

# Das 'Vocabulaire Internationale de Diplomatie': Perspektiven domänenspezifischer Vokabularien

Georg Vogeler

Zentrum für Informationsmodellierung in den  
Geisteswissenschaften der Universität Graz  
Institut für Dokumentologie und Editorik

<ZIM<sup>ig</sup>/>



# VOCABULAIRE INTERNATIONAL DE LA DIPLOMATIQUE

# Vocabulaire Internationale de Diplomatique (VID) 1980->2004



# Vocabulaire Internationale de Diplomatie

- Commission Internationale de Diplomatie arbeitete am Vocabulaire seit 1980
- Ziele
  - Normalisierung der Begrifflichkeit
  - Verständnis fremdsprachiger Texte
  - Sicherstellen richtiger Nutzung der Produkte der Diplomaten durch Dritte.
- Erschienen 1. Aufl. 1994, 2. Aufl. 1997.
- Im Folgenden abgekürzt als VID

# Ressource zur Erschließung mittelalterlicher Urkunden?

- VID als klassisches kontrolliertes Vokabular
- Kritik
  - Nicht immer die besten Begriffe
  - Schlechte Berücksichtigung der deutschen Diplomatie (i.e. Urkunden, die im deutschen Sprachraum entstanden sind)

# Digitalisierungsstrategie

- Primär nicht Text (z.B. Scan > TEI)
- Sondern Datenbank
  - abstraktes Konzept mit Identifikator und Definition
  - Bezeichner in verschiedenen Sprachen
- Basis für derzeitige Volltextansicht:  
<http://www.cei.lmu.de/VID/VID.php>

# Konversionen

- TEI <<http://www.cei.lmu.de/VID/tei/>>
  - <div> mit <desc> und <term>
  - <list> mit <label> und <item>
  - <taxonomy>
- SKOS <<http://www.cei.lmu.de/VID/skos>>
  - turtle
  - RDF

# Vorgehen der Konversion: Automatisch – händisch

- OCR
  - mit Nachkorrektur
- Datenbankstruktur
  - Skripting und von Hand
- Binnenverweise
  - über reguläre Ausdrücke und von Hand

# Anwendungsszenarien

1. Erschließung von Urkunden
2. Integration in existierende SKOS-Ressourcen
3. Referenz für Datenaustausch
4. Schließen

# Erschließung von Urkunden

- Beschlagwortung von Urkundentypen, z.B. „cédule“
- **VID No. 411:** Une cédule est:
  - \_ a. Soit un \*billet par lequel une autorité fait connaître sa volonté à un de ses dépendants (...)
  - \_ b. Soit, émanée d'un magistrat de l'ordre judiciaire pour un justiciable, une cédule de citation, d'évocation, etc. (...)
- **VID No. 412:** Types de documents dits "cédules" dans diverses chancelleries:
  - \_ a. A la Curie pontificale, les cédules consistoriales (lat.: *cedulae consistoriales*) (...)
  - \_ b. En Angleterre, la *schedule* est, en général, un parchemin complétant un autre document (...)
  - \_ c. En Espagne, la *Real cédula* est un document abrégé où les souscriptions sont indiquées (...)
  - \_ d. De la *scheda* ou *schedula* dérive le sens de \*minute, à l'époque carolingienne (...).

⇒ **Cédule\_VID\_412Ad != Cédule\_VID\_412Ab**

# Integration in existierende SKOS-Ressourcen

- English Heritage: MDA Object Type Thesaurus  
des National Monuments Record Centre

<[http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes\\_no=144](http://thesaurus.english-heritage.org.uk/thesaurus.asp?thes_no=144)>

z.B. Written Communications > Written Item >  
**Document** > Bulla

- UNESCO-Thesaurus

<<http://www2.ulcc.ac.uk/unesco>>

z.B. MT 5.30 Information sources > **Documents** >  
Primary Documents, Manuscripts

# Referenz für Datenaustausch

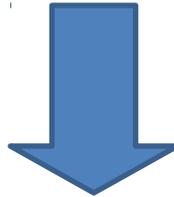
- Urkundendigitalisierungsprojekte
  - monasterium.net: CEI
  - Codice diplomatica della Lombardia medievale: Eigenes Encoding
  - Virtuelles Deutsches Urkundennetzwerk: EAD
  - Acte originaux de France etc.: TEI
- `<tei:author>` = `<cei:issuer>` = `<cdlm:rogatario>`  
= `<ead:persname role="vid:VID_16">`

# Semantic Web

```
SELECT ?urkunde ?notariusPP ?scribe count(?urkunde) as
numOfUrkunde
WHERE {
    ?urkunde vid:VID_3 regpont:charter ;
        ?vid_schreiber ?scribe ;
        vid:VID_286b ?notariusPP .
    OPTIONAL { ?vid_schreiber skos:narrower vid:VID_274 . }
    OPTIONAL { ?vid_schreiber a vid:VID_274 . }
    GRAPH ?s { ?scribe a regpont:person . }
    GRAPH ?n { ?notariusPP a regpont:person . }
    FILTER (?s != ?n)}
GROUP BY ?notariusPP ?scribe
HAVING count(?urkunde) = 1
```

# Schließen

class(vid:272\_redacteur) vid:343\_dictare class(vid:3\_acte\_écrits)  
„Reinald von Dassel“ vid:343\_dictare mgh:DLD\_125



mgh:DLD\_125 isInstanceOf vid:3\_acte\_écrits  
„Rainald von Dassel“ isInstanceOf vid:272\_redacteur  
pnd:118597922 dc:creator mgh:DLD\_125

# Domänenspezifische Vokabularien in der Datenaggregation

- CEI/TEI@VID/EAD@VID nach VID
  - cei:issuer => vid\_16 (Aussteller)
  - ead:unitid => vid\_33 (Archivsignatur)
- VID nach Dublin Core
  - vid\_16 (Aussteller) => dc:creator
  - vid\_272 (Konzipient), vid\_273 (Kanzleischreiber), vid\_274 (Schreiber), vid\_275 (ingrossator) ... => dc:contributor
  - vid\_20 (Regest) => dc:title
  - vid\_560 (Ausstellungsdatum) => dc:date

# Fazit

- Domänenspezifische Vokabularien
  - ergänzen übergreifende Vokabularien
  - dienen dem Datenaustausch zwischen Projekten
  - ermöglichen formalisierte Definitionen
- Domänenspezifische Vokabularien ermöglichen es, wissenschaftliche Detailbeschreibungen in übergreifende Such- und Auswertungskontexte zu übernehmen.