

---

# Virtuelle Integration von Datenquellen mit einer Graph-Datenbank (Projekt ArcoFaMa)

Prof. Dr. Petra Sauer  
Dipl.-Inf. Frank Herrmann  
M.Sc. Jan Matusewicz



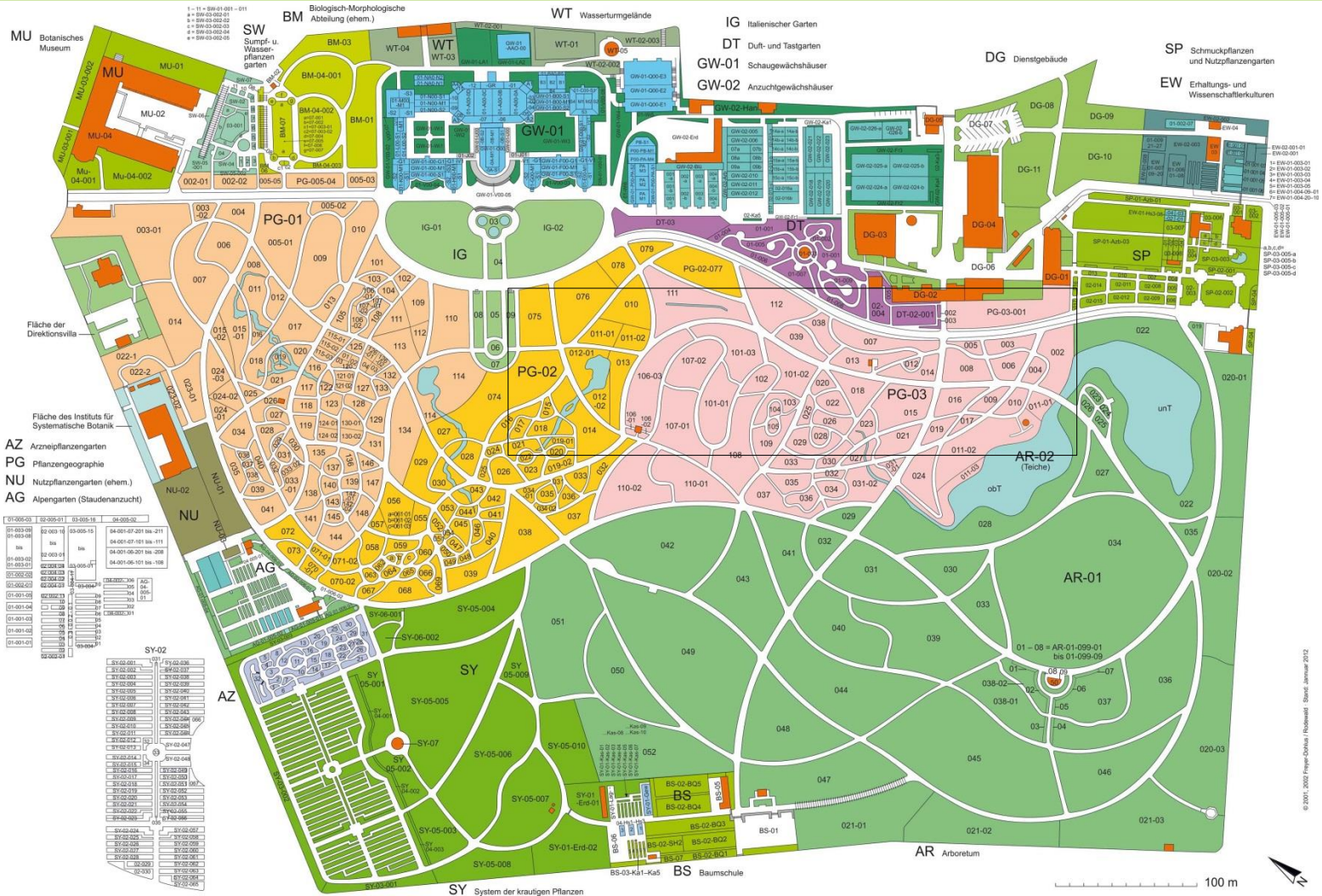




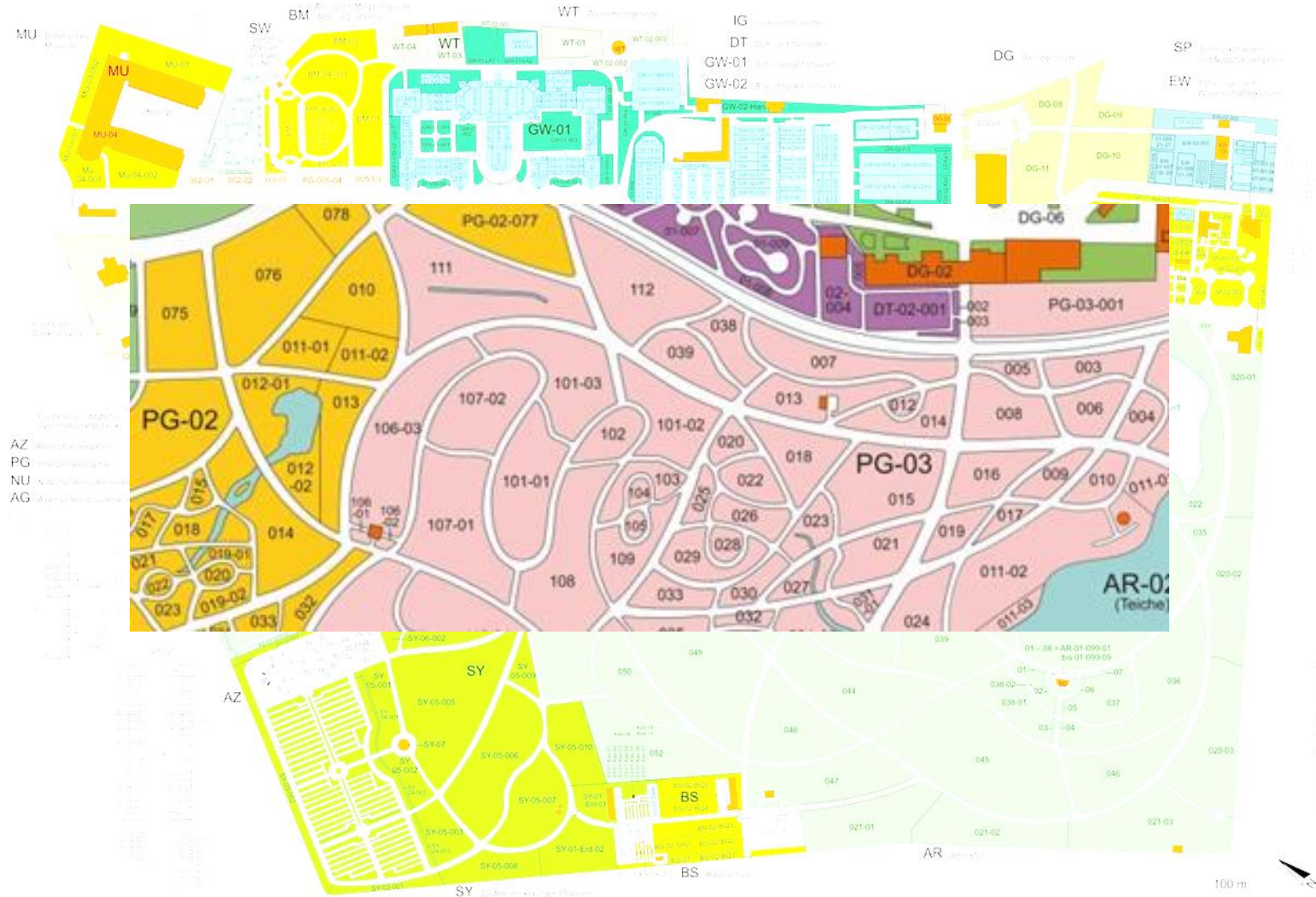
# Geoinformation im Kontext des Botanischen Gartens



# Geoinformation im Kontext des Botanischen Gartens



# Geoinformation im Kontext des Botanischen Gartens



## 4000 Bäume in Baumdatenbank



Gattung/Art	Baum	BG-Ort	Kategorie	Alter	Kronenvitali	Baumhöhe [i	Kronendurch	Stammdm. [i Ka
<i>Ostrya virginiana</i>	5	PG-03-001	Pflanzengeo	63	3	über 10	9	50
<i>Celtis occidentalis</i> L.	6	PG-03-001	Pflanzengeo	63	3	über 10	10	49
<i>Aesculus glabra</i> Willd.	7	PG-03-001	Pflanzengeo	78	2	über 20	9	85
<i>Aesculus glabra</i> Willd.	8	PG-03-001	Pflanzengeo	68	2	über 15	8	60
<i>Ostrya virginiana</i>	9	PG-03-001	Pflanzengeo	48	3	über 15	9	39
<i>Fraxinus americana</i> L.	10	PG-03-001	Pflanzengeo	63	2	über 15	13	56
<i>Quercus rubra</i> L.	11	PG-03-001	Pflanzengeo	68	3	über 15	13	55
<i>Quercus rubra</i> L.	12	PG-03-001	Pflanzengeo	68	3	über 15	12	60
<i>Quercus rubra</i> L.	13	PG-03-001	Pflanzengeo	63	3	über 15	9	49
<i>Liriodendron tulipifera</i> L.	14	PG-03-001	Pflanzengeo	63	2	über 20	10	48

Entfernung zu Baum 10?

Genauer Standort ?

Weg zum Ausgang ?

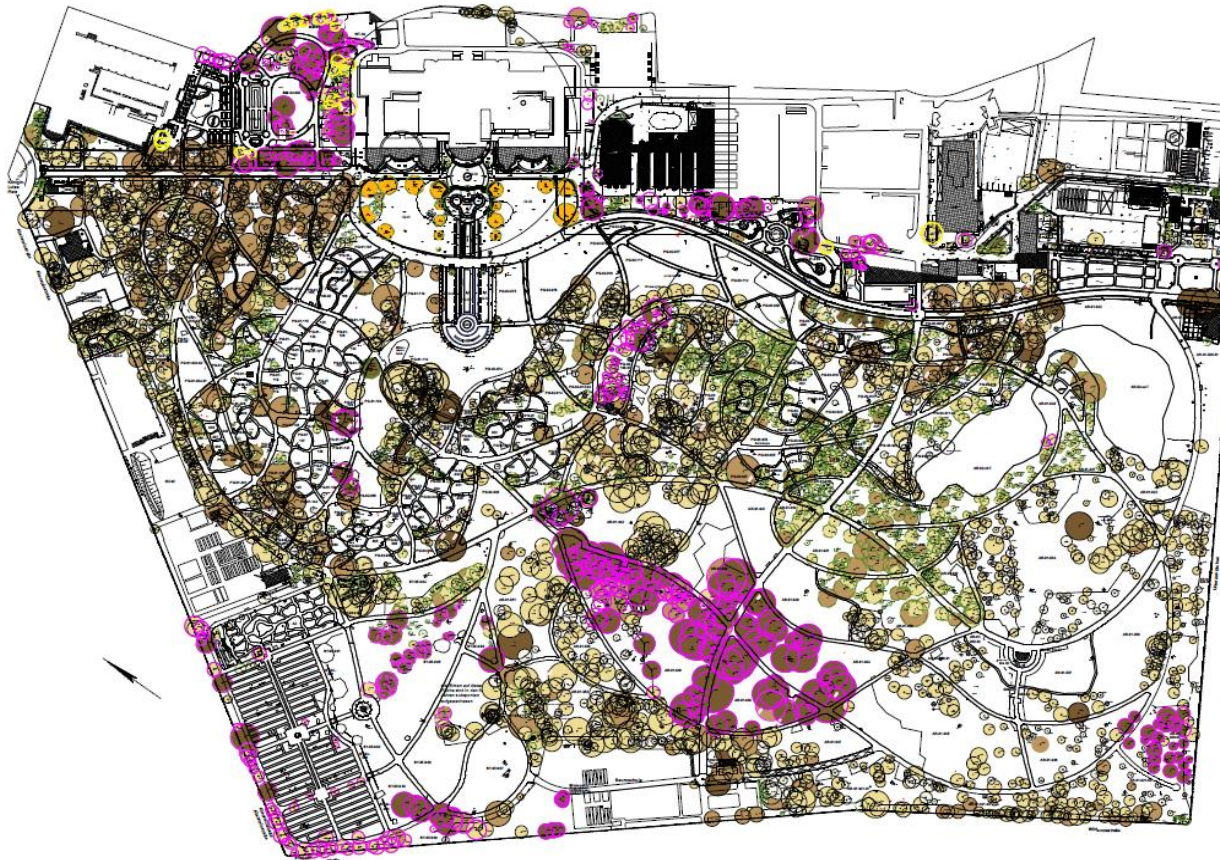


Nächster Hydrant ?

Gattung/Art	Baum	BG-Ort	Kategorie	Alter	Kronenvitali	Baumhöhe [m]	Kronendurch [m]	Stamm dm. [Kc]
<i>Ostrya virginiana</i>	5	PG-03-001	Pflanzengeo	63	3 über 10	9	50	
<i>Celtis occidentalis</i> L.	6	PG-03-001	Pflanzengeo	63	3 über 10	10	49	
<i>Aesculus glabra</i> Willd.	7	PG-03-001	Pflanzengeo	78	2 über 20	9	85	
<i>Aesculus glabra</i> Willd.	8	PG-03-001	Pflanzengeo	68	2 über 15	8	60	
<i>Ostrya virginiana</i>	9	PG-03-001	Pflanzengeo	48	3 über 15	9	39	
<i>Fraxinus americana</i> L.	10	PG-03-001	Pflanzengeo	63	2 über 15	13	56	
<i>Quercus rubra</i> L.	11	PG-03-001	Pflanzengeo	68	3 über 15	13	55	
<i>Quercus rubra</i> L.	12	PG-03-001	Pflanzengeo	68	3 über 15	12	60	
<i>Quercus rubra</i> L.	13	PG-03-001	Pflanzengeo	63	3 über 15	9	49	
<i>Liriodendron tulipifera</i> L.	14	PG-03-001	Pflanzengeo	63	2 über 20	10	48	



## 300 CAD-Dateien



- BAUMBESTAND  
ALTERSSTRUKTUR**
- Einstufung nach d.b.g / Ö&P**
- Gehölze aus dem alten Botanischen Garten stammend
  - Gehölze der Erstanlage, 1899-1914
  - Gehölze der Kriegs- und Zwischenkriegszeit, 1915 - 1945
  - Gehölze der Nachkriegszeit, 1946 - 1973
  - Gehölze ab 1974
- Einstufung nach Herr Marquardt**
- Alter von Herr Marquardt bestimmt, jedoch nur für Teilbereiche
  - dsgl. Zuordnung nicht zweifelsfrei
  - lt. Herr Marquardt subspontane Verbreitung (ohne Alterszuordnung bzw. Zuordnung nach Ö&P / d.b.g.)
- Stand 04.08.2009
- Nadelgehölz

**VORABZUG**

**BOTANISCHER GARTEN  
BERLIN-DAHLEM** Pflegeamt

**Auftraggeber:** Landes-Immunort Berlin  
Gartenbauamt  
Königsplatz 47 10715 Berlin-Mitte

**Eigenkriterium / Zeichner:** Botanischer Garten und  
Städtisches Museum Berlin-Dahlem (StM)M  
Zweibereichung der Freien Universität Berlin  
Königsplatz 47, 10715 Berlin

**Planung / Inhalt:** TOPOS  
Bodenkarte, Blatt 28  
13115 Berlin  
Tel.: 030 1 564 40 40  
mailto:mg@postplanung.de

**Planinhalt:** ÖKOLOGIE & PLANUNG  
Bodenkarte Blatt 28  
13115 Berlin - Siegel  
Tel.: 030 1 564 40 40  
Ökologie-Planung@fu-berlin.de

**Planinhalt:** **BAUMBESTAND  
ALTERSSTRUKTUR**

**Datum:** 08.04.2009  
**Zeichner:** G.F.  
**Jahr:** 2009

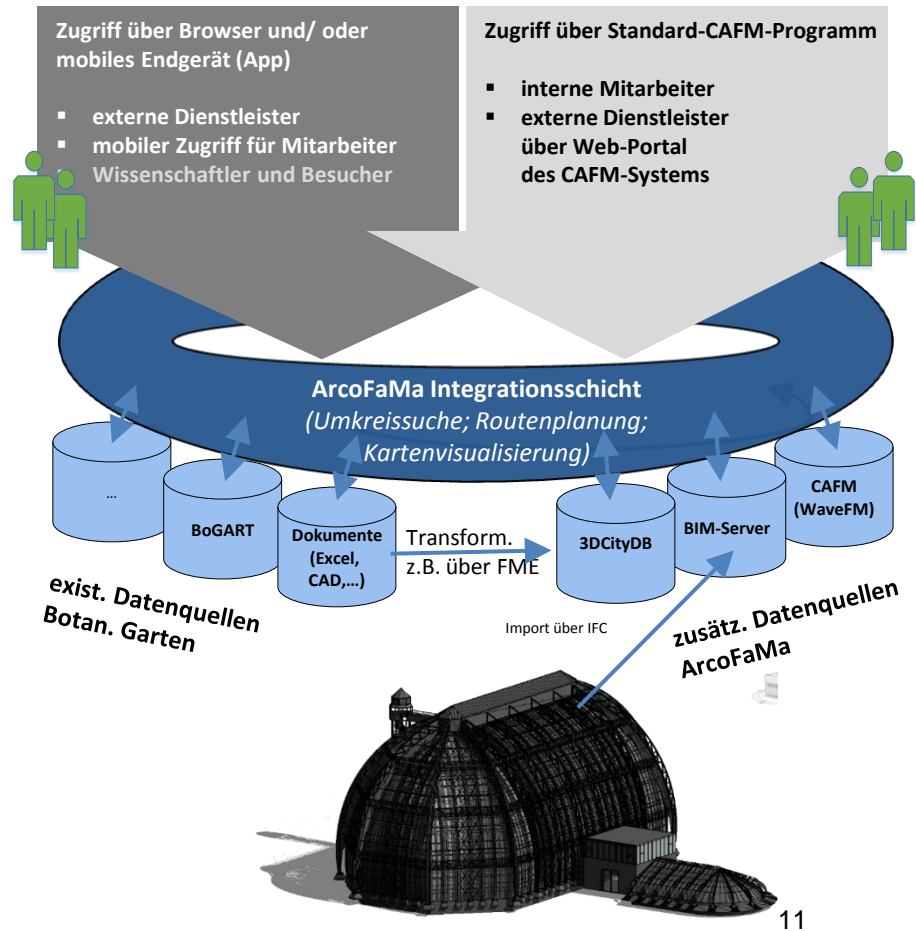
**Plannummer:** ÖF\_BG\_121  
**Maststab 1:0:** 1:1.000

# Projekt ArcoFaMa: Zielstellung und Ansatz

- Ziel
  - Geometrien und Geodaten aus proprietärem CAD-Format in offenes, interoperables Format überführen
  - Geodaten u.a. für das Facility Management nutzbar machen
  - Integration mit Datenbeständen des Botanischen Gartens (BGBM)
- Problematik
  - überwiegend isolierte Datenhaltung von Geo- und FM-Daten
  - Überwiegend isolierte Datenhaltung von Geo- und BGBM-Daten
- Ansatz
  - Regelbasierte Transformation von Geodaten aus CAD-Daten
  - Einheitliche Datenhaltung in Geodatenbank
  - **Aufbau einer Integrationsplattform**, um heterogene Daten gemeinsam zu verarbeiten
  - Entwicklung von Anwendungen, die die Integrationsplattform nutzen

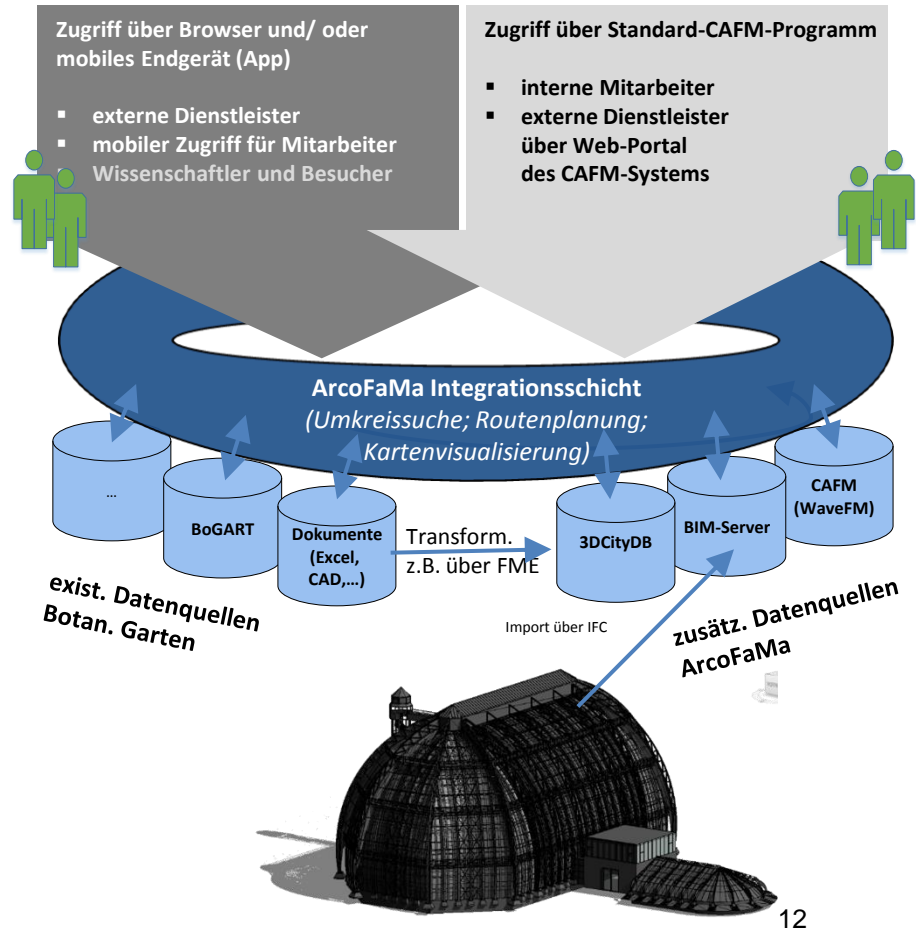
## Herausforderungen

- Verteilung der Datenquellen
- Autonomie der Datenquellen
- Beachtung der Heterogenität
  - Datenmodell (relational, OO, XML)
  - Strukturell (Schema)



## Entscheidungen

- Virtuelle Integration
  - Verknüpfte Informationen verbleiben in den ursprünglichen Datenbanken
  - Anfrageschnittstellen nutzen
  - Keine schreibenden Zugriffe auf Datenquellen
  
- Metadatenbasierter Ansatz



- Informationen zur Beschreibung anderer Daten
- Im Kontext einer Integrationsplattform:
  - Welche Datenquellen sind verfügbar?
  - Wie erreicht man diese Datenquellen? (Technologie, Zugriff)
  - Wie sind die Datenquellen strukturiert? (Abbildung)
  - Wie können die Datenquellen miteinander verknüpft werden? (Klassifizierung Knoten / Kanten)
  - Welche impliziten Daten können mit einbezogen werden? (Informationsanreicherung)
- Notwendigkeit des Einsatzes eines effizienten Datenspeichers mit flexiblen Verwaltungsfunktionalitäten

## Property-Graphen

- kombinieren mehrere Besonderheiten:
  - Knoten verfügen über Eigenschaften (Schlüssel–Werte-Paare)
  - Kanten sind gerichtet und verfügen über einen Namen
  - Kanten können über weitere Eigenschaften verfügen
- Property-Graphen sind gewichtete, gerichtete und benannte Multigraphen
- Hypergraphen erlauben die Verbindung von Knoten mit mehr als einer Kante
- M/N – Abbildungen können somit vereinfacht dargestellt werden

## Prozessierung der Daten

- Native Processing Engine – beherrscht indexfreie Adjazenz
- Wesentliches Merkmal der effizienten Prozessierung
- Knoten haben direkten Verweis auf den Nachbarknoten
- Kein ‚Nachschlagen‘ in einem globalen Index nötig
- Großer Geschwindigkeitsvorteil gegenüber non native Graph-DB
- Neo4j ist bezüglich der Speicherung und Prozessierung eine native Graph-Datenbank

## Neo4j – Open-Source-Graph-DB

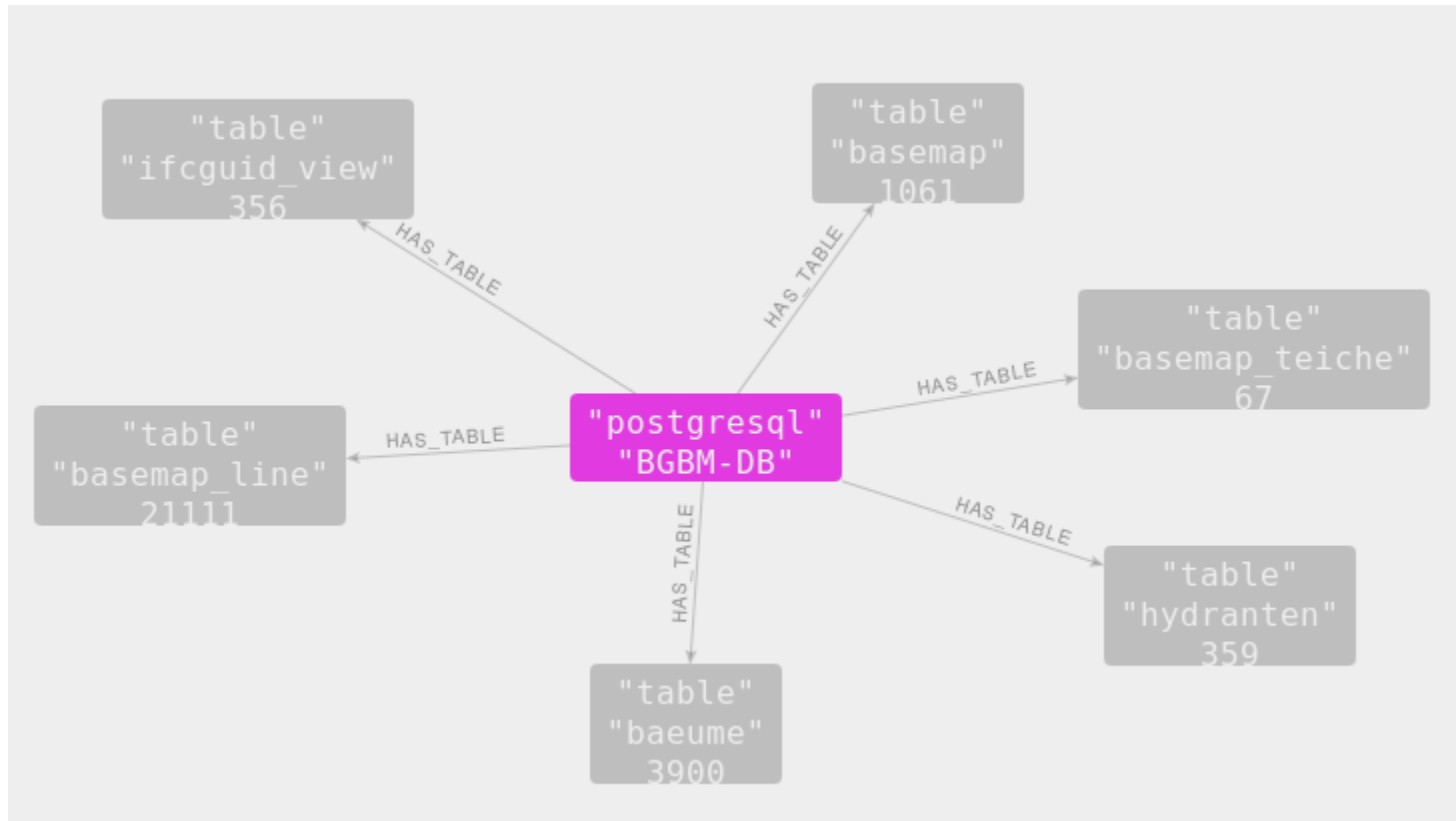


- Native Processing Engine – beherrscht indexfreie Adjazenz
- Seit 2003 in Entwicklung, Firma Neo Technology
- Implementiert in der Programmiersprache Java
- Version 1.0 im Jahre 2010, aktuell Version 2.0.1
- Eigens entwickelter Persistenzmechanismus für die Speicherung und Verwaltung
- Persistenzschicht nutzt per Java NIO Blöcke fester Größe zur Speicherung von Knoten und Kanten



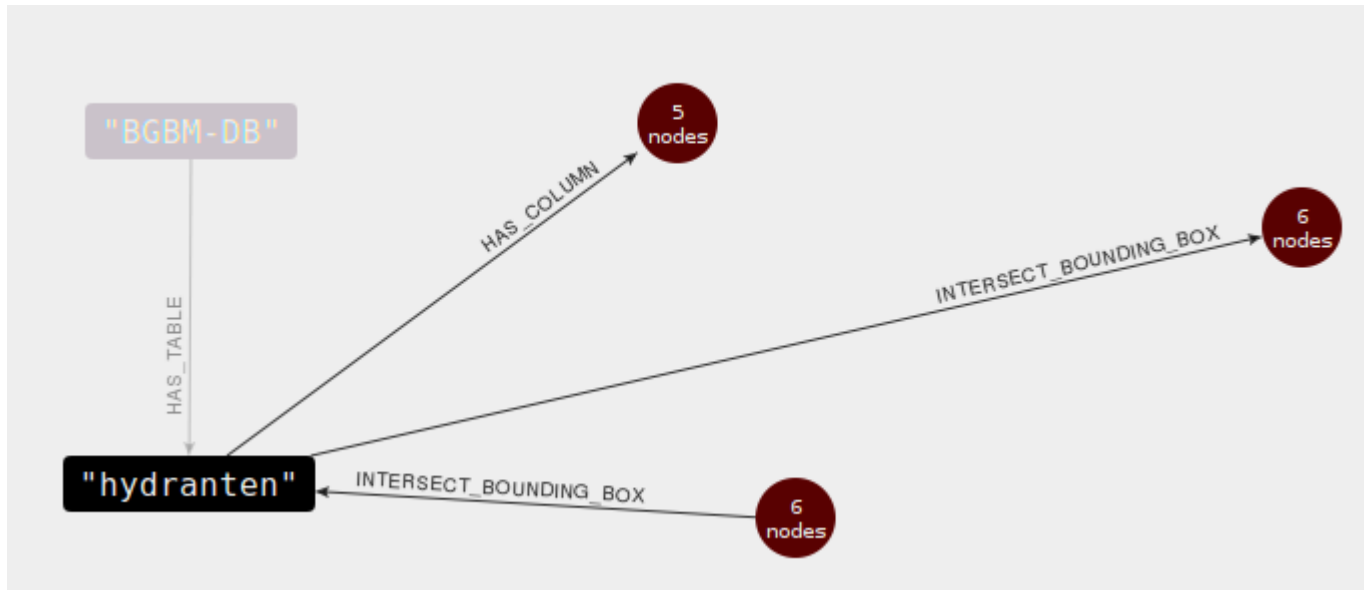
- Neo4j – Open-Source-Graph-DB
- Unterstützt Transaktionen per JTA, vollständige ACID Eigenschaften / Garantien
  - *Atomarität, Konsistenz, Isoliertheit, Dauerhaftigkeit*
- Datenzugriff per
  - Java API
  - Abfragesprache Cypher
  - Traverser Framework
  - REST Schnittstelle
- Abfrage, Verwaltung und Visualisierung über eine Webapplikation

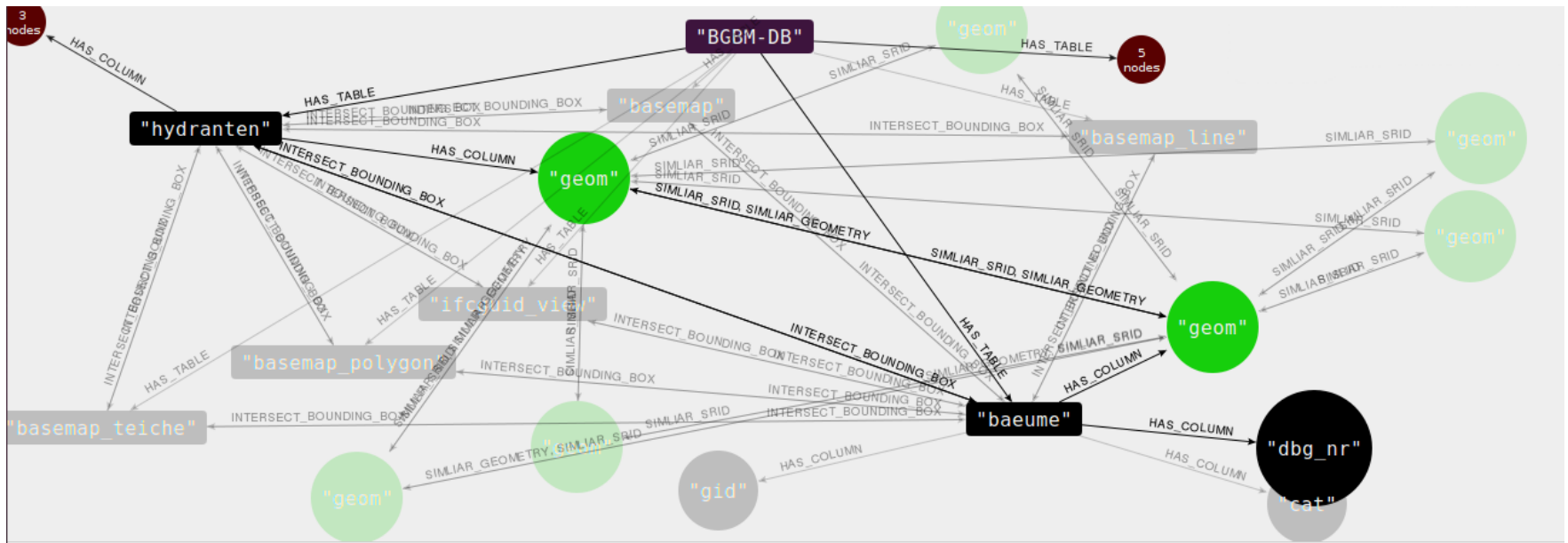


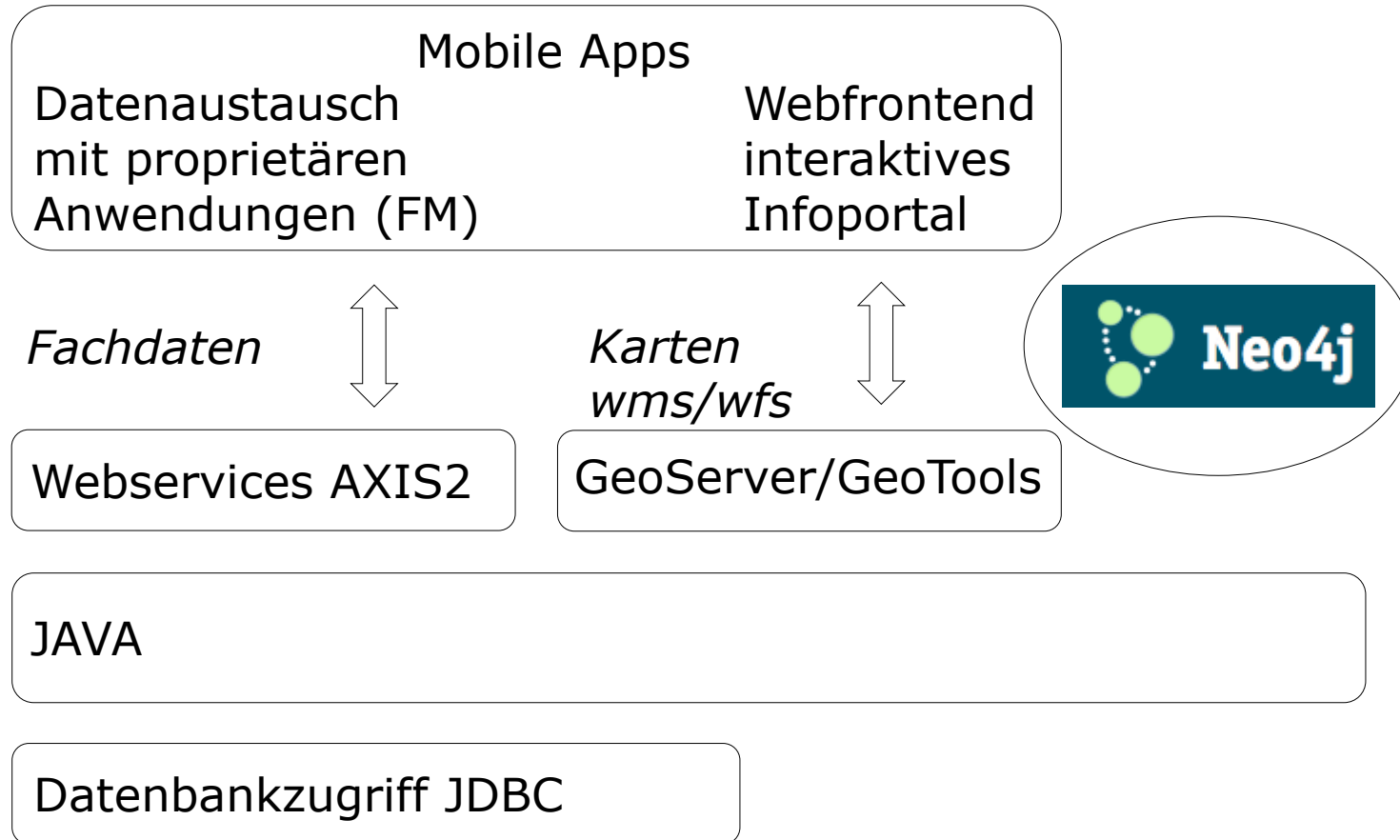


## Postgresql / Postgis

- Räumlicher Datentyp Geometry (Geography)
- Abfrage nach Geometrietyp und SRID (Koordinatenreferenzsystem)
- Eigenschaften werden im Knoten gespeichert
  - Geometrietyp (POINT, POLYGON etc.)
  - SRID
  - Tabellenknoten hält die BoundingBox-Parameter
- Identifikation möglicher Relationen über Tabellen hinweg
  - Gleichheit der SRID und Geometrietyp
  - Überschneidung/Überlappung der von den BoundingBoxen abgedeckten Zonen







- **CAFМ-System**
  - Wave Facilities (Loy und Hutz)
- **ArcoFaMa Integrationsplattform**
  - GeoServer ([www.geoserver.org](http://www.geoserver.org))
  - NoSQL-Datenbank Neo4j (Metadaten)
  - Tomcat 6
- **Verknüpfte Datenquellen**
  - PostgreSQL 9.3 (Geodatenbank)
    - PostGIS 2.0.4
    - 3DCityDB-Schema
  - BIM-Server 1.2 ([www.bimserver.org](http://www.bimserver.org) )
  - BoGart-Datenbank (MS SQL-Server)
  - Wave Facilities-Datenbank (MS SQL-Server)
- **Schnittstellenstandards**
  - Industry Foundation Classes (IFC 4)
  - CityGML 2.0
  - OGC-Services (WebMapService, WebFeatureService)
  - JSON (Webservices)
- **Software zur Konvertierung**
  - FME (Feature Manipulation Engine); Safe Software Inc. (Transformation)
  - OpenJump (Java GIS – Rotation, Koordinationssystem)
  - GrassGIS 7 (Koordinatenumrechnung)

- Umfangreiche Aufgabenstellung
- Umfangreiche empirische Untersuchungen
- Best Practice – Ansätze finden
- Baustellen :
  - Transformation CAD-Daten / CityGML
  - Anbindung FM-Daten im IFC-Format
  - Weiterentwicklung Integrationsplattform



# Kontaktdaten

- Prof. Dr.-Ing. Markus Krämer  
Hochschule für Technik und Wirtschaft Berlin  
Tel. +49 30 5019 4236  
[markus.kraemer@htw-berlin.de](mailto:markus.kraemer@htw-berlin.de)
- M.Sc. Benjamin Peris  
Hochschule für Technik und Wirtschaft Berlin  
Tel. +49 30 5019 3630  
[benjamin.peris@htw-berlin.de](mailto:benjamin.peris@htw-berlin.de)
- Prof. Dr. Petra Sauer  
Beuth Hochschule für Technik Berlin  
Tel. +49 30 4504 2691  
[sauer@beuth-hochschule.de](mailto:sauer@beuth-hochschule.de)
- Dipl.-Inform. (FH) Frank Herrmann  
Beuth Hochschule für Technik Berlin  
Tel. +49 30 / 4504-3880  
[fherrmann@beuth-hochschule.de](mailto:fherrmann@beuth-hochschule.de)
- Dipl.-Ing. und Dipl.-Kfm. (FH) Karsten Schomaker  
Botanischer Garten und Botanisches Museum  
Berlin-Dahlem  
Freie Universität Berlin  
[k.schomaker@bgbm.org](mailto:k.schomaker@bgbm.org)  
[www.bgbm.org](http://www.bgbm.org)
- Dipl.-Inform. (FH) Thomas Kalweit  
Aviant GmbH  
[thomas.kalweit@aviant.de](mailto:thomas.kalweit@aviant.de)  
<http://www.aviant.de/>
- Holger Fell  
Fell & Kernbach GmbH  
[holger.fell@fell-kernbach.de](mailto:holger.fell@fell-kernbach.de)  
<http://www.fell-kernbach.de/>