Francesca Caratti
Arne Recknagel
Matthias Lindinger

WiTTReader

Facsimile-Reader with extracted and highlighted WiTTFind-hits

Outline

- What does interoperability mean?

- What is the current status?

- Live demo of the current status

- What's new?
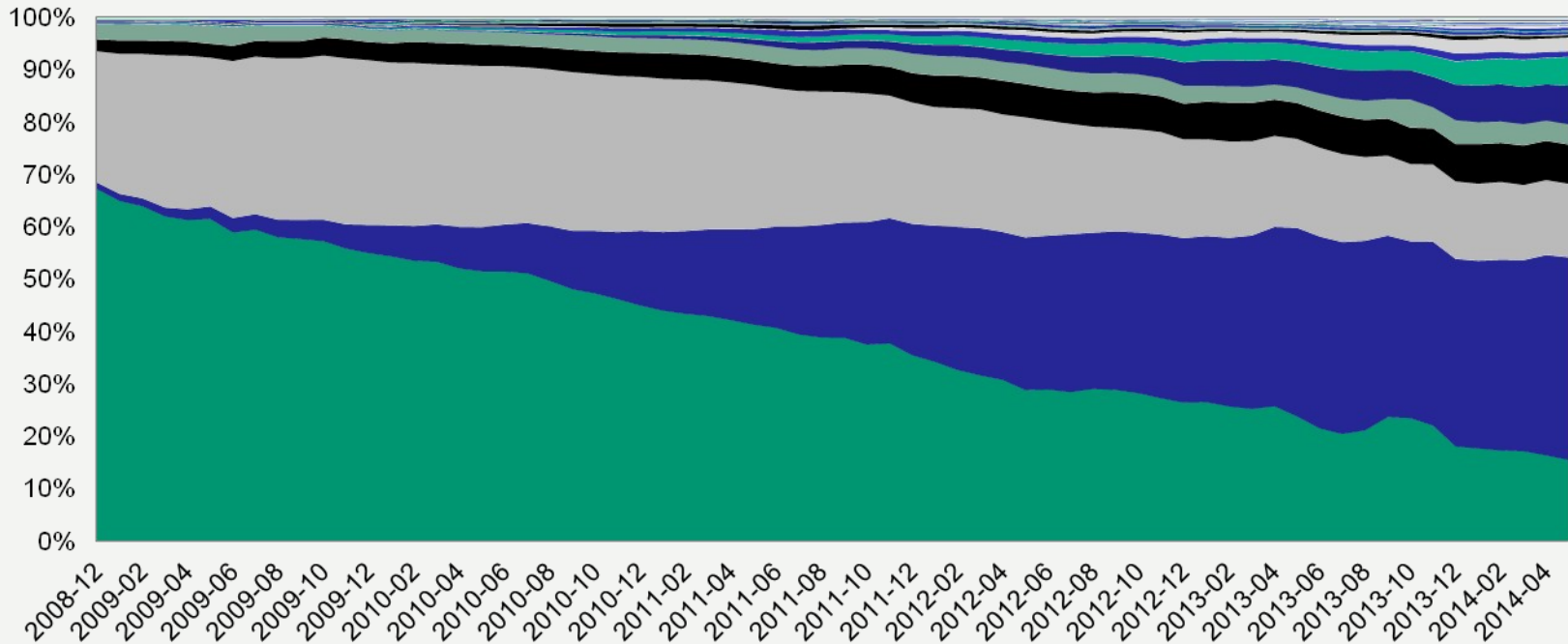
- Live demo of the new technologies

Interoperability

Application   —?→   ~~User~~

- For the last few years, diversity in IT has been on a steep rise.

- This means that interoperability is a more important concern than ever.

- For our purposes, it can be roughed down to this:

  - Established open standards

  - Developers who enforce them

- Example of such a standard: JQuery, abstracts client processes in order to improve interoperability. It is supported by all important browsers at this point.

Usage Share of Web Browsers

The Problem of Internet Explorer

- Even though its percentage share is falling, Internet Explorer is very much in use.

- Why is this a problem?

  - Internet Explorer is known for its slow update frequency, leading to a variety of security issues, as well as its bad performance.

  - Even though the newest version, IE 11, is considered to be better than its predecessors, it still lags behind its main competitors, Firefox, Chrome and Safari.

  - Additionally, many users are still equipped with one of the older versions, which do not display many websites correctly.

The Problem of Internet Explorer

- For all these reasons, and many more, some companies and web developers, among them Google and Facebook, have finally decided to no longer support IE 7.

- This just showcases the importance of interoperability – not even its still significant market share, neither the fact that it was made by one of the most powerful organizations in the world could save it when its compatibility was just too low.
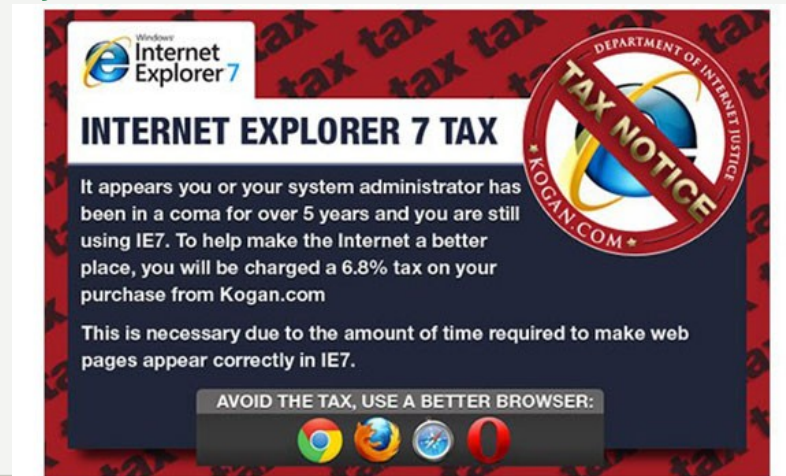
How did we do?

- Try it out! We tested it on many popular current browsers (Chrome, Firefox, IE 11, Safari, Opera and the Android web browser), and we still have to find an incompatibility of any kind.

Current status

- Fully implemented in PHP

- Presentation of a single siglum with all corresponding pages

- Static size

- Index of coordinates in a simple text file

Coordinates

- A list of coordinates for all the siglen

```
Ts-213,0r[1]/Ts-213,0r.jpg/style="top:244px; left:330px; height:39px; width:250px;"
Ts-213,1r[1]/Ts-213,1r.jpg/style="top:249px; left:143px; height:39px; width:705px;"
Ts-213,1r[2]/Ts-213,1r.jpg/style="top:376px; left:134px; height:99px; width:847px;"
Ts-213,1r[3]/Ts-213,1r.jpg/style="top:518px; left:134px; height:70px; width:847px;"
Ts-213,1r[4]/Ts-213,1r.jpg/style="top:593px; left:134px; height:178px; width:847px;"
```

**Live demo of the
current status**

New technologies

- Implemented in Javascript with jQuery and turn.js

- Creation of a collection with all WiTTFind-hits included

- Availability of all the TS-213 pages at once

- Dynamically load only of the required pages

- Resizable

- Index of coordinates in a JSON file

- New Helpbox

- Posibillity to swith between the hits

Centrum für Informations- und
Sprachverarbeitung (CIS)
Wittgenstein Scholarship 2014

LUDWIG-
MAXIMILIANS-
UNIVERSITÄT
MÜNCHEN

LMU

## Collection of all WiTTFind-hits

- Central part: 'Ts-213_X_ciswab.xml.wfc.out'

- Contains information the search engine returns after each query

- 'n' attribute of the hit-tag → Siglum information

- Flexible Design
    - Easy implementation of additional services (no pipeline)
    - Easy possibility to add information to the .out - file
    - → Without interfering with existing utilities

```
<hit n="Ts-213,i-r[7]_1" f="Ts-213,i-r" abnr="7" satznr="15" no="0">
 <match from="10562" to="10571" pos="7" token="verstehen" tag="VVINF" id="35613" gc="+V+refl+tr" ic=":1mGc:1mGi:3mGc:3mGi:OI" />
</hit>
<hit n="Ts-213,ii-r[14]_1" f="Ts-213,ii-r" abnr="39" satznr="91" no="1">
 <match from="43494" to="43503" pos="4" token="verstehen" tag="VVFIN" id="35613" gc="+V+refl+tr" ic=":1mGc:1mGi:3mGc:3mGi:OI" />
</hit>
<hit n="Ts-213,ii-r[14]_2" f="Ts-213,ii-r" abnr="39" satznr="92" no="2">
 <match from="43914" to="43923" pos="2" token="verstehen" tag="VVFIN" id="35613" gc="+V+refl+tr" ic=":1mGc:1mGi:3mGc:3mGi:OI" />
</hit>
```

Highlighting

- Siglum → Page and paragraph of the hit

- CSS-script does the highlighting

- Runs for all currently shown excerpts
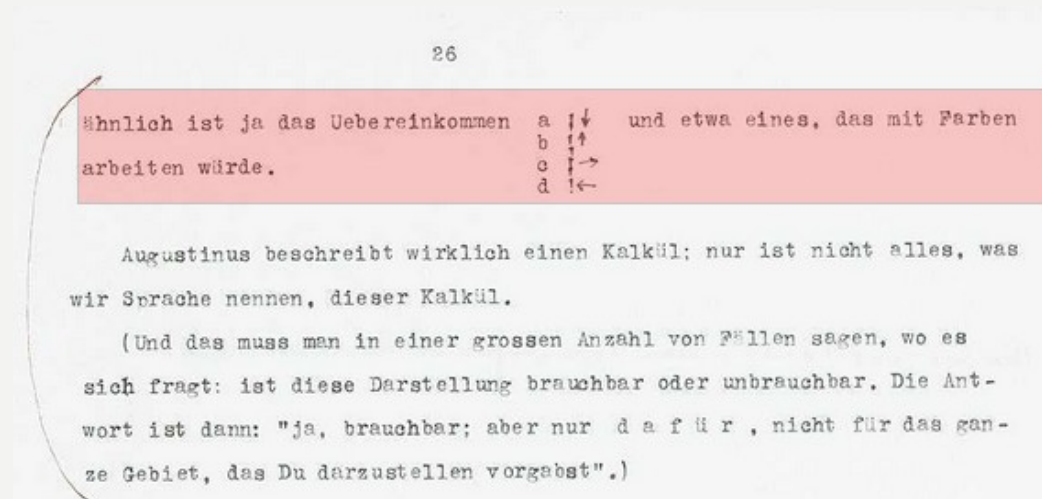
Choices

- Why highlight a paragraph instead of a word?
  - Because Wittgenstein wrote in short parapgraphs (2-3 sentences, very seldomly more than 5), and his work is only understandable in context

- Why not join the pipeline of content presention for the reader?
  - Because changes there could effect the reader, which is supposed to stand independently

- Why not ask the search engine only for the information we need right now?
  - Because the search engine is the only module that needs to have access to the actual data. By minimizing the accesses to it, we can turn it into the only service which needs to be run on our servers – all other work can then be finished locally

Index in JSON format

- Much better accessability than with a plain text file

```
"coordinates": [
  {
    "siglum": "Ts-213,i-r[1]",
    "coordinates": {"top": "6%", "left": "36%", "height": "3%", "width": "24%"}
  },
  {
    "siglum": "Ts-213,i-r[2]",
    "coordinates": {"top": "8%", "left": "12%", "height": "3%", "width": "81%"}
  },
  {
    "siglum": "Ts-213,i-r[3]",
    "coordinates": {"top": "11%", "left": "12%", "height": "3%", "width": "69%"}
  },
```

Helpbox with user guide

**Live demo of the
new technologies**

Possible next steps

- User language dependent

- Correspondent edition to the shown facsimile