

---

Cost Workshop, 4.5.2010, Maynooth

---



TextGrid, TextGridLab  
and the  
Graphic Link Editor (Text – Image)



---

Michael Leuk, Simon Rettelbach

- TextGrid Project Description
- Introduction TextGridLab
- Presentation Graphic Link Editor
- Graphic Link Editor: features coming



- Phase I: TextGrid - Community Platform - 2006-2009
- Phase II: TextGrid – Virtual Research Environment - 2009-2012
- Joint project between 10 academic and commercial partners:  
Goettingen State and University Library (Project Management),  
Universities of Kaiserslautern, Munic, Paderborn, Trier and  
Würzburg, University of applied Sciences of Worms, Max Planck  
Digital Library, Institute for the German Language in Mannheim  
and Daasi International in Tübingen.
- Part of the D-Grid initiative funded by the German Federal  
Ministry of Education and Research (BMBF)



- Aim: Develop and establish a Virtual Research Environment in the Humanities
- Build and bring forward collaborative structures:
  - TextGridLab: Tools and Services
  - TextGridRap: Repository, long-term archive for research data
- First Grid Community Project in the humanities
- Core disciplines (Phase I and II):
  - Edition philology, Linguistics, Musicology, Art history



→ Generic services:

→ authentication, user/permission management, versioning, annotation with metadata, ...

→ Specific tools and services:

→ XML-Editor, Lemmatizer, Dictionary Grid, Graphic link editor (Text-Image)

→ Community Building:

→ Workshops and trainings

→ Survey of demands, bring users feedback to the developers

→ Expert monitoring of the project through an empirical survey (PhD proposal at Uni Trier)



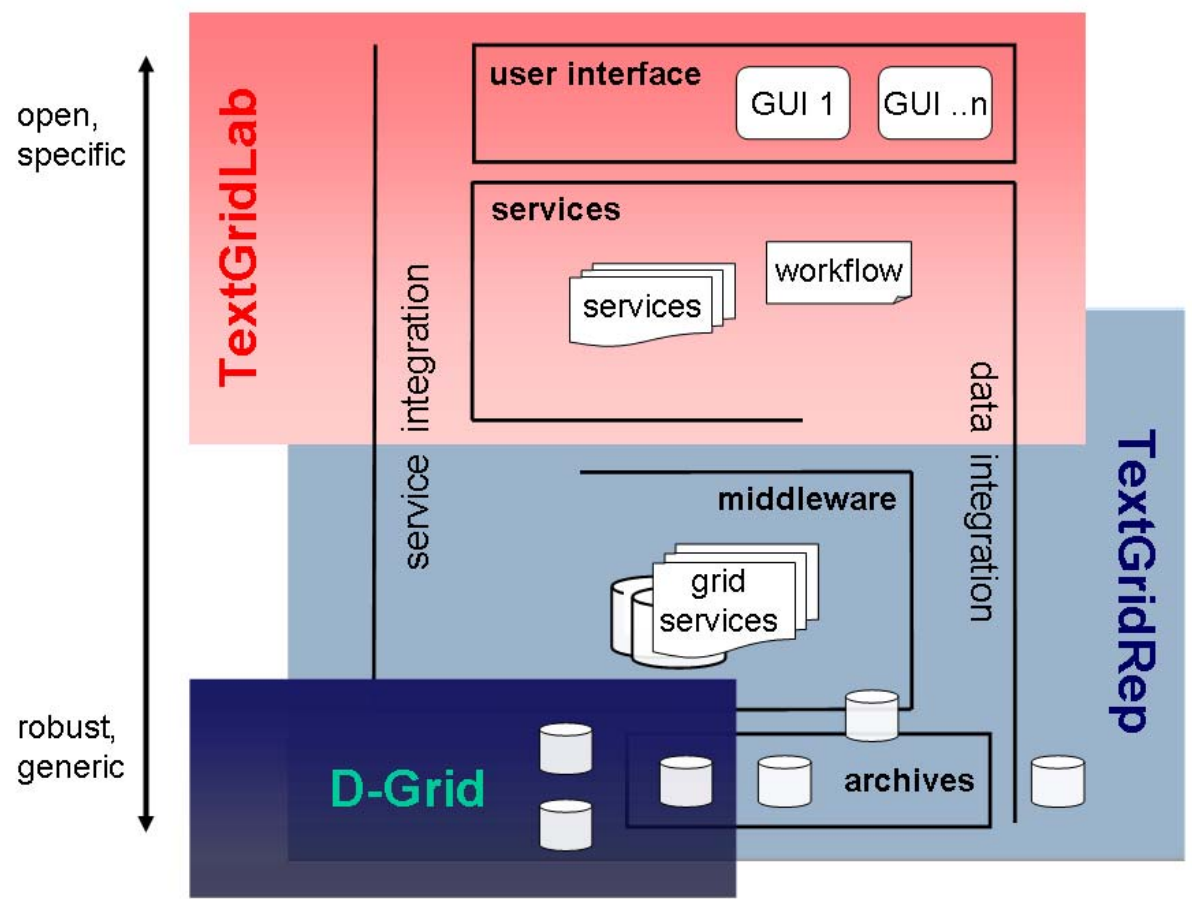
## TextGrid: Project Description



- TextGrid is part of the WissGrid Grid for Science together with:
  - AstroGrid (astronomy), C3Grid (climate research), HepGrid (high-energy physics), MediGrid (medical science)
- Aims of WissGrid:
  - consolidation of the scientific Grid communities
  - development of lasting generic structures
  - organisational model, business concepts, blue prints for forthcoming scientific Grid communities







- Frontend to the virtual research environment
- Use is (to a large extent) intuitive
- allows scholarly work with texts: analysing, editing, annotating, publishing, ...
- integrated access to all tools and services
- "portable software", location-independent access to the TextGrid via internet
- developed further continuously
- implemented with Eclipse (open-source development environment)



Generic services:



User and Project Management



Project Browser / Navigator



Search Tool

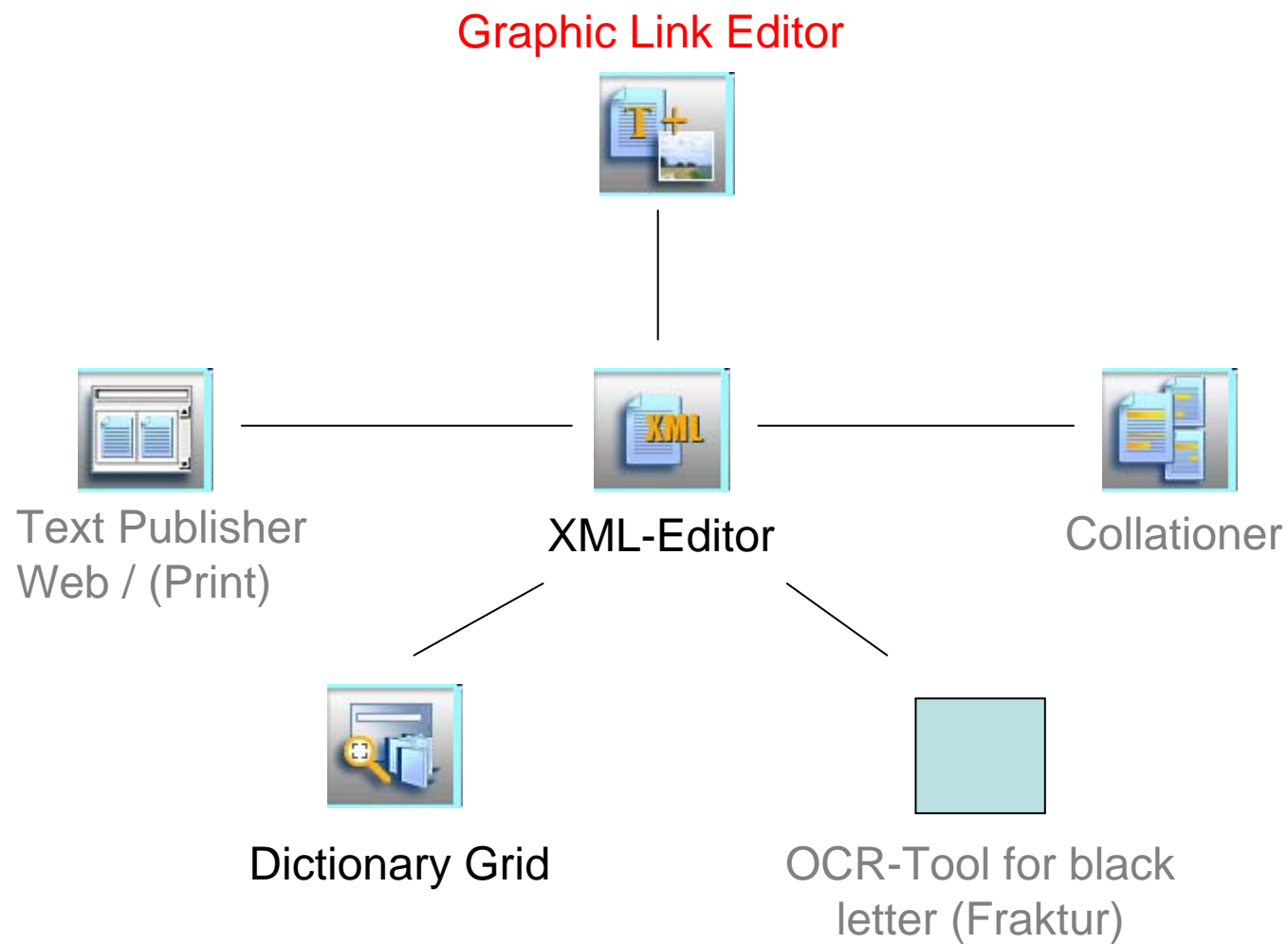


Metadata Annotator



- Specific Services and Tools – use cases:
- Since TextGrid I:
  - Edition philology: next slide...
  - Linguistics: XML-Editor, Streaming-Editor, Tokenizer, Lemmatizer
- In TextGrid II:
  - Art history: Integration and extension of DigiLib
  - Classical Philology: adapt edition tools to editing glosses
  - Musicology: GUI for editing and displaying music (in MEI encoding)





TextGridLab

Beta version...

TextGridLab 1.0

February 2011



## Graphic Link Editor: Standoff Markup

```

<TEI>
  <teiHeader/>

  <facsimile>
    <svg>
      <svg:g id="layer1">
        <svg:image width height xlink/>
        <svg:shape x y with height id="shape1"/>
        <svg:text writing-mode id="text1"/>
      </svg:g>
    </svg>
  </facsimile>

  <body>
    <linkGrp facs="#layer1">
      <link targets="#shape1 uri...anchor1 uri...anchor2"/>
      <link targets="#text1 uri...anchor1 uri...anchor2"/>
    </linkGrp>
  </body>
</TEI>

```

Image  
.tif, .jpg, .bmp



Transcription  
.xml (TEI)

```

<anchor1/>
<anchor2/>

```



... features coming:

- link elements which already have id-attributes to shapes without inserting <anchor>
  - no writing authorities needed on transcription file
- rotate image view
- 
- Tokenizer automatically tags words and attributes id's
- automated selection of lines and words in the facsimile using the segmentation tool of the OCR
  - (semi)automated linking of text and image
- image-catalogue for handwriting samples

