Reflecting Knowledge Diversity

31.05.2012

The RENDER Project - Motivation

- WWW offers the possibility to publish, to share and to retrieve huge amounts of information
- large number of viewpoints, opinions, and background information
- increasing number of filtering mechanisms to pre-process the available data according to personalized preferences or settings
- difficult to get an all-embracing overview of a topic

→ information is inaccessible or it is a time consuming process

The RENDER Project

- develop methods, techniques, software and data sets which enable users to understand, to describe, to process and to make use the diversity of knowledge and information

- To verify the scalability and use of the research, results will be applied in three case studies

The RENDER Project

- duration 10/10 – 09/13
- 7 partners from 5 countries
- budget of 4.193.565 €
- Coordinated by KIT
- 384 person months
- 7 work packages
- 54 deliverables

Diversity in Wikipedia

WP articles are usually written by multiple editors, who may be biased towards a certain point of view

→ necessary that either editors can transcend their personal point of view, or that a multitude of editors covers the significant points of view

Diversity is a necessity for high-quality Wikipedia articles!

Diversity aspects for Wikipedia

The most important aspects of diversity:

- Neutrality
- Fact coverage
- Timeliness

Besides these parameters RENDER takes editor behaviour and interaction (KIT) into account.

Wikipedia Use Case - Goals

- The goal of Wikimedia's case study is to support Wikipedia editors in maintaining and improving the site, and to support readers in understanding the quality and biases of a given article.

- Tools and extensions to support editors in the management, understanding, and decision-making about complex and heated controversies on Wikipedia.

- We want Wikipedia to offer high quality articles on both highly visible and as well as on more obscure topics.

Wikipedia Use Case Scenarios

Main Goal: Improvement of the quality, the value and the trustworthiness of Wikipedia by supporting Wikipedia users (readers and editors)

- UC1: Display warnings to the reader when detecting bias
- UC2: Notify authors that an article needs to be updated
- UC3: Lower the barrier for readers to extend and/or correct articles

Provide the analysis results for researchers

Supporting Tools

Tools to support readers/ editors/ administrators:

- **Article Statistics Quality Monitor** (ASQM):
    - in particular to support the readers
- **Task List Generator**:
    - editors concerning problems of the content
    - administrators concerning editor behaviour and interaction

Article Statistics Quality Monitor

Article Statistics Quality Monitor

- Statistics and quality/ diversity measures for a certain article
- User can decide which metrics or values she wants to observe
- Need of modular architecture to extend with further analysis tools and/or available scores

- By providing these information the user/reader can understand the collaborative editing process, detect lacks in an article and gets the information to cure these problems

Article Statistics Quality Monitor (ASQM)

Statistics:

Status: Featured article since 19/01/2007
First edit: 2002-07-09 16:11 (Ben-Zin)
Recent edit: 2012-05-28 14:15 (Matthiasb)
Number of references: 57
Number of pictures: 53

Number of edits: 2618
Number of unique editors: 1113 (526 IPs)

Number of views (last 30 days): 28844
Number of views (yesterday): 1204   ↑ less...

RENDER Analysis:

Fact coverage: View page ratings
Neutrality: View page ratings
Timeliness: View page ratings
Editor Interaction: more information   ↑ less...

Further Assessment Scores:

Wikibu.ch: 10/10   more information
Article feedback score:
 * Trustyworthy: 4.0
 * Objective: 4.3   more information
 * Complete: 3.6
 * Well-written: 4.3
WikiTrust: View page ratings

http://de.wikipedia.org/wiki/Potsdam

Task List Generator

Task List Generator

Discuss this page

Select a flaw: biased article ▲▼   Learn more about this tool

Search by topic:

Select a category: Add a further category

Show 50 | 100 | 250 | 500 biased articles   GO

Biased articles

Verotoxin-producing Escherichia coli   Neutrality score: X   More information
Michael Jackson's health and appearance   Neutrality score: Y   More information
2018 Winter Olympics   Neutrality score: Z   More information
...   ...

- User can choose one specific topic, a preferred category or the intersection of categories
- User can select a specific lack she wants to cure e.g. all articles with a neutrality problem

Diversity Analysis Tools

Within the RENDER project we and our partners developed the following tools:

- LEA (LinkExtrActor)

- Change Detector

- Wikipedia Map

- Corpex

RENDER Toolkit for Knowledge Diversity in Wikipedia

LEA   **Change Detector**   **Wikipedia Map**   **Corpex**

Home

Info

Feedback

RENDER-Toolkit

Hi and welcome to the RENDER Toolkit Page!

Here you can test and play with different tools. These are developed in the context of the research project RENDER (Reflecting Knowledge Diversity). In this project, Wikimedia Deutschland is a use case partner.
The tools provide an insight into the different analyses of diversity in Wikipedia, which are implemented within the project. For more Information about the project, please visit RENDER on Meta-Wiki or the project website.

The whole toolkit as well as the tools are in a continuous process of development. We are looking forward to receiving suggestions, feedback or ideas on the discussion page concerning RENDER on Meta-Wiki or by email.

RENDER is funded by

http://toolserver.org/~RENDER/toolkit/

Further Diversity Analysis Tools

News Query Tool (JSI)

- Combining with Change Detector results → providing sources for additional information (e.g. URL, keywords, publishing date)

Political Bias Detector

- Data set of parliamentary minutes provided by OKFN: 17. legislative period, sessions 33 – 166, 215.765 speeches

- → Computation of word frequency and expression analysis for each party
- → Signal a warning if characteristic political words occur in an article

Opinionated Wikipedia Checker (JSI)

- we offered neutrality data sets (DE and EN) as training data

Open Questions - Frontend

- Is there any (documented) experience with layout and design aspects we should consider?

- RENDER Toolkit design and functionality

- How should we provide these tool results?

  - Gadget links to external page or embedded within the article page?

- List Generator - Testing with small groups of authors e.g. 2 or 3 WikiProjects or further ideas?

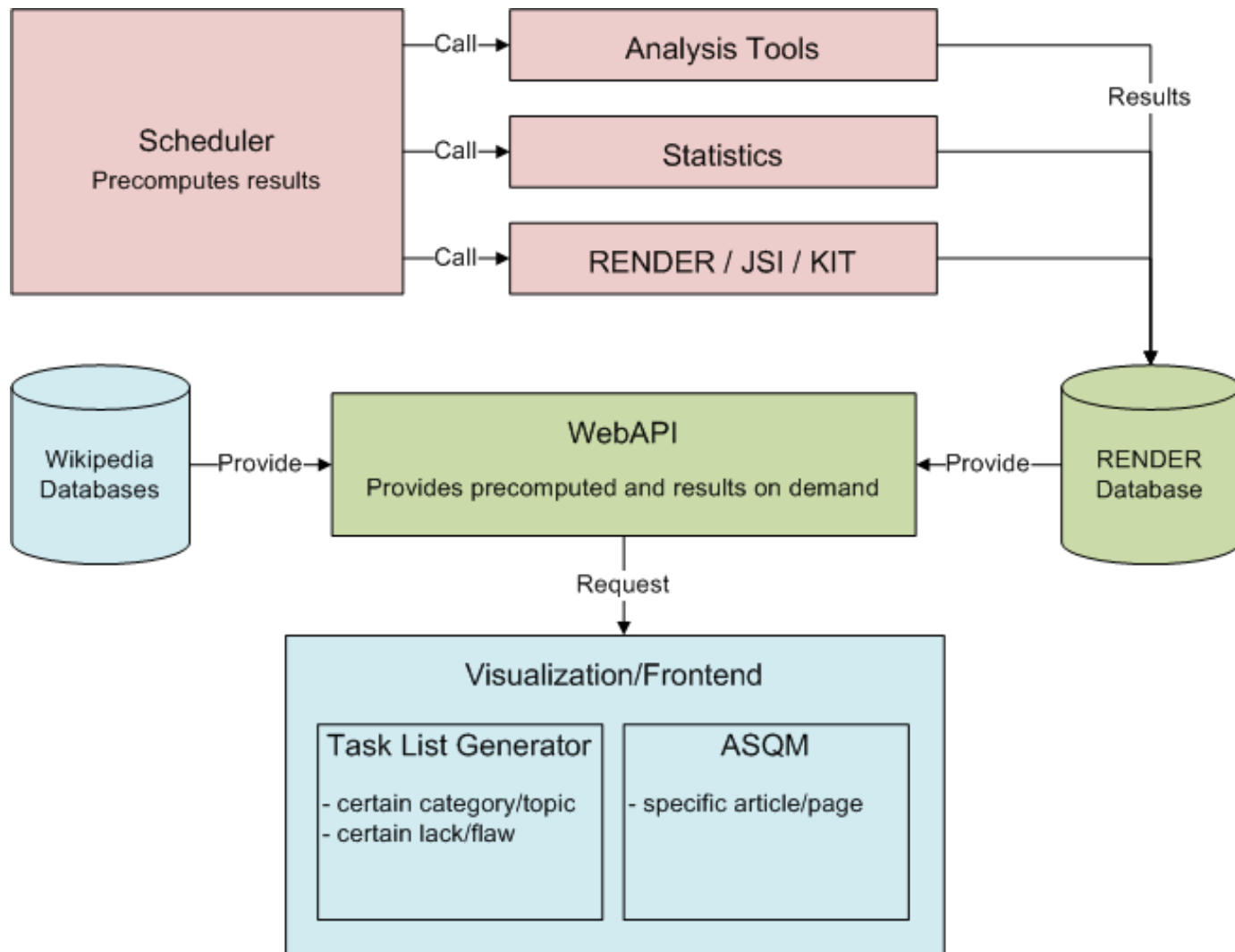- ASQM - Inclusion and testing for readers (without an account)?

Open Questions - Backend

- Discuss our architecture draft
  - Practicability of WikiLabs for our concerns?
    - General roadmap of feature enhancements (saved snapshots in particular)
    - Access to live databases (or replications) possible?
- Additional available analysis tools?
- Data of Article Feedback Tool
  - No dumps available since July 2011
  - Possibilities to access?

Appendix – RENDER Architecture (draft)

Thanks for your attention!