (Freitags ebenfalls geringer ...)
- **Fazit:** Nutzung der Webdienste überwiegend beruflich, nicht privat

### Anzahl der Zugriffe nach ServiceType



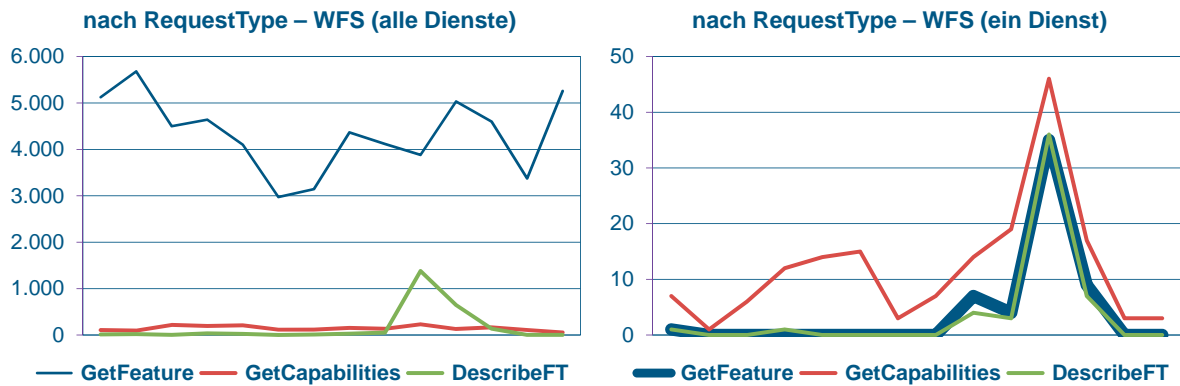
- WMS am weitesten starksten genutzt
- **Offen:** Wird der WFS nur zum einmaligen Herunterladen genutzt?
- **Offen:** Sind den Nutzern Downloaddienste hinreichend bekannt?

### Anzahl der Zugriffe nach RequestType – WMS



- GetLegendGraphic: einmalig bei WMS-Aufruf durch Portal/GIS
- **Offen:** Anzahl der GetLegendGraphic-Aufrufe als Hinweis für die Anzahl der Nutzungen bzw. Nutzer?

### Anzahl der Zugriffe nach RequestType – WFS



- GetFeature (tatsächlicher Datenabruf): bei „allen Diensten“ erwartungsgemäß; bei „einem Dienst“ kaum tatsächliche Datenabrufe
- **Fazit:** Nutzung von Downloaddiensten befördern

### POST-Abfragen

- Parameter werden nicht per URL übergeben -> nicht im Logfile
- ServiceType nicht ableitbar (ggf. aus ServiceName)
- Nutzung von WFS kaum über GET realisierbar

ISK			INSPIRE		
service_name	service_type	Anzahl	service_name	service_type	Anzahl
dog_lgb200_bebb_v3	n/a	550356	bobasensorp	wfs	7
lgb	n/a	8074	borelief	wfs	4
dnmvg	n/a	950	boschwerm	wfs	3
dnm2012	n/a	704	boartsubstr	wfs	2
gaz	n/a	586	boerosion	wfs	1
dog_lgb200	n/a	319	bokarten	wfs	1
easytousegazetteer_dog2	n/a	69	bohumuskohl	wfs	1
dnmbs	n/a	31	strassennetz	wfs	1
proxy	n/a	6			
brw_wms2011	n/a	2			

- **Fazit:** POST-Anfragen (bei den INSPIRE-Diensten) kaum genutzt



### Sonstige Auswertungen


- Top-Liste der meistgenutzten Dienste (HighScore)
- Analysen zur Entwicklung über die Zeit
- Detektion von „Ausreißertagen“

### Spezielle Auswertungen


- Analyse von fehlerhaften Aufrufen (über http-Statuscode 4xx)
- Auswertung der Referrer:  
In welchen Geoportalen sind meine Dienste eingebunden?
- Auswertung der abgerufenen Koordinatenreferenzsysteme
- Räumliche Auswertung von WMS-Abfragen („bbox“+„crs“) ...


## Digitale Orthophotos Land Brandenburg – DOP20


- DOP20: Pixelkantenlänge 0,2 m



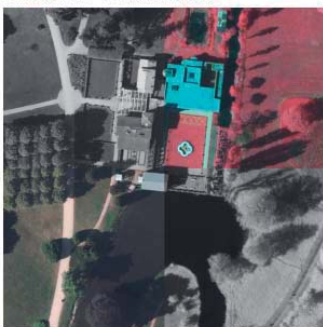
**GEOBROKER**  
Der Internetshop der LGB

  
Bestellübersicht

  
Downloadcenter

  
Warenkorb

Startseite » DOP - Produktmetadaten



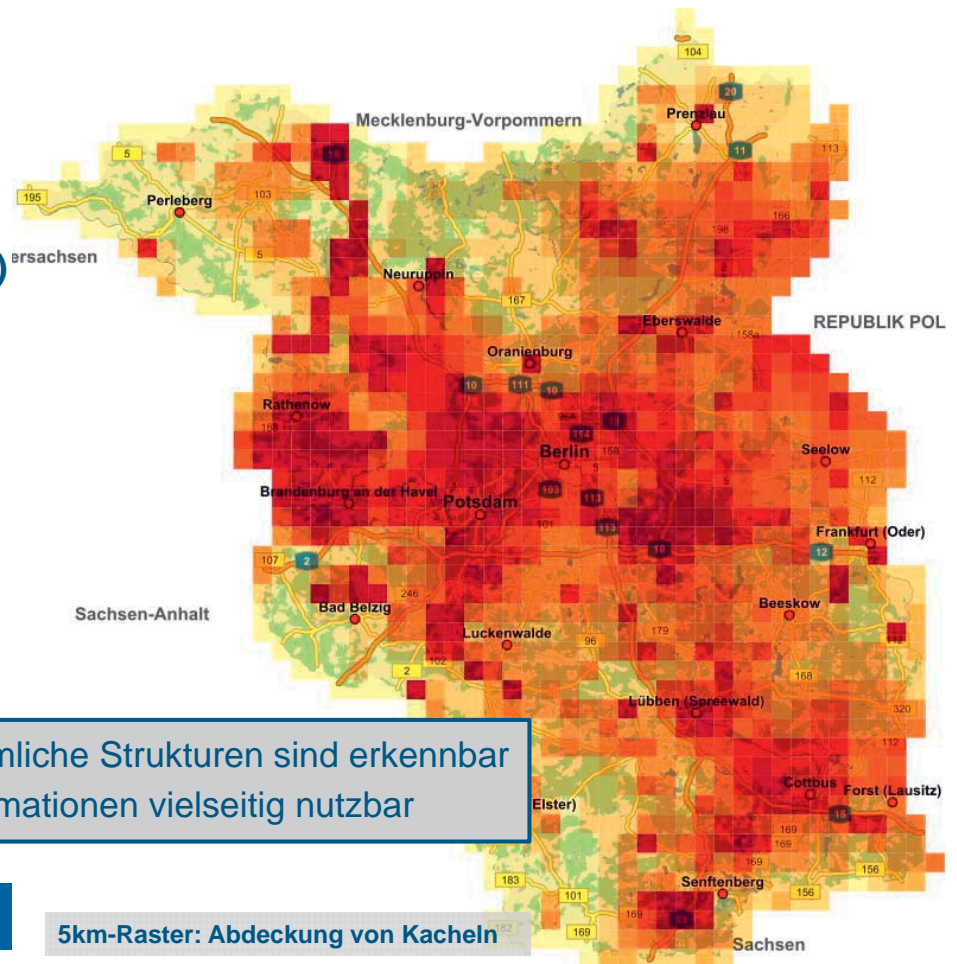
### Digitale Orthophotos Land Brandenburg - DOP

Digitale Orthophotos sind digital entzerrte Luftbilder und bilden alle zum Aufnahmezeitpunkt luft sichtbaren Objekte und Sachverhalte parallelperspektivisch ab. Sie besitzen eine hohe Informationsdichte zu ökologischen, phänologischen, geographischen und anderen Themen. Die DOP sind Farbbilder mit Infrarotkanal. Der Ausgabemaßstab beträgt 1 : 5 000. Die Bodenauflösung bzw. Pixelgröße beträgt 0.2 m x 0.2 m (635 dpi = 250 L / cm). Auslieferung in Farbe mit Infrarotkanal (RGBl, 32 Bit), in Farbe (RGB, 24 Bit), in Farbinfrarot (CIR, 24 Bit), in schwarz-weiß (8 Bit) oder in schwarz-weiß Infrarot (8 Bit), blattschnittfrei; ETRS89 (GRS80, UTM-Abbildung) bzw. Transformation in andere Bezugssysteme möglich. Die Datenabgabe über den Geobroker erfolgt blattschnittfrei. Wenn die Größe des Gebietes in eine Richtung 2000 m überschreitet wird das Gesamtgebiet automatisch in 2000m x 2000m Kacheln aufgeteilt.

[Weitere Informationen zum Produkt erhalten Sie auf dieser Seite »](#)

## Ergebnisse

08.02.2016  
DOP20 (WMS)



17.03.2016 | Georg Ulbrich

5km-Raster: Abdeckung von Kacheln

## Nutzungsauswertung für Geowebdienste

20 / 21

### Fazit

- *Out-of-the-box* Lösung für Auswertung von Geowebdiensten fehlt
- Flexible Auswertemöglichkeiten durch eigene Implementierung
- Fachliche Interpretation der Ergebnisse notwendig
- Ergebnisse liefern Grundlage für INSPIRE-Monitoring und Berichte an Kunden der INSPIRE-Zentrale (ab Q2/2016)
- Nutzung von WFS-Diensten befördern!

### Ausblick

- Implementierung von *on-the-fly* Visualisierung der Ergebnisse
- Integration in Geschäftsprozesse

17.03.2016 | Georg Ulbrich

Nutzungsauswertung für Geowebdienste

**Vielen Dank  
für Ihre  
Aufmerksamkeit!**

