

# Enabling Collaborative Information Management on Federated Data Sources

Analysis, Design and Prototypical Implementation for MS-SharePoint and MS-Exchange

## Referenten

Thomas Reschenhofer  
Bernhard Walzl

## Betreuer

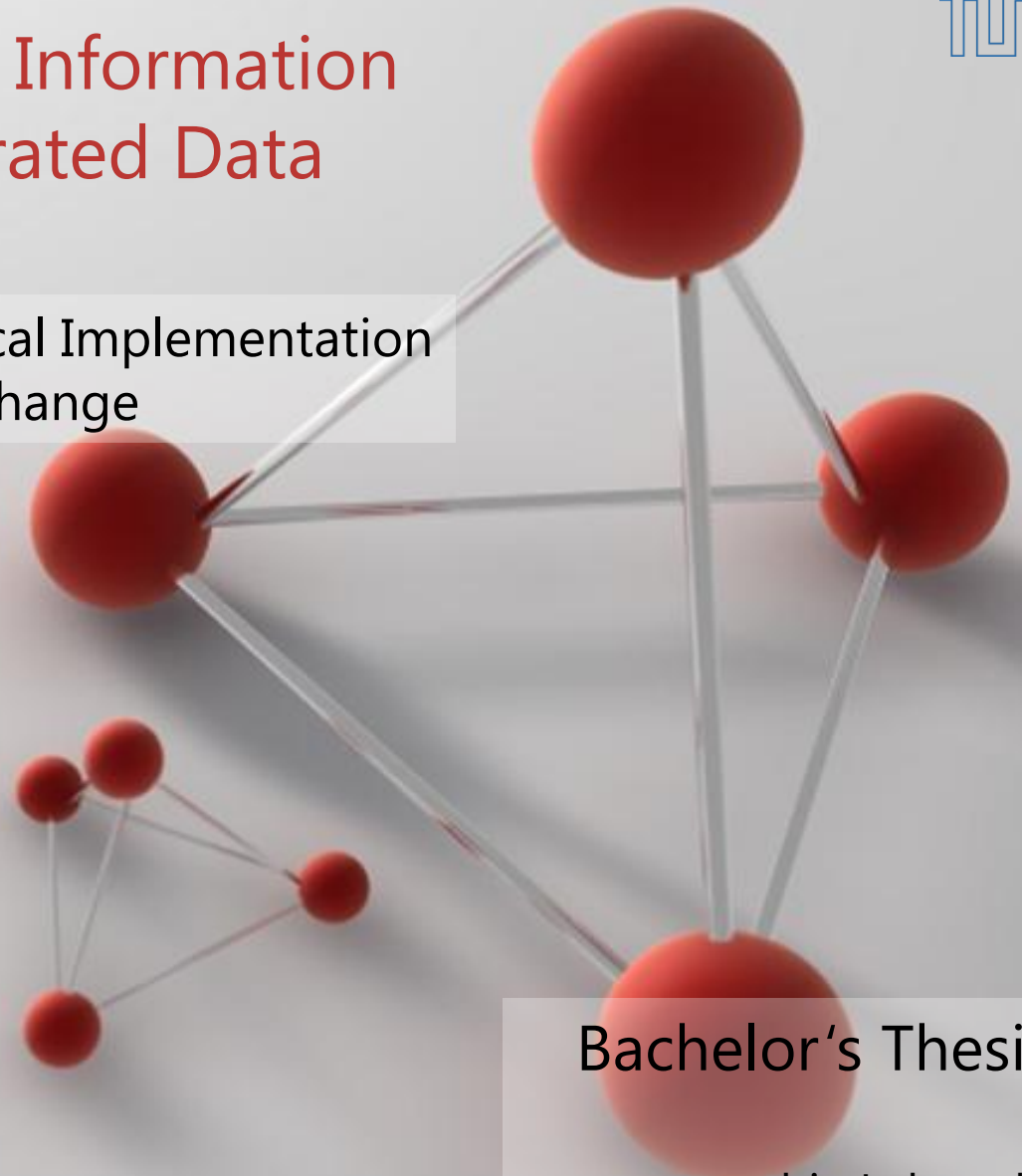
Christian Neubert  
Dr. Thomas Büchner

## Aufgabensteller

Prof. Dr. Florian Matthes

Bachelor's Thesis

sebis-Lehrstuhl  
Sommersemester 2011  
Abschlussvortrag



# Agenda

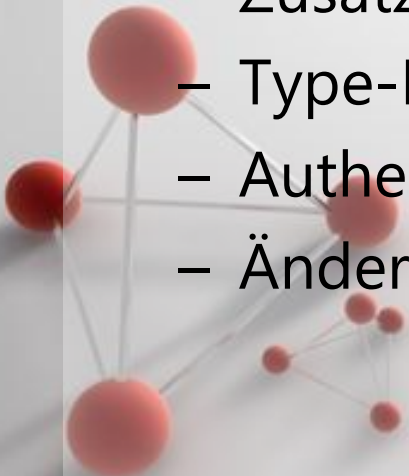
- Motivation und Problemstellung
- Einführung
  - Datenintegration
  - Hybrids
- Systeme
  - MS SharePoint
  - MS Exchange
- Schnittstellen zu externen Datenquellen
- Repräsentation externer Daten
- Fazit

# Motivation und Problemstellung

- Ausgangssituation
  - Viele Informationssysteme/Datenquellen
    - MS SharePoint
    - MS Exchange
  - 1 zentrales System
    - Tricia
- Ziel
  - Integration der Daten in Tricia
- Wieso Datenintegration?
  - Vermeiden von Redundanzen
  - Erleichterter Zugriff auf Information

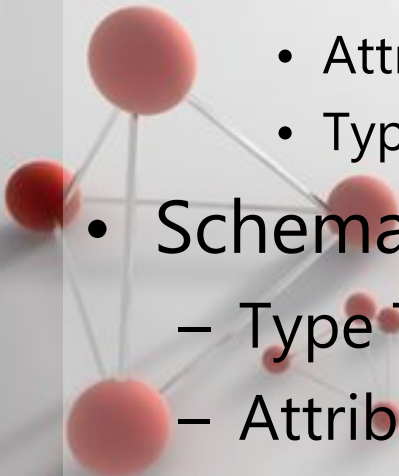
# Datenintegration

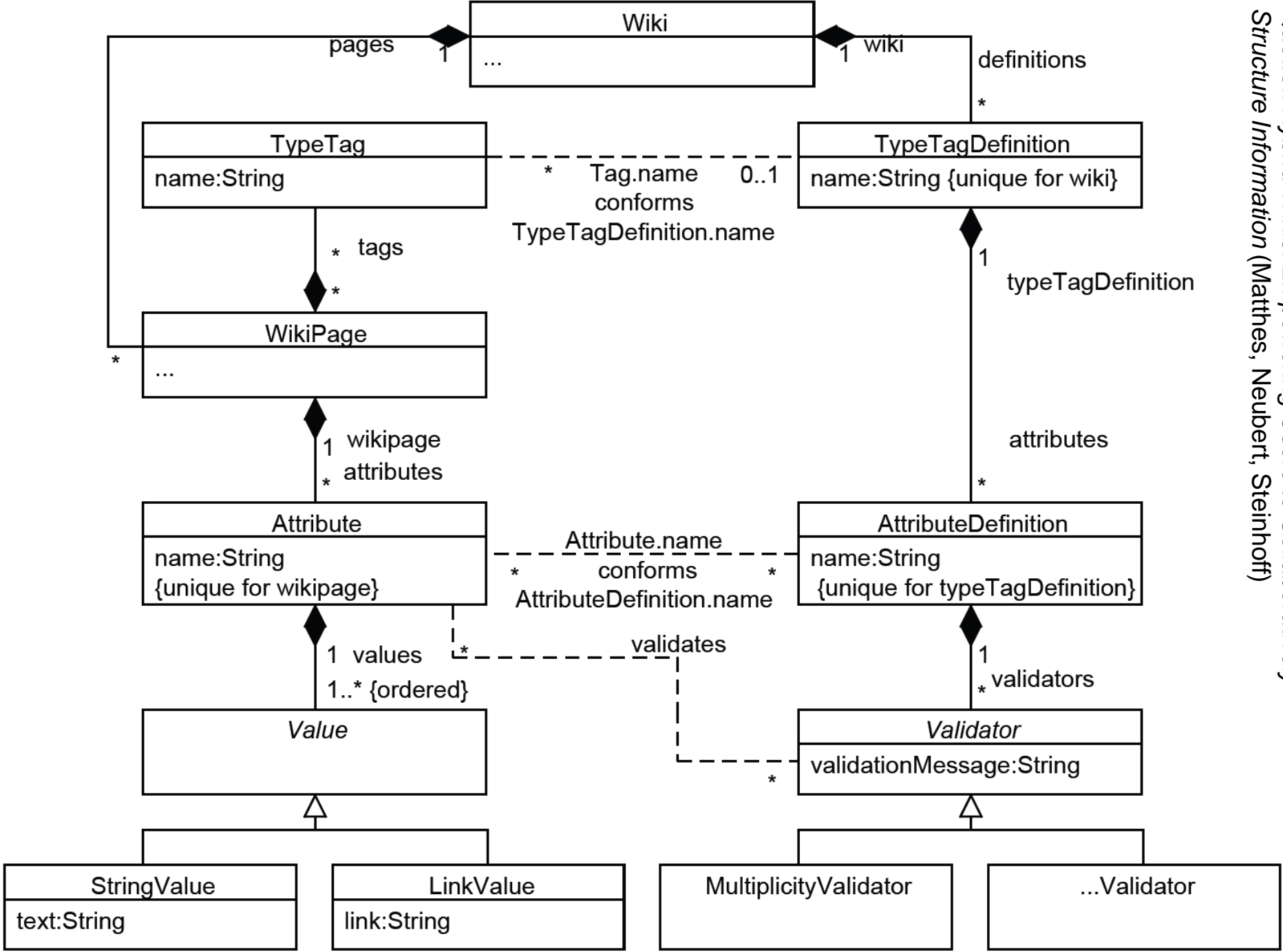
- Förderierte Datenquellen
- Anforderungen
  - Effizientes Importmanagement
  - Flexibilität
  - Benachrichtigungen durch Fremdsysteme
  - Zusätzliche Attribute
  - Type-Mapping
  - Authentifizierung
  - Änderungshistorie



# Hybrids

- „Etwas Gebündeltes bzw. Gekreuztes“
- Strukturieren unstrukturierter Informationsobjekte
  - Unstrukturierte Informationsobjekte
    - z.B. Wikiseiten
  - Strukturierung durch
    - Attribute (Name-Wert-Paar)
    - Typen (als sog. „Type Tag“)
- Schemabeschreibung
  - Type Tag Definition
  - Attribut Definition





# Microsoft SharePoint


- SharePoint
  - ECM-System von Microsoft
  - Enthält Listen und Bibliotheken
  - OData-Producer ab Version 2010




- OData (Open Data Protocol)
  - Webservice Protokoll
  - Baut auf HTTP, AtomPub und JSON auf
  - Ermöglicht CRUD-Operationen auf Entitäten
  - Client Bibliotheken für Java, JavaScript, .NET, ...




# Microsoft SharePoint - Beispiel

Site Actions 
System Account ▾

List Tools
Browse


Bachelor's Thesis ▸ People ▸ All contacts ▾

Bachelor's Thesis
Search this site... 

**Libraries**


Shared Documents


**Lists**



People


Sample ODATA Queries

---

 Recycle Bin

 All Site Content

| Last Name   | First Name | Job Title                 |
|---|------------|---------------------------|
| <a href="#">Buechner</a>  | Thomas     | Senior Research Assistant |
| <a href="#">Matthes</a>  | Florian    | Full Professor            |
| <a href="#">Neubert</a>   | Christian  | Research Assistant        |
| <a href="#">Reschenhofer</a>  | Thomas     | Student                   |
| <a href="#">Waltl</a>  | Bernhard   | Student                   |

 [Add new item](#)



# Microsoft SharePoint - Beispiel

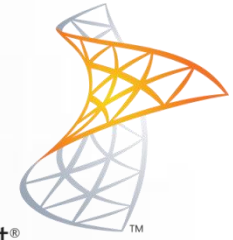
```

- <feed xml:base="http://sharepoint2010:1234/_vti_bin/listdata.svc/"
  xmlns:d="http://schemas.microsoft.com/ado/2007/08/dataservices"
  xmlns:m="http://schemas.microsoft.com/ado/2007/08/dataservices/metadata" xmlns="http://www.w3.org/2005/Atom">
  <title type="text">People</title>
  <id>http://sharepoint2010:1234/_vti_bin/listdata.svc/People</id>
  <updated>2011-06-22T21:46:15Z</updated>
  <link rel="self" title="People" href="People" />
   <entry m:etag="W/"1"">
    <id>http://sharepoint2010:1234/_vti_bin/listdata.svc/People(3)</id>
    <title type="text">Neubert</title>
    <updated>2011-05-25T10:49:10+02:00</updated>
    - <author>
      <name />
    </author>
    <link rel="edit" title="PeopleItem" href="People(3)" />
    <category term="Microsoft.SharePoint.DataService.PeopleItem"
      scheme="http://schemas.microsoft.com/ado/2007/08/dataservices/scheme" />
    - <content type="application/xml">
      - <m:properties>
        <d:LastName>Neubert</d:LastName>
        <d:FirstName>Christian</d:FirstName>
        <d:JobTitle>Research Assistant</d:JobTitle>
      </m:properties>
    </content>
  </entry>
  + <entry m:etag="W/"1"">
  + <entry m:etag="W/"17"">
  + <entry m:etag="W/"1"">
  + <entry m:etag="W/"1"">
</feed>

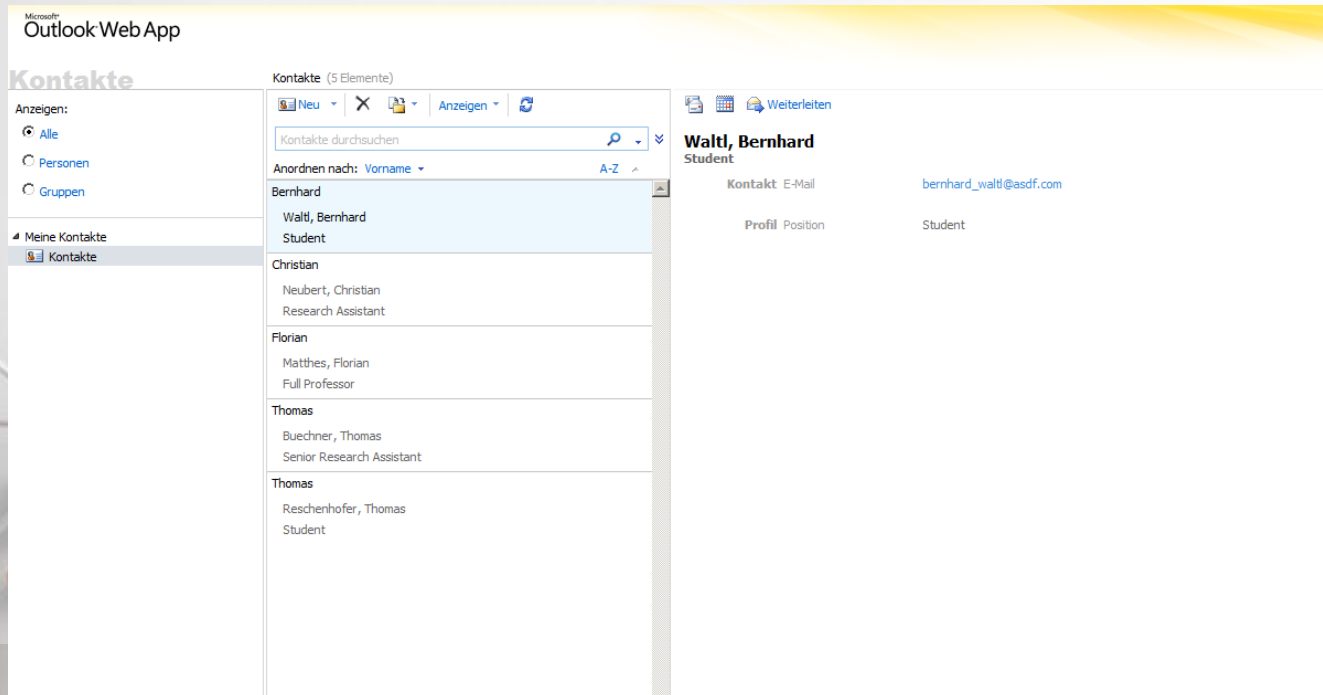
```

# Microsoft Exchange

- Funktionsumfang
  - E-Mail
  - Kontakte und Adressen verwalten
  - ...

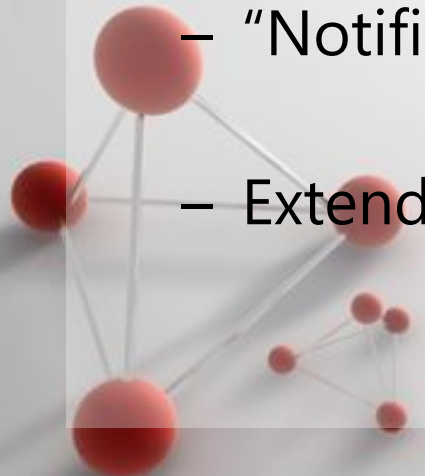


Microsoft®  
**Exchange**



# Microsoft Exchange

- Zugriff über Webservices
  - Geeignete API
  - Abrufen und Importieren von Informationen
    - insbesondere Kontaktdaten und E-Mails
  - “Notificationservice” von MS Exchange
  - Extended Properties möglich

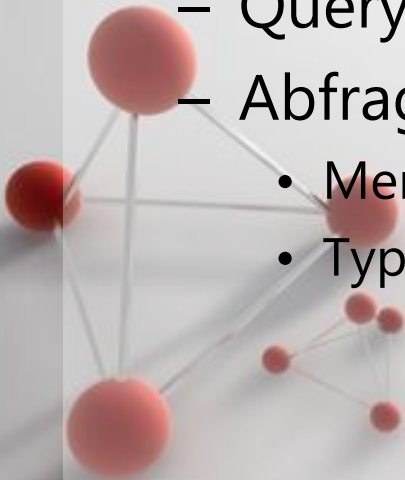


## Gesamtarchitektur – 2 Schichten

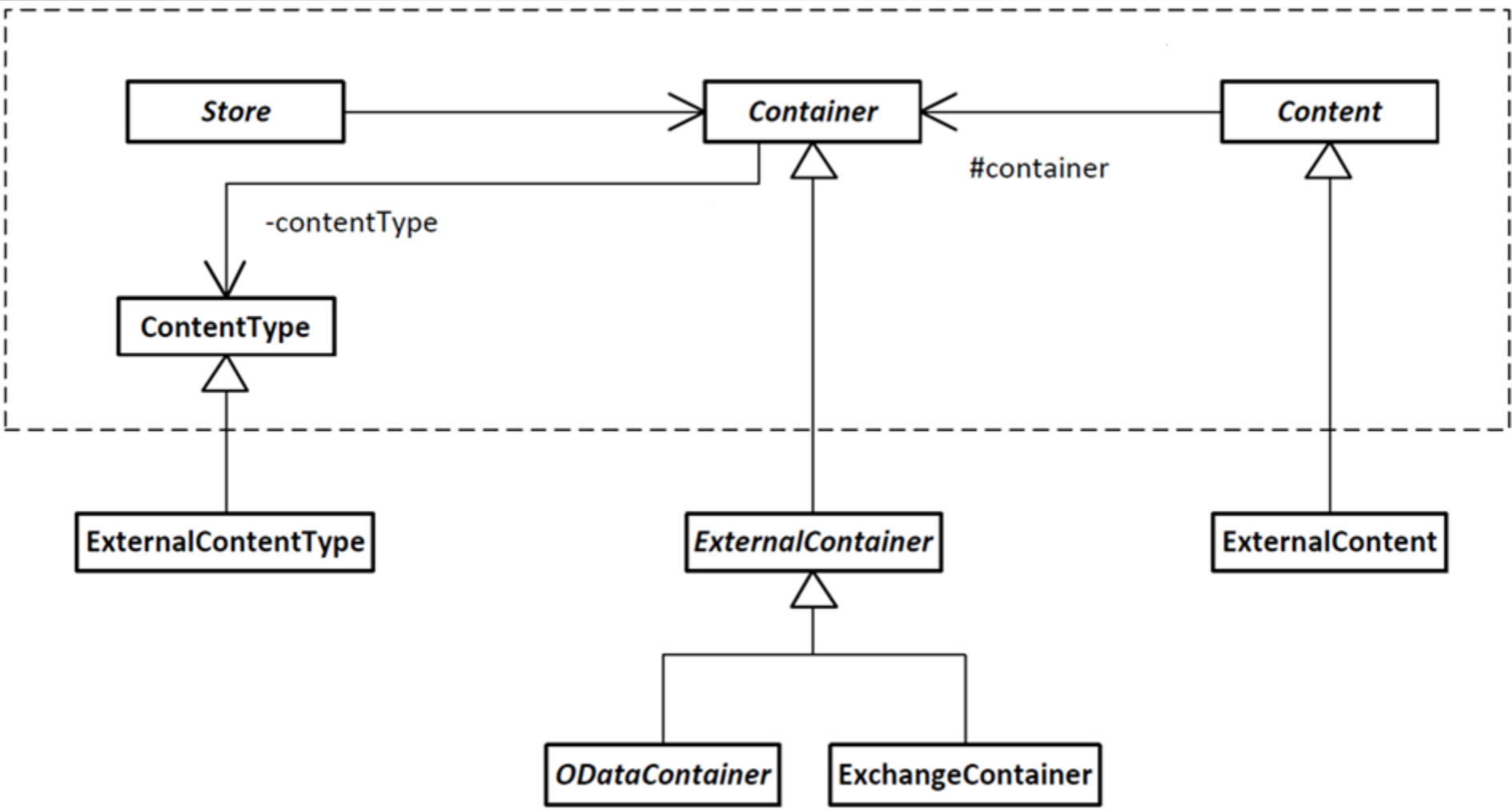
- Schicht I: Generische Schnittstellen zu Datenquellen
  - Store-Ebene
  - Schnittstelle zu MS SharePoint
  - Schnittstelle zu MS Exchange
- Schicht II: Repräsentation externer Daten im Datenmodell
  - Repräsentation des Schemas
  - Repräsentation eines Datensatzes
  - User Interface

## Schicht I: Generische Schnittstelle

- Kommunikation mit Datenquelle
  - CRUD-Methoden
    - Create
    - Read
    - Update
    - Delete
  - Query (Filter)
  - Abfrage des Schemas
    - Menge der Attribute
    - Typen der Attribute

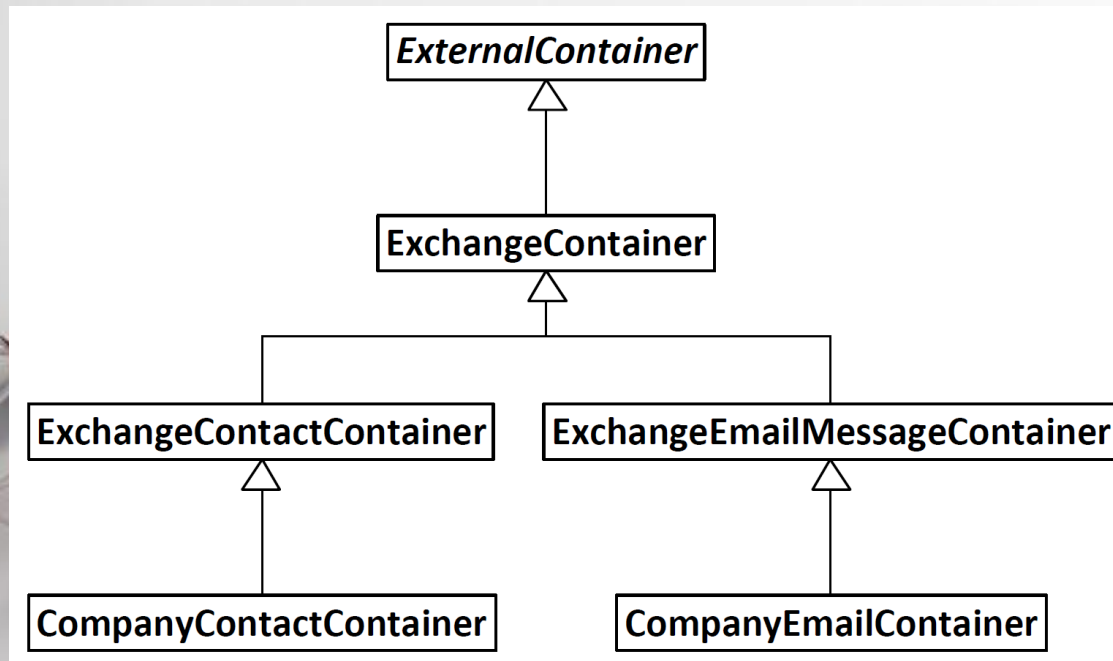


# Erweiterung auf Store-Ebene



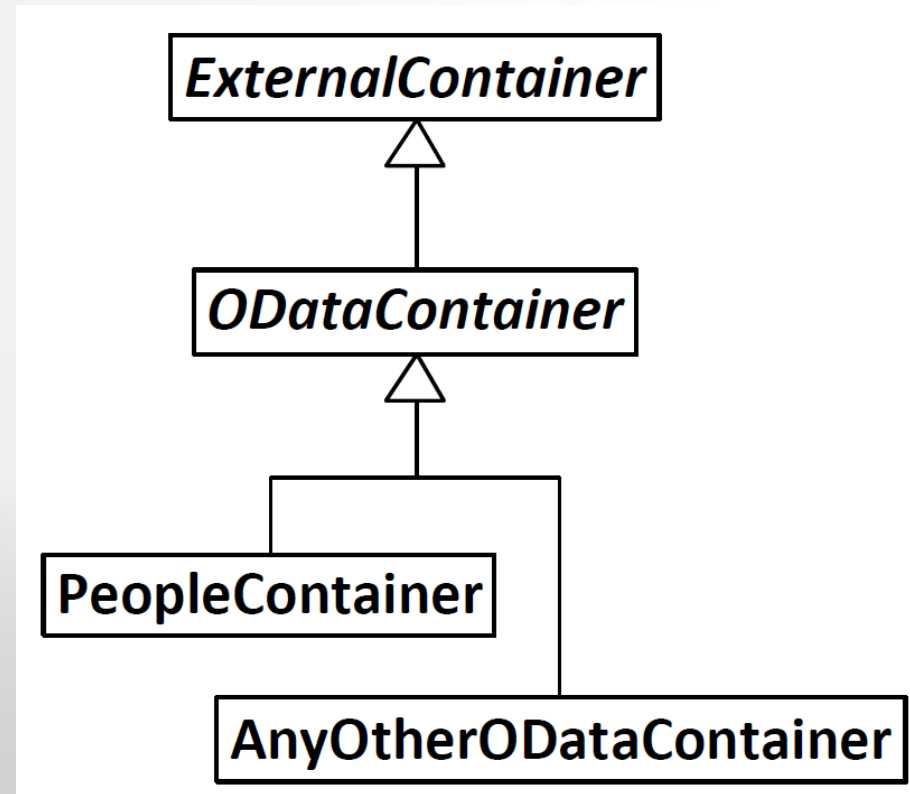
## Schnittstelle zu MS Exchange

- Zugriff über Webservices: EWS Java API
- Ein Itemtyp (Kontakt, Email, ...) pro Containerklasse



## Schnittstelle zu MS SharePoint

- Spezialisierung einer OData-Schnittstelle
  - Client-Framework: odata4j
  - Beliebige OData-Producer integrierbar
- ODataContainer
  - Implementiert CRUD-Methoden
  - Benötigt Service-URL





## Schicht II: Repräsentation externer Daten – Evaluierung (1/2)

### Built-In-Typen

- Darstellung als Ausprägung eines Built-In-Typen
  - Attribute als Membervariablen
- Datentypen
  - String, Date, Boolean, Integer, TimeStamp, DateTime, Email, ...

### Hybrids

- Darstellung als Hybrid
  - Attribute als HybridProperties
- Datentypen
  - Nur Typvalidatoren
  - Boolean, Date, Enumeration, Link

## Schicht II: Repräsentation externer Daten – Evaluierung (2/2)

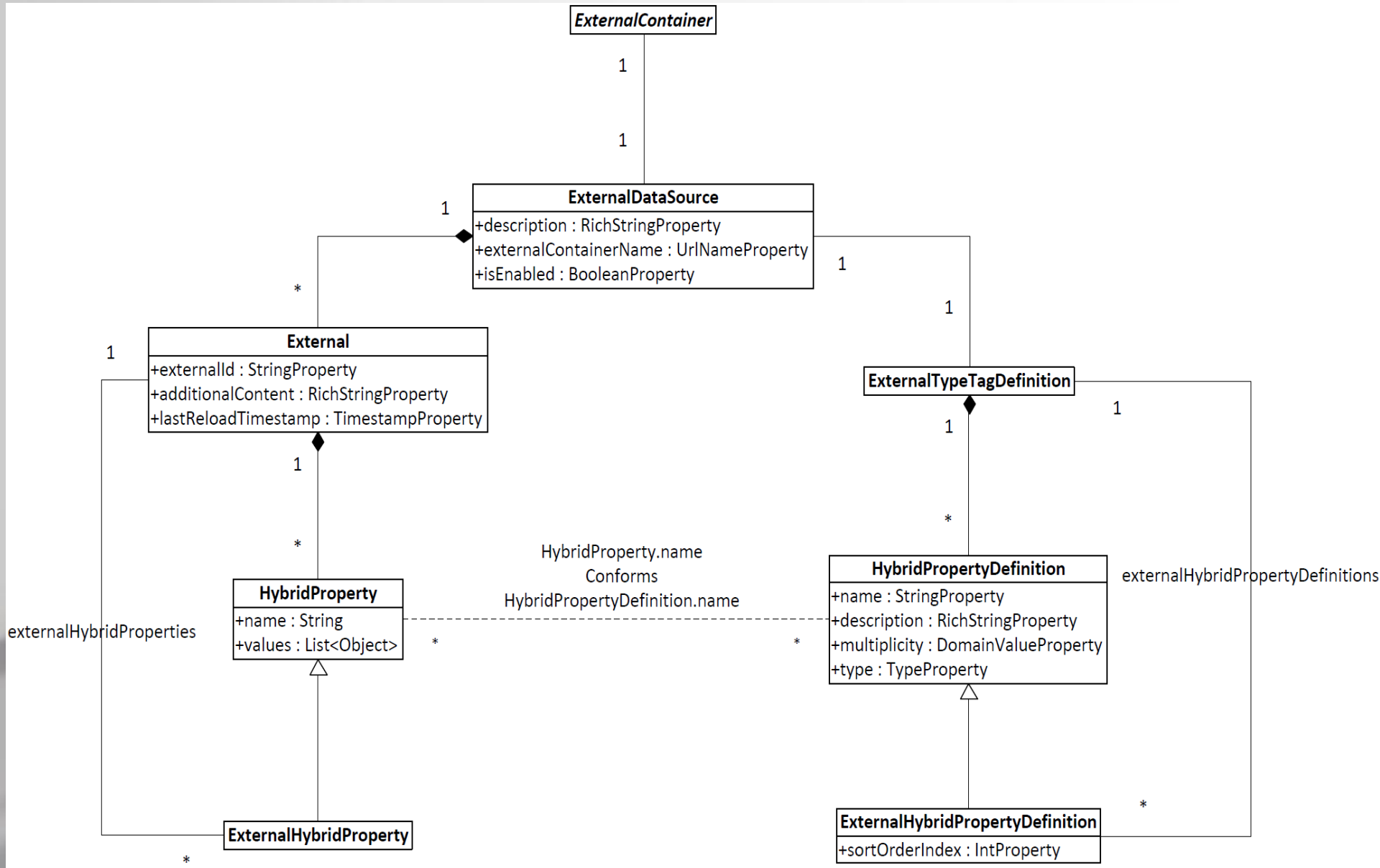
### Built-In-Typen

- Flexibilität
  - Schemaänderung zur Laufzeit nicht möglich
  - Bei neuer Datenquelle neue Entität
- GUI
  - Hoher Aufwand
  - Für jede Entität eigene View notwendig

### Hybrids

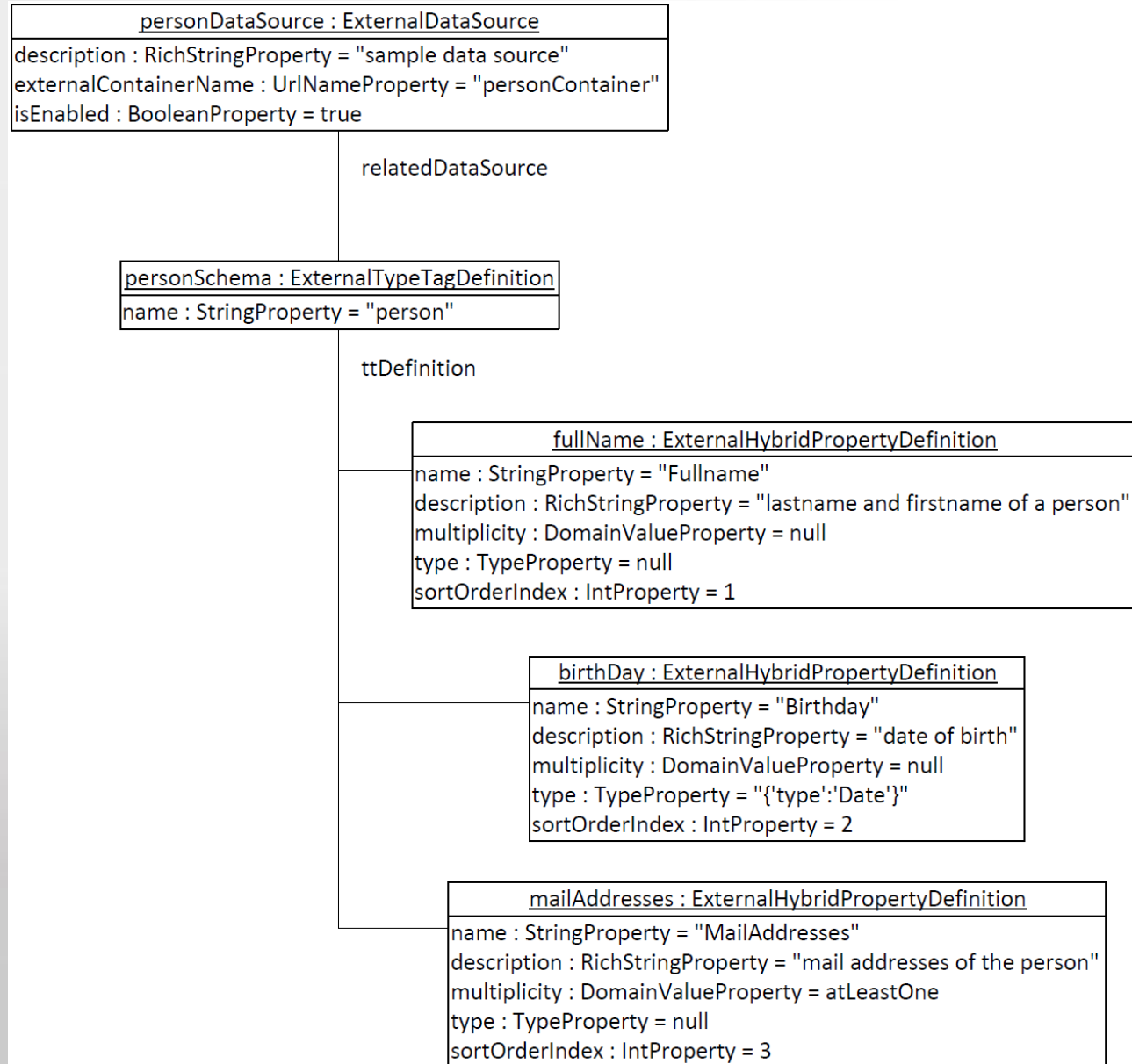
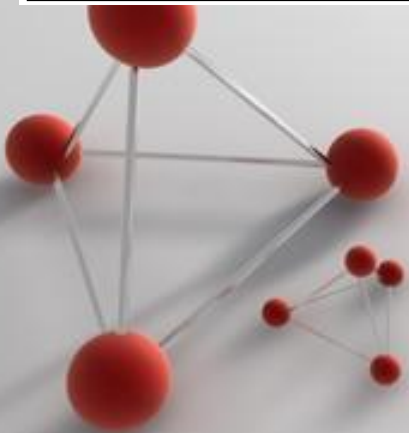
- Flexibilität
  - Schemaänderung jederzeit möglich
  - Kein Aufwand bei neuer Datenquelle
- GUI
  - Geringer Aufwand
  - Verwenden der Hybrid Wiki Funktionalität

# External Modell



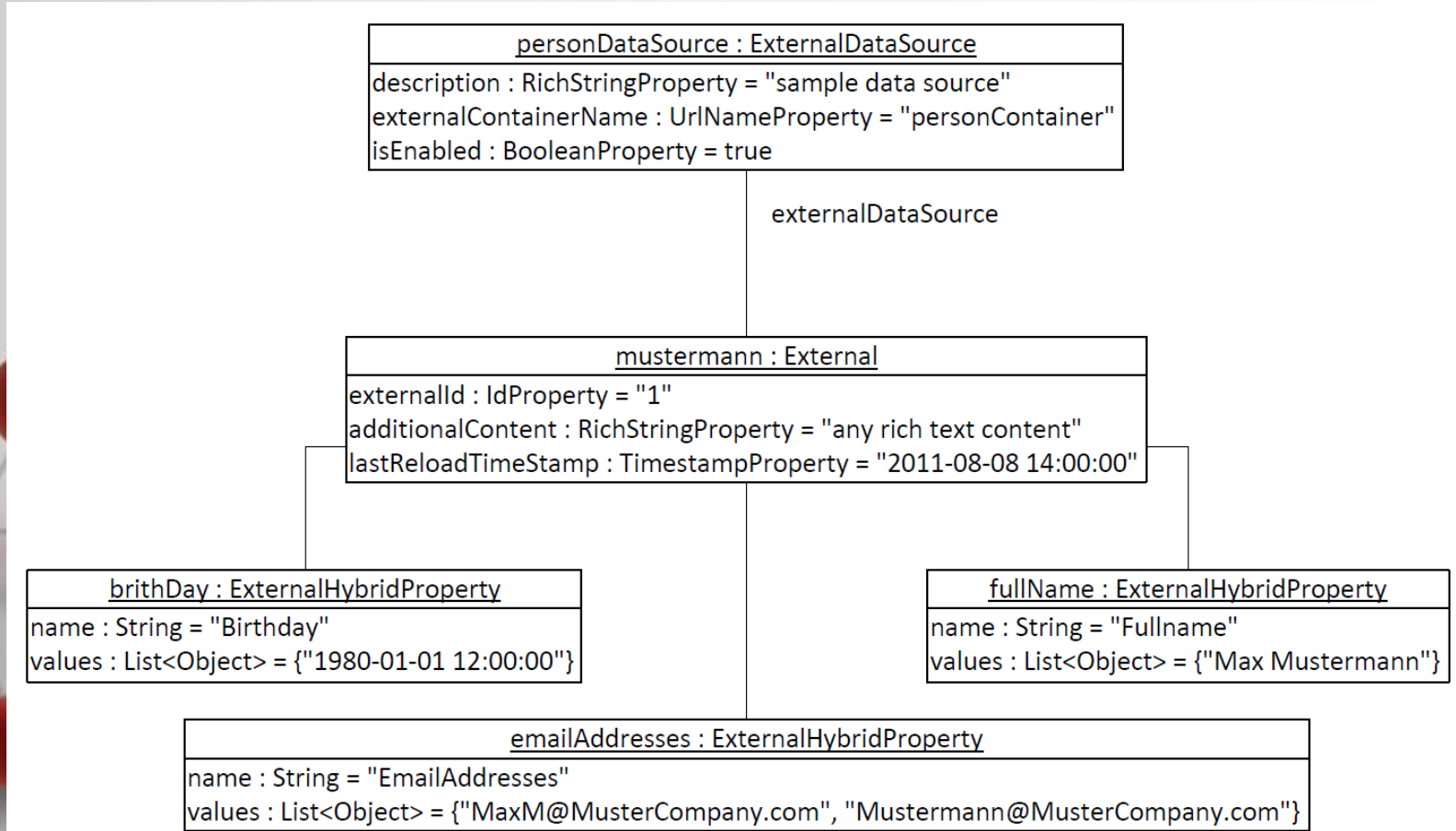
# Darstellung des Schemas - Beispiel

| Person                    |
|---------------------------|
| +Fullname : String        |
| +Birthday : Date          |
| +MailAddresses : String[] |



# Darstellung eines Datensatzes - Beispiel

| <u>mustermann : Person</u>  |
|---|
| Fullname : String = "Max Mustermann"  |
| Birthday : Date = "1980-01-01 12:00:00"   |
| MailAddresses : String[] = {"MaxM@MusterCompany.com", "Mustermann@MusterCompany.com"} |




## Laden externer Daten


- Laden aller externen Datenquellen
  - Laden des Schemas
    - Nur Attribute, welche in Tricia verwendet/angezeigt werden
  - Laden aller Items aus externen Systemen
    - Laden aller durch das Schema vorgegebenen Attribute
- Zyklisch
- Bei Benachrichtigung
- Manuell



# Ergebnis - Livedemo

**infoAsset**
bernhard


---

 » Search

**Contents matching 'thoma'**

▼ Last modification

Any Date

▼ Content type

no label provided. (4)

▼ Types

exchange contacts (2)




sharepoint contact (2)

▼ Special

Contains Invalid Links

Contains Invalid Values





Search results for 'bernhard':

-  **Waltl Bernhard**  
 from 'Exchange Contacts'
-  **Waltl Bernhard**  
 from 'SharePoint People'
-  Hello world!  
 from 'Exchange Emails'


[exchange contacts](#) [sharepoint contact](#)

---

Results 1 - 4 of 4

-  **Reschenhofer Thomas**  
 DateTimeReceived DateTimeSent Department DisplayCc DisplayName **Thomas** Reschenhofer DisplayTo  
[exchange contacts](#)
-  **Buechner Thomas**  
 DateTimeReceived DateTimeSent Department DisplayCc DisplayName **Thomas** Buechner DisplayTo  
[exchange contacts](#)
-  **Reschenhofer Thomas**  
 08:21:00 GMT CreatedById 1073741823 EMailAddress **thomas.reschenhofer@mytum.de** FaxNumber FirstName  
[sharepoint contact](#)
-  **Buechner Thomas**  
 :56:45 GMT CreatedById 1073741823 EMailAddress FaxNumber FirstName **Thomas** FullName HomePhone  
[sharepoint contact](#)

1 (total result pages: 1)



Tag Filter

▶ Attribute Filter

## Ausblick / Erweiterungsmöglichkeiten

- Single Sign-On
- Berechtigungsverwaltung
- Erzeugen und Löschen externer Datensätze
- Erzeugen und Löschen externer Attribute
- No-Code-Erweiterbarkeit
- Benutzerfreundliche Attributnamen
- Datenquelle-abhängige Aktionen und Views





# Enabling Collaborative Information Management on Federated Data Sources



Diskussion und  
Fragen