The Durm Project

- [Cultural Heritage Data](#)
- [Durm](#)

toc_collapse=0; Table of Contents

1. Overview

The Durm project, carried out from 2004-2006 at the Institute for Program Structures and Data Organization (IPD) at the University of Karlsruhe, Germany, investigated the use of advanced semantic technologies for cultural heritage data management. The goal was to support end users, in particular users from building history and architecture, with tools that go beyond classical information retrieval techniques. Experiments were carried out on the historical *Handbuch der Architektur* (Handbook on Architecture).

2. Project Results

Some of the major project results include the Durm corpus, the Durm Wiki, and the Durm German Lemmatizer, as well as a number of publications.

2.1. The Durm Wiki

One of the major results from this project is our strategy for transforming cultural heritage data into a Wiki format. A [public version](#) (albeit with a restricted set of features due to the shared web hosting environment) is available online.

2.2. The Durm Corpus

As part of the project, we digitized a complete book of the encyclopedia and transformed it into several formats. Both [scanned page images and the marked-up text](#) are freely available.

2.3. The Durm German Lemmatizer

Due to the rather special language of the Durm corpus, namely 100-year old architectural texts, we needed a lemmatizer that can accurately compute lemmas even for outdated, technical terminology. Since freely available resources for German are rather scarce, we developed a [self-learning context-aware lemmatizer for German](#), which is available under an open source license.

2.4. Publications

A complete list of publications from this project is available [here.](#)

3. Project Members

Project supervision:

- [Peter C. Lockemann](#) (IPD, University of Karlsruhe, Germany)
- [René Witte](#)

The following group members contributed to this project:

Tom Gitzinger
  Index Generation, Wiki/NLP Connection
[Thomas Kappler](#)
  Durm Wiki, Durm Corpus, Wiki Bot
[Ralf Krestel](#)
  Automatic Summarization, Ontology Population
Qiangqiang Li
  NLP Pipelines, Ontology Population
Praharshana Perera
  Index Generation, German Lemmatization.

4. Acknowledgments

5. Feedback

For questions, comments, etc., please use the [Durm Forum](#).

Semantics for the Masses

Source URL (retrieved on *2025-12-05 22:03*): https://www.semanticsoftware.info/durm