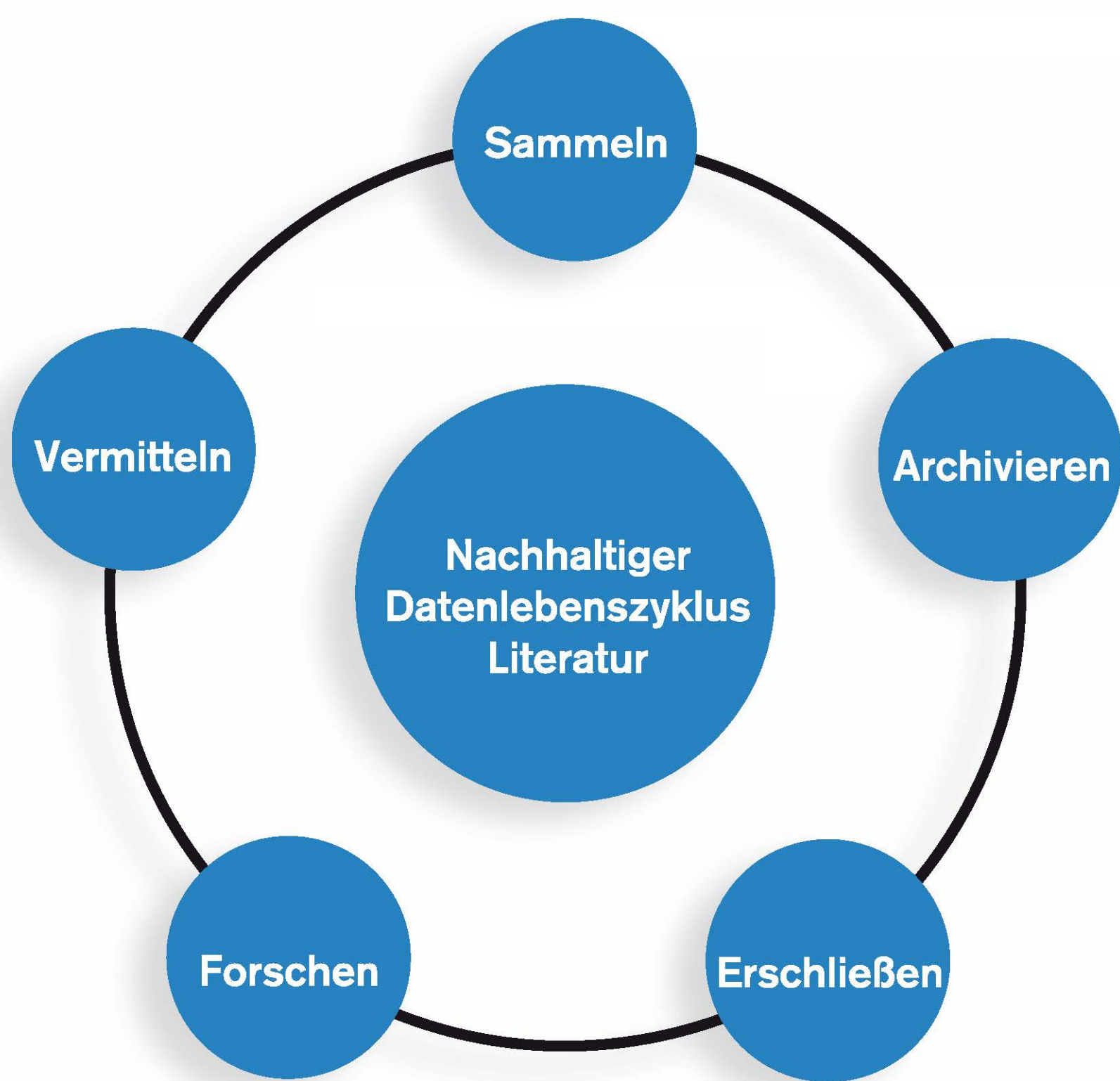


Infrastruktur, Methodik und Kompetenzvermittlung für einen Datenlebenszyklus Digitaler Literatur



Science Data Center for Literature

Volodymyr Kushnarenko², Claus-Michael Schlesinger³, Nina Buck¹, Andreas Ganzenmüller¹, Mona Ulrich³, Heinz Werner Kramski³, Jan Hess³, Alex Holz³, André Blessing⁴, Pascal Hein², Kerstin Jung⁴, Nicolas Schenk³, Thomas Bönisch¹, Roland S. Kamzelak³, Jonas Kuhn⁴, Gabriel Viehhauser²



Aufgaben und Ziele

Aufbau eines nachhaltigen Datenzentrums zur Sammlung, Langzeitarchivierung, Erschließung, Erforschung und Vermittlung digitaler literarischer Quellen mit Hilfe moderner computerbasierter Methoden:

- Modellierung von Born-Digitals und Literatur im Netz
- Integration und Entwicklung von computergestützten Methoden für die Textanalyse und für die Arbeit mit digitalen Objekten
- Forschung und nutzerorientierte Entwicklung auf der Grundlage von Fallstudien zu digitaler Literatur
- Schnittstellen zur Integration von Quellen, Methoden, Werkzeugen und Best Practices in die Arbeitsumgebung der Plattform
- Nachhaltige Verwaltung und Sicherung der Daten nach den FAIR-Prinzipien

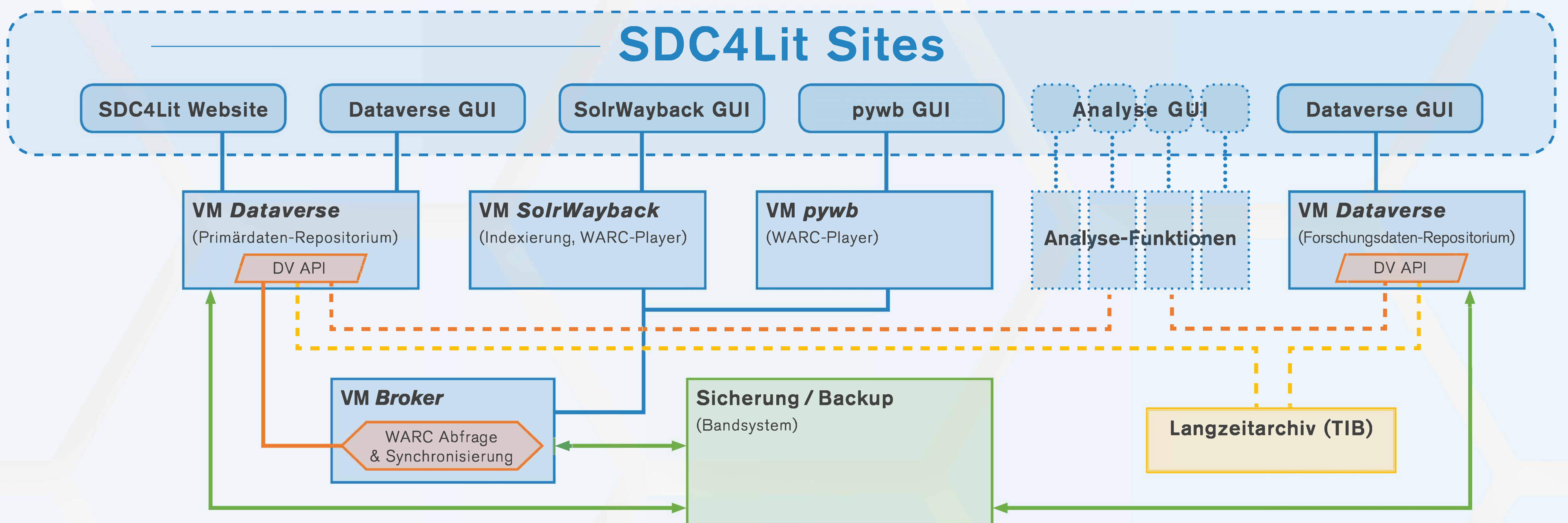
Methodik

- **Forschungsgegenstand:** genuin digitale literarische Materialien wie Netzliteratur, digitale Vor- und Nachlässe und literarische Computerspiele
- **Digitale Literatur:** neue Ansätze für die Archivierung, Bereitstellung und Erforschung aufgrund der besonderen Ästhetik und Medialität notwendig
- **Digitale Materialien:** nicht nur Speicherung, sondern Entwicklung von neuen Mechanismen und Methoden für die Rekonstruktion, Wiedergabe und Analyse erforderlich
- **Arbeit mit Forschern, Studenten und Archivaren:** Anwendung von bestehenden und neu entwickelten Methoden, Vermittlung von erarbeiteten Know-hows und Erfahrungen, Bereitstellung von mehreren Anleitungen sowie Organisation von Seminaren, Workshops und Vortragsreihen zu projektbezogenen Themen

Kompetenz

- **Durchführung von Fallstudien:** Techniktagebuch Analyse [<https://techniktagebuch.tumblr.com>], Linknetzwerk Hyperfiction, Archivierung von Blogs, etc.
- **Jupyter-Notebooks:** Skripte für die Kombination von verschiedenen Named Entity Recognition (NER) Systemen
- **Entwicklung von warc2graph:** Extraktion der graphischen Datenstruktur aus den WARC-Dateien [<https://github.com/dla-marbach/warc2graph>]
- **Interdisziplinäre Zusammenarbeit:** Expertisen aus den Bereichen Archiv, Literaturwissenschaft, Höchstleistungsrechnen, Maschinelle Sprachverarbeitung und Digital Humanities
- **Infrastruktur:** Vereinigung der digitalen Materialien, Methoden und Tools sowie Anleitungen von dazugehörigen Analysemethoden

Infrastruktur



Projekt Partner

- 1 Deutsches Literaturarchiv Marbach (DLA)
- 2 Höchstleistungsrechenzentrum Stuttgart (HLRS), Universität Stuttgart
- 3 Institut für Literaturwissenschaft/Digital Humanities (ILW/DH), Universität Stuttgart
- 4 Institut für Maschinelle Sprachverarbeitung (IMS), Universität Stuttgart

Danksagung

Das Projekt SDC4Lit wird gefördert vom Bundesministerium für Wissenschaft, Forschung und Kunst in Baden-Württemberg



Dieses Poster ist unter CC BY 4.0 lizenziert