

---

# Potential(e) hyperdiplomatischer Transkriptionen deutschsprachiger mittelalterlicher Texte

(Projekt CoReMA –  
Cooking Recipes of the Middle Ages, fwf I-3614)  
Astrid Böhm / Helmut W. Klug / Christian Steiner

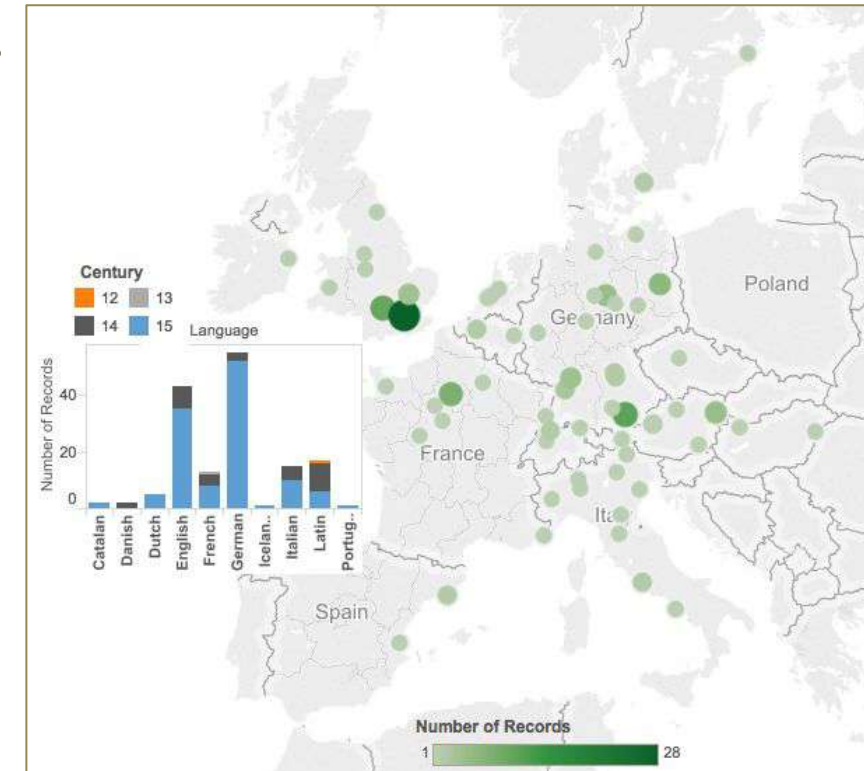
---

# Aviso

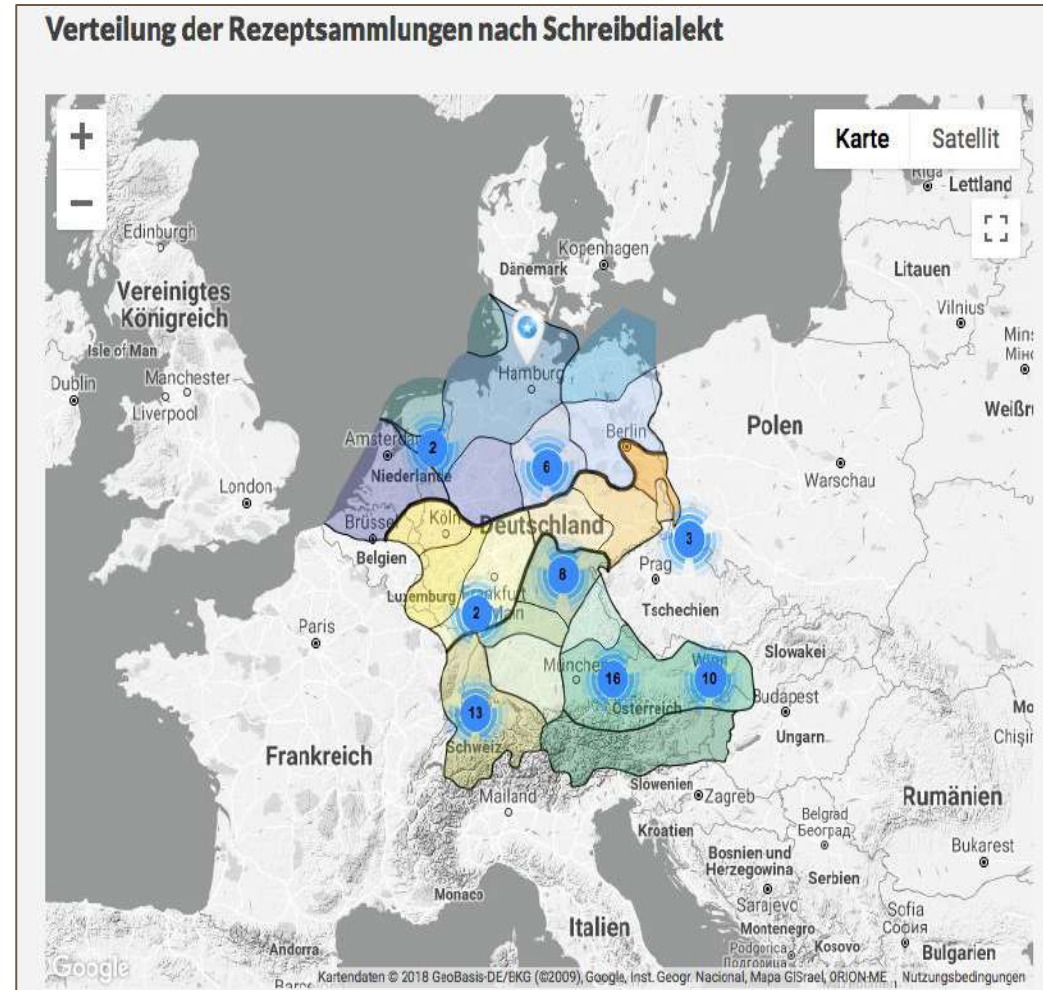
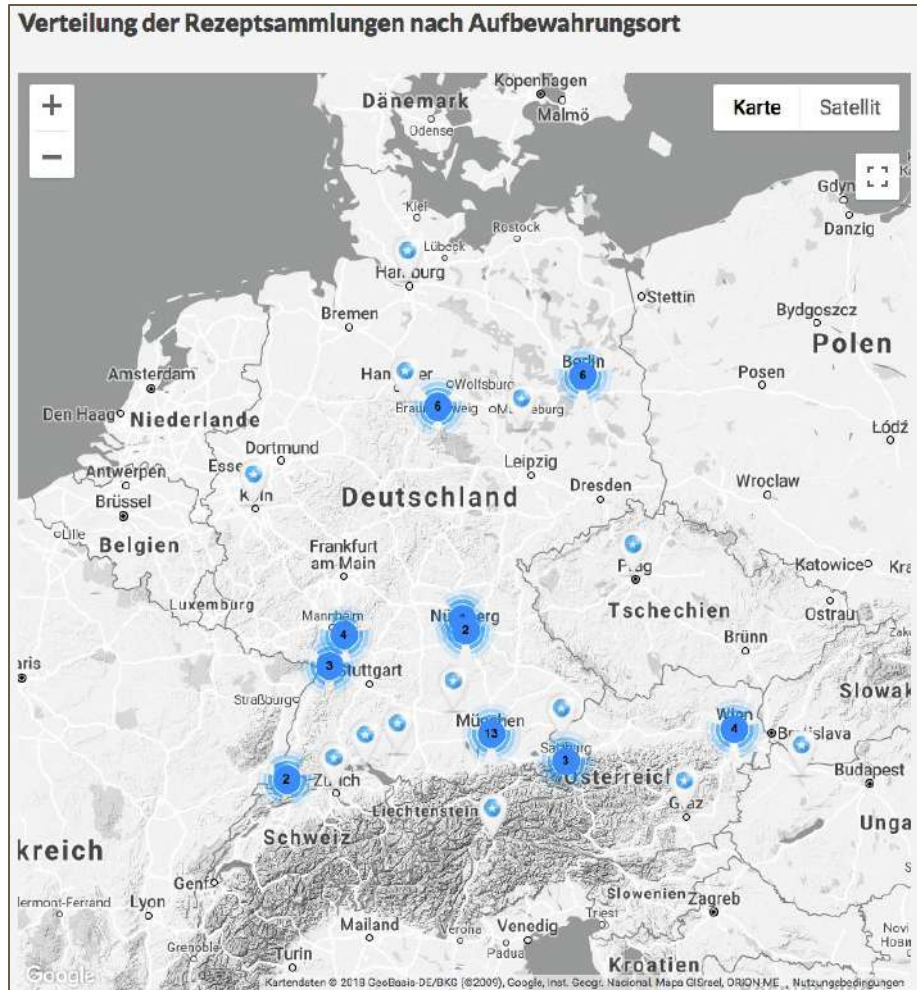
- Kontextualisierung: Projektvorstellung
- Überblick: Transkriptionsworkflow und Ergebnisse
- Theoretische Verankerung
- Potentiale

# CoReMA – Cooking Recipes of the Middle Ages Corpus, Analysis, Visualisation

- Kochrezeptüberlieferung mlat., afrz., fnhdt.
- Handschriftliche Überlieferung bis 1500
- Ca. 80 Manuskripte mit ca. 8000 Rezepten
  - 60 dtspr. Sammlungen



# Dtspr. Rezeptsammlungen



# Analyse

- Aufarbeitung der Texte
    - Erst-Transkription sowie Neu-Transkription bestehender Editionen
    - Modellierung in XML/TEI
    - Metadaten (kodikologisch, philologisch)
  - Annotation (Zutaten, Kochanweisungen, ...)
  - Semantic Web-Technologien
    - Ontologie
- ⇒ **Sprachübergreifender Vergleich**
- Essgewohnheiten, Einflüsse, Migrationsbewegungen

# Visualisierung

- GAMS
- Bild-Text-Verknüpfung
- Räumliche und zeitliche Visualisierung
- Statistische Auswertungen
- Ergebnisse



# Ergebnisse

- Open access (CC-BY)
  - Langzeitarchivierung GAMS
  - Weiternutzung der gewonnenen Daten
  - Erweiterung andere Sprachen, andere Perioden
  - Eröffnung neuer Forschungsfragen
    - Paläographie, historische Graphematik, Historiolinguistik, Kulturgeschichte, Kulinarhistorik, Wirtschafts- und Sozialgeschichte ...
- ⇒ **Transkriptions- und Publikationsworkflow**

# Workflow: Quellenerfassung

- Transkription und Basiskollationierung mit Transkribus
- XSL-Transformationen des Transkribus Exports (TEI-XML) in div. Formate
  - HTML zur Kollationierung
  - XML-Basisfassung (diplomatisch)
  - normalisierte Textfassung (Basis für semant. Enrichment)
- manuelle Korrekturdurchgänge
  - Silbentrennung <w>
  - Stellenkommentar <note>
  - Mark-up von Texteinheiten <seg>
- Ergänzung um Metadaten
  - <msDesc>
- Semi-automatische semantische Annotation
  - Phrasen <title>, <opener>, <instruction>, ...
  - Named Entities <ingredient>, <dish>, <tool>, ...



# Ergebnis(se)

- hyperdiplomatische Textfassung
  - <abbr>, <g> (XML-Download)

```
<pb xml:id="MS1609_161v" n="161v"/>

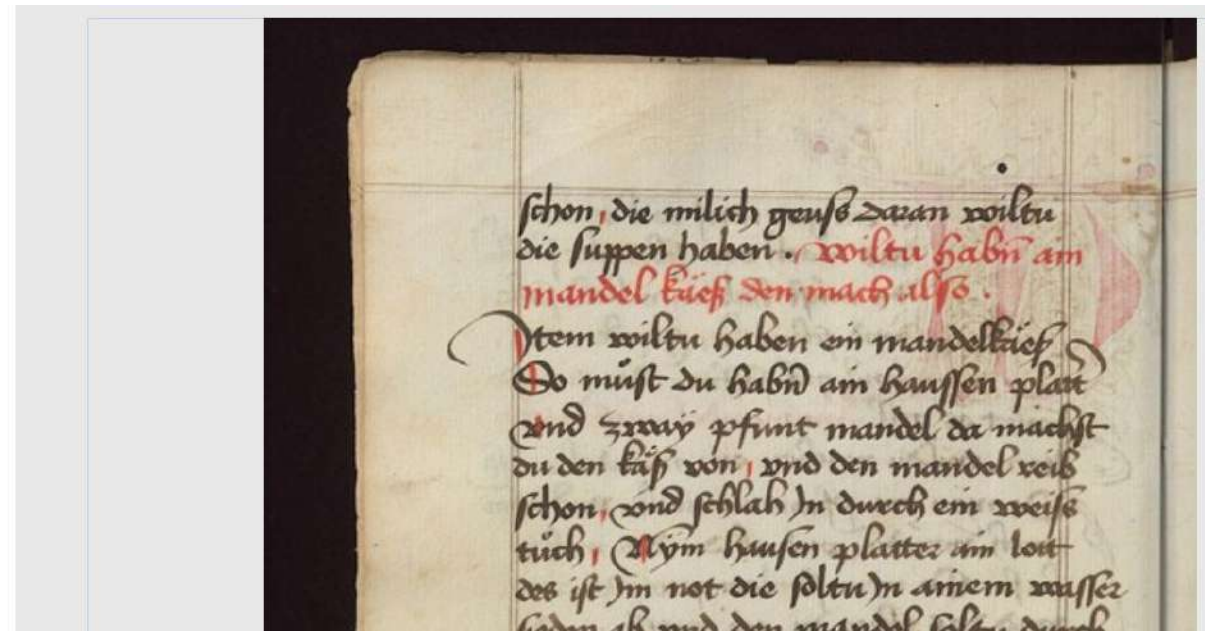
<lb n="N001"/><hi rend="textColour:RED;"><g ref="#verbar"/></hi>Swarcz vnd velt darnach hera
<lb n="N002"/>padenn vnczt das dy<g ref="#combuml"/> wa<g ref="#combuml"/>rcz heraus velt <g
<lb n="N003"/><hi rend="textColour:RED;"><foreign>Aqua que d<g ref="#inodot">i</g>c<g ref="#
<lb n="N004"/><hi rend="textColour:BLACK;"><g ref="#verbar"/></hi>Das gulde<g ref="#inodot">
<lb n="N005"/>mu<g ref="#slong">s</g>capluet <hi rend="textColour:RED;"><g ref="#fracsol"/><
<lb n="N006"/>ras <abbr>korn<ex>er</ex><am><g ref="#combccomma_er"/></am></abbr> <hi rend="te
<lb n="N007"/>das alles <mod ana="TR.RP.DL.WR.ST.IN.MM.MG.LF"><del><g ref="#inodot">i</g>n c
<lb n="N008"/>weins daran als des gewurcz i<g ref="#slong">s</g>t <abbr>v<del><ex>er</ex><am>
<lb n="N009"/><note>Lat.-dt. Rezept am linken Rand auf Höhe der Zeilen 8-15 eingetragen: <fo
<lb n="N010"/>vnd las es also <g ref="#slong">s</g>tenn pis an den xv tag an dem <abbr><g re
<lb n="N011"/>tag nym das heraus vnd seich den wein darab <abbr>du<ex>r</ex><am><g ref="#com
<lb n="N012"/><g ref="#slong">s</g>chon tuech vnd <g ref="#slong">s</g>tos gewurcz chlain <g
<lb n="N013"/>ay<g ref="#combuml"/>n mu<g ref="#combuml"/>es vnd leg es dan <abbr>w<g ref="#
<lb n="N014"/><hi rend="textColour:RED;"><g ref="#verbar"/></hi>dan das gewurcz vnd wein aus
<lb n="N015"/>las es aus trieffen <g ref="#inodot">i</g>n ein glas mit <abbr>l<g ref="#inodo
<lb n="N016"/>n<g ref="#inodot">i</g>cht da uon mu<g ref="#combuml"/>g <hi rend="textColour:
<lb n="N017"/>man <hi rend="textColour:RED;"><g ref="#verbar"/></hi>In legenn <hi rend="text
<lb n="N018"/>aus getrouft <hi rend="textColour:RED;"><g ref="#verbar"/></hi>So gehalt den
```

# Ergebnis(se)

- editorische Fassung:
  - normalisiertes ASCII Zeicheninventar + editorische Eingriffe: Abkürzungen, Revisionen, Rubrizierung, <note> (Online in Text-Bild-Synopse)

[17v]

- 01: schon die milich geuss daran wiltu  
 02: die suppen haben . **wiltu haben ain**  
**03: mandel käesz den mach also .**  
 04: Item wiltu haben ein mandelkaeesz  
 05: So müst du haben ain haussen platter  
 06: vnd zway pfunt mandel da machst  
 07: du den kaesz von vnd den mandel reib



# Ergebnis(se)

- normalisierte Fassung
  - normalisiertes ASCII Zeicheninventar, ohne: Abkürzungen, Revisionen (Textbasis für semantische Annotation)

```

<pb xml:id="MS1609_011r" n="011r"/>
<ab type="recipe" xml:lang="deu-enh" n="1" xml:id="MS1609.1"><opener>Item ildw
  machen ein <title>guet <ingredient en="egg" commodity="Q93189"
    >air</ingredient> von <ingredient en="almond" commodity="Q39918"
    >mandel</ingredient></title></opener> so <instruction>nymb den
    <ingredient en="almond" commodity="Q39918"
  >mandel</ingredient></instruction> vnd <instruction>Stoß in
  chlain</instruction> vnd <instruction>streich In mit <ingredient en="wine"
    commodity="Q282">weinn</ingredient> durch</instruction> vnd
  <instruction>thue dar ein <ingredient en="ginger" commodity="Q15046077"
    >Imberr</ingredient></instruction> vnd <instruction>ain wincziger dunner
  <ingredient en="Fish stock (food)" commodity="Q30748629">visch <dish
    en="" commodity="Q275068">prue</dish></ingredient> dar vnder
  <alternative reason="instruction">ob dw der <ingredient
    en="Fish stock (food)" commodity="Q30748629">visch <dish en=""

```

# Ergebnis(se)

- Lesefassung / Rezeptfassung
  - normalisierte Fassung strukturiert nach Rezepten, Fließtext, semantisch annotiert (Online Lesefassung, )

## **\*Ain mandel suppen\* (Nr. 3)**

Ain mandel (almond) suppen Item ain mandel (almond) suppen wiltu die machen die geet von mandel milich (almond milk) zue nymm ain semlein (white bread) prot (bread) das pae

---

[17v]

schon die milich (milk) geuss daran wiltu die suppen haben .

# Transkription (dtspr. Kochrezepttexte)

- Transkribus (Universität Innsbruck) <http://transkribus.eu/>
- Plattform zur Erkennung und Transkription historischer Dokumente
- Infrastruktur frei zugänglich, hochgeladenen Daten nicht
  
- Bild-Textverknüpfung
- Seitenbasierte Struktur
- XML/TEI-Export

# Transkription








- hyperdiplomatische Transkription
  - *Basistranskription mittels propriäterem Code*
- Transkriptionsregeln
- 10%-Studie jeder Handschrift

# Prinzipien

- lautliche Unterschiede
- Bedeutungsunterschiede
- Gliederungsprinzipien
  - Rubrizierungen, Initialen, Schriftarten
- Abkürzungen
- Revisionen

**⇒ Projektspezifische Einschränkung auf den kleinstmöglichen Auszeichnungsumfang!**

# Codierungstabelle Beispiele

<u>Graphauszeichnung</u>	
<a> einstöckiges a	
<a2> zweistöckiges a	
<s> rundes s	 
<§> langes s (Schaft-s)	
<r> gerades r	
<r2> rundes/Rotunda r	

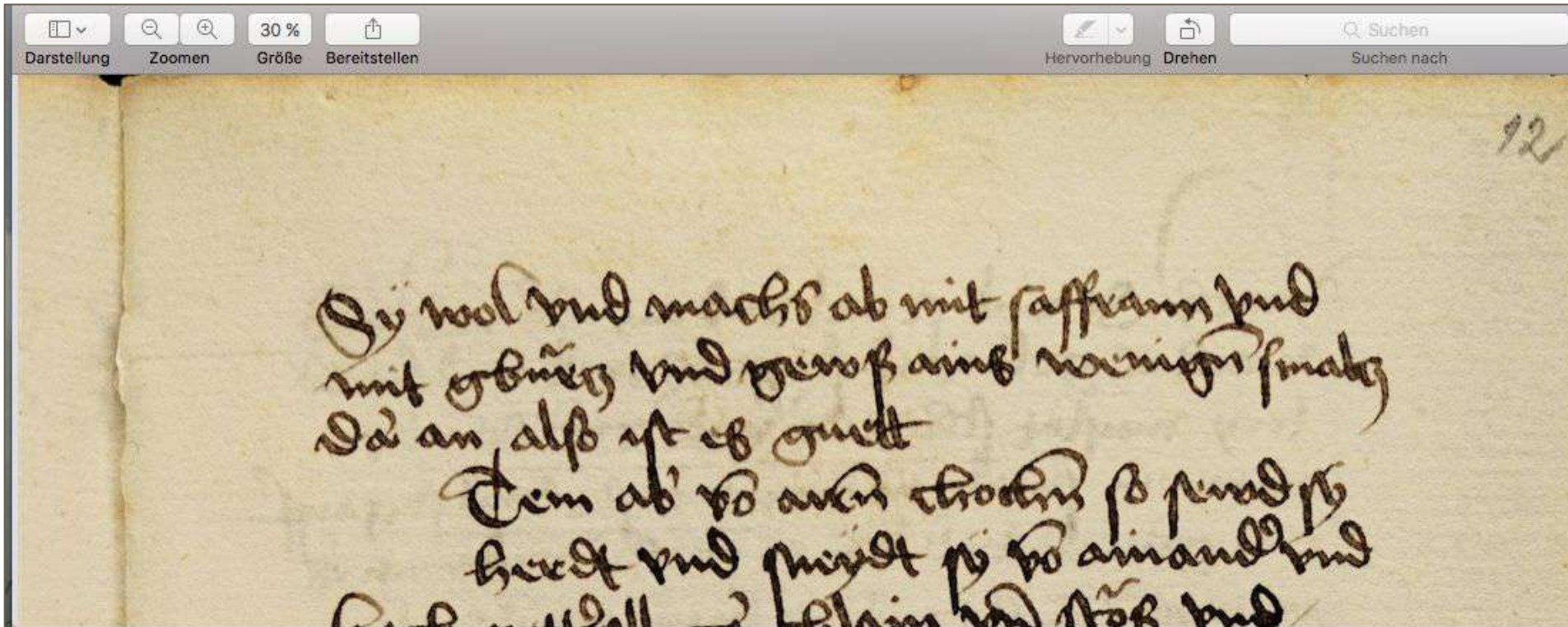
Kürzungsstrich („Nasalstrich“)		
-1	(e)	 lt(e)m
-2	(n)	 aine(n)
-3	(m)	 ny:(m)
-4	(d)	 vn(d)
-5	(en)	 mach(en)
Kürzungshaken		
'1	(r)	 da(r)
'2	(er)	 d(er)  wass(er)
'3	(ir)	 w(ir)t



# Annotationen

- Abkürzungen <abbr>
- Textfarbe < textStyle textColour>
- Unterstreichungen < textStyle underlined>
- Initialen < hi>
- Revisionen
  - Streichung, Radierung/Rasur, Bleichung < del>
  - Einfügungen < add>
- Editorische Rekonstruktion < supplied reason="...">
- Kommentare < note>

```
s_3' xml:id='Graz_UB_Ms_1609_0012r.jpg' n='3' />
s_3_r1'>
#facs_3_r1l1' n='N001' /> Sy: wol vnd machs ab mit §affrann vnd
#facs_3_r1l2' n='N002' /> mit gbu~rcz vnd gewiß ains <choice><expan></expan><abbr>wen
#facs_3_r1l3' n='N003' /> <choice><expan></expan><abbr>da&apos;1</abbr></choice> an
#facs_3_r1l4' n='N004' /> <supplied reason='missing'>I</supplied> Tem <choice><expan>
#facs_3_r1l5' n='N005' /> herdt vnd sney:dt §y: <choice><expan></expan><abbr>vo-2</a
#facs_3_r1l6' n='N006' /> hack <choice><expan></expan><abbr>pett&apos;2§i2ll</abbr><
#facs_3_r1l7' n='N007' /> Slachs durch mit <choice><expan></expan><abbr>guett-1n</ab
#facs_3_r1l8' n='N008' /> gs ab / an die air vnd gi2b §y: al§o chaldt
#facs_3_r1l9' n='N009' /> <supplied reason='missing'>I</supplied> Tem wil man he:rte
#facs_3_r1l10' n='N010' /> §ol <choice><expan></expan><abbr>mach-1n</abbr></choice>
#facs_3_r1l11' n='N011' /> §toß den vnd §lachs in mit wein durch
#facs_3_r1l12' n='N012' /> vnd ny:mb ho~ni2g <choice><expan></expan><abbr>da&apos;1<
#facs_3_r1l13' n='N013' /> gbu~rcz ab vnd ra~§t §y ym §malcz vnd thu:e
#facs_3_r1l14' n='N014' /> §y in dy: gru!e!n varb al§o <choice><expan></expan><abbr>
#facs_3_r1l15' n='N015' /> vnd guett
#facs_3_r1l16' n='N016' /> <supplied reason='missing'>I</supplied> Tem wi2lldw mach a
```



- N001: Sÿ wol vnd machs ab mit saffrann vnd
- N002: mit gbürcz vnd gewß ains wenig/en smalz
- N003: da/an also ist es guett
- N004: /Tem ab/er vo/n air/en choch/en so sewd fy
- N005: herdt vnd sneydt fy vo/n ainand/er vnd

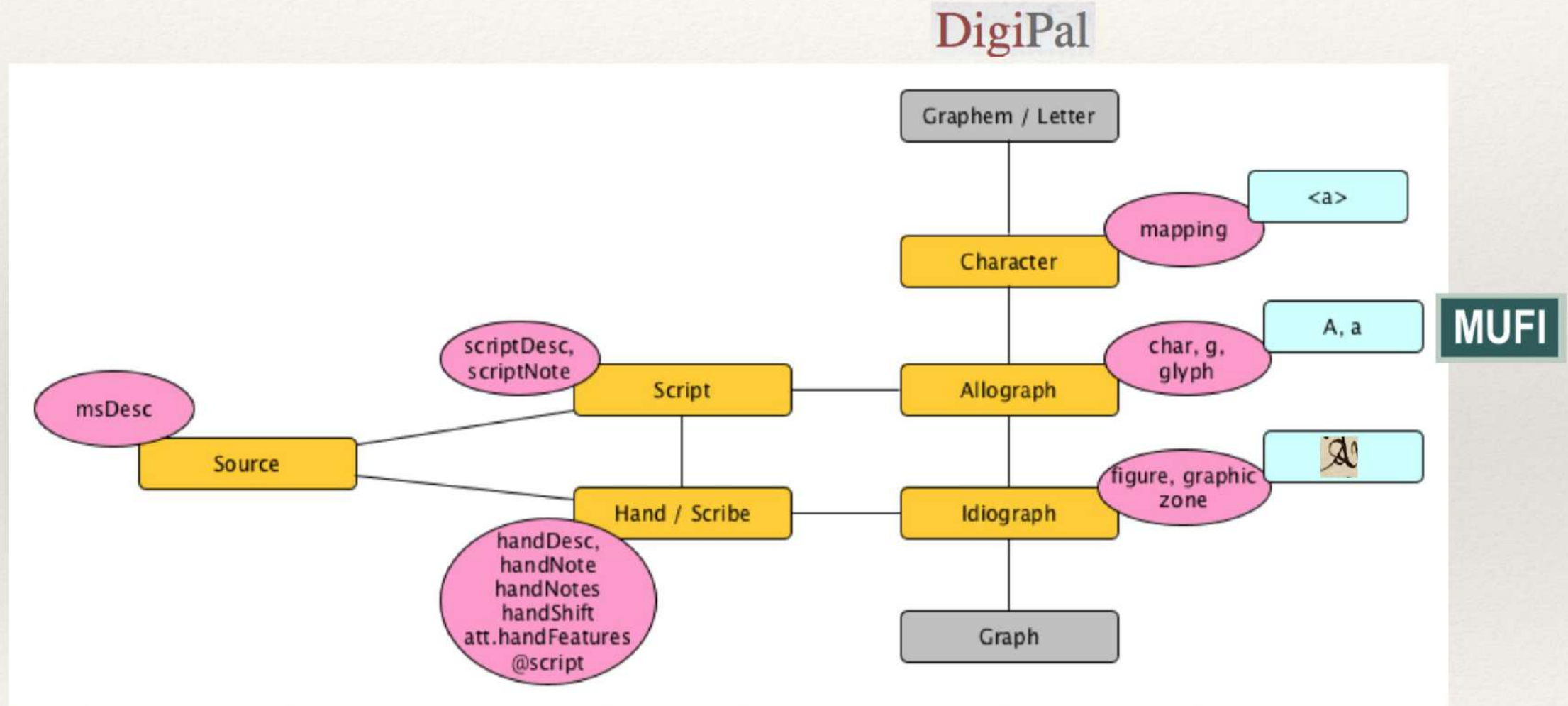
# Datentransformationen

1. Pre-processing
  - a. Transkribus TEI-Probleme ausbessern, Metadaten organisieren
2. Glyph orientierte Transformationen
  - a. <abbr>, <ex>
  - b. <g>
3. toHtml
  - a. Glyph, Abbreviaturen, automatische Auflösungen, Normalisierung
4. Normalisierung
  - a. Sonderinventar entfernen
  - b. Normalisierungsregeln anwenden

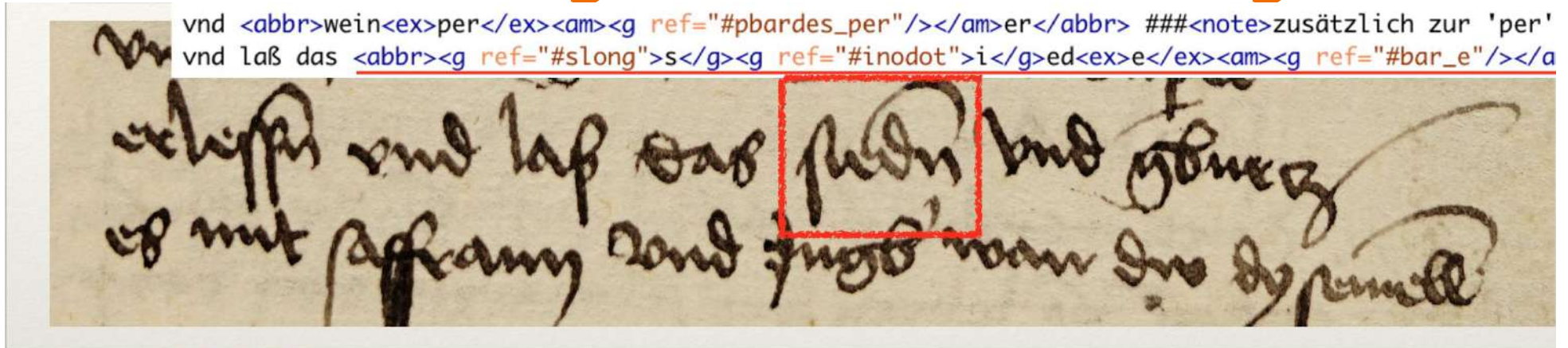
# Hyperdiplomatische Transkription

- Verortung des Textes am digitalen Faksimile
  - Transkribus: Textregionen
- Systematisierung der Schriftzeichen und der Textauszeichnung
  - Schriftzeichen
    - Hierarchie DigiPal
    - Beschreibung: MUFI

# Zeichenbeschreibung /-modellierung



# Zeichenkodierung: Textmodellierung



```
<char xml:id="i">
  <charName>LATIN SMALL LETTER I</charName>
  <mapping/>
</char>
<glyph corresp="#i" xml:id="inodot" ana="allograph"
  resp="https://bora.uib.no/handle/1956/10699" source="p.48">
  <glyphName>LATIN SMALL LETTER DOTLESS I</glyphName>
  <mapping type="normalized">i</mapping>
  <mapping type="transcription">i2</mapping>
  <mapping type="unicode_codepoint" subtype="LatExtA">0131</mapping>
  <mapping type="unicode_codepoint_hex" subtype="html_entity">&#305;</mapping>
  <mapping type="unicode_codepoint" subtype="symbol">i</mapping>
</glyph>
```

# Projekthomepage

demnächst unter

<https://gams.uni-graz/context:corema>

ähnliche Editionsrichtlinien im Projekt 'Mittelalterlabor'

<https://gams.uni-graz.at/context:malab>





# Potentiale

- Allgemein
  - *Frei zugängliche Daten (CC-BY)*
  - *maschinelle Verarbeitungsmöglichkeiten*
- Linguistik
  - *Historische Graphematik*
    - Unterscheidung auf Graphebene (Vokal- und Konsonantengrapfie)
  - *Historiolinguistik*
    - Interpunktion, “Linguistik der Streichungen”
    - Entwicklung Mündlichkeit-Schriftlichkeit, Fnhdt. Schriftsprache, Schreibsprachenbestimmung
    - Lexikografie
- “Paläografie”
  - *Schreiberhandanalyse*
- Text- und Literaturwissenschaft
  - Textgenese

---

(Projekt CoReMA –  
Cooking Recipes of the Middle Ages, fwf I-3614)  
Astrid Böhm / Christian Steiner / Helmut W. Klug

**VIELEN DANK FÜR IHRE AUFMERKSAMKEIT**

**Kontakt:** [astrid.boehm@uni-graz.at](mailto:astrid.boehm@uni-graz.at)

**Kontakt:** [helmut.klug@uni-graz.at](mailto:helmut.klug@uni-graz.at)

**Folien:** [http://bit.ly/corema\\_hyperdip](http://bit.ly/corema_hyperdip)