

A Common PID Kernel Information Profile for the German Helmholtz Association of Research Centres

Thomas Jejkal, Andreas Pfeil, Jan Schweikert, Anton Pirogov, Pedro Videgain Barranco, Florian Krebs, Christian Koch, Gerrit Günther, Constanze Curdt, Martin Weinelt

In the concept of FAIR Digital Objects, PID Kernel Information is key to machine actionability of digital content. Strongly relying on Data Types and stored in a PID record directly at the PID resolution service, PID Kernel Information allows to be used by machines for fast decision making. As a first step into the direction of standardizing PID Kernel Information, the RDA Working Group on PID Kernel Information has defined a first proposal of a core Kernel Information Profile (KIP) together with a list of seven guiding principles helping to decide on which information could be part of a KIP and which information should be stored elsewhere.

The Helmholtz Association

- Largest association of (18) large-scale research centers in Germany (<https://www.helmholtz.de/>)
- Covers a wide range of research fields with 43k employees
- Initiated Helmholtz Metadata Collaboration Platform in 2019 to harmonize metadata handling
- Goal is to realize a FAIR metadata platform across all research fields of the Helmholtz Association

RDA WG on PID Kernel Information

- Provides recommendations on PID Kernel Information allowing smart machine actionable decisions
- Introduces guiding principles for defining KIP attributes
- Compiles Draft Kernel Information Profile
- Presents architecture and workflow on how to store and evaluate PID Kernel Information
- Available at DOI: 10.15497/RDA00031

Motivation and Goal

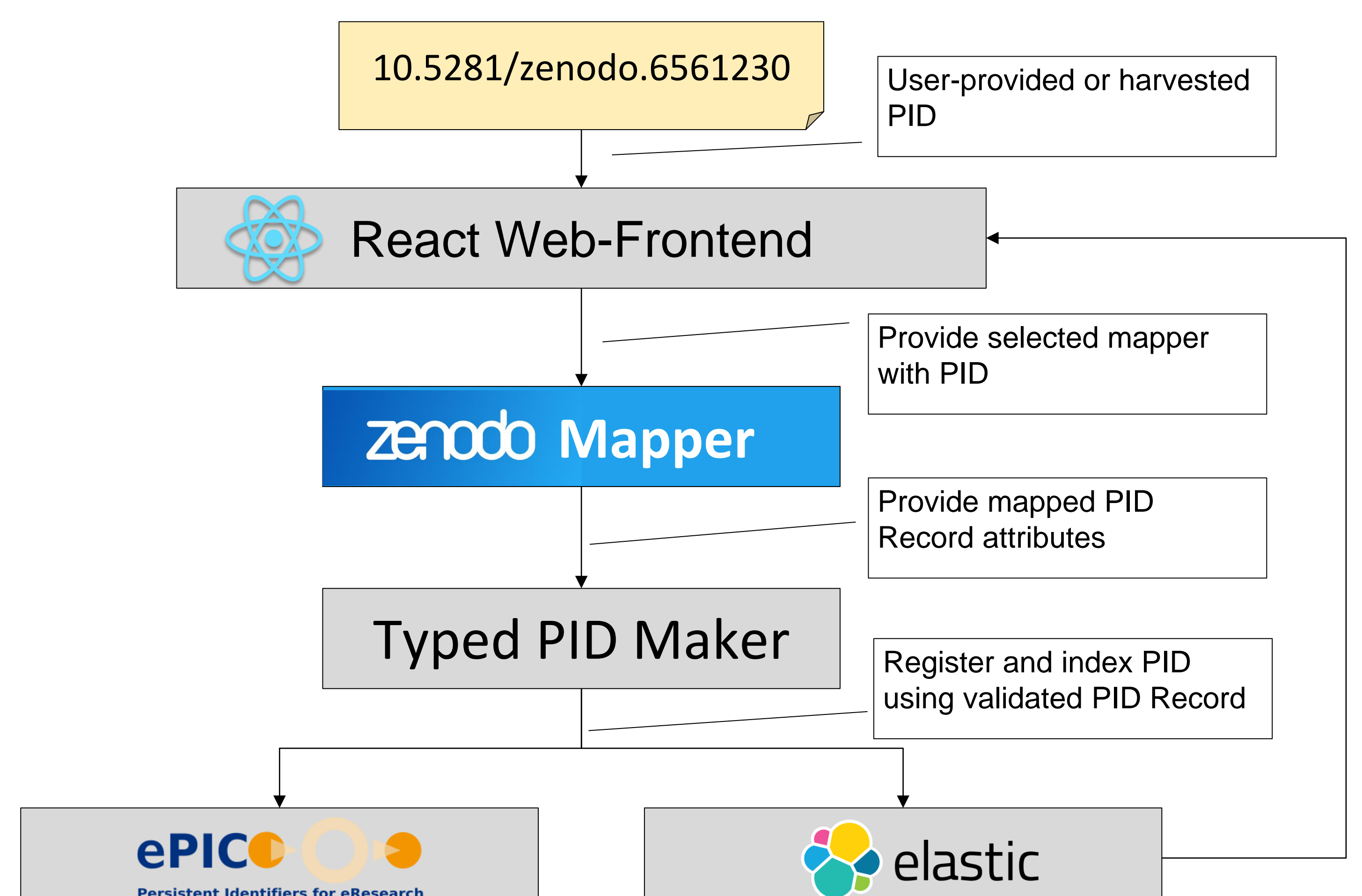
- FAIR Digital Objects identified as top-level commonality across all research fields
- Evaluate Draft Kernel Information Profile for applicability in HMC and extend sensibly
- Agree on a base KIP that can be applied to all digital artifacts produced by the Helmholtz Association
- Show applicability by implementing a demonstrator for (semi-) automatic creation of FDOs from existing repositories

Results and Next Steps

- Adopted most elements of RDA Draft KIP
- Skipped PID attribute due to unclear usage
- Renamed 'etag' attribute to 'checksum'
- New attributes to provide contextual information
- Working on demonstrator for (semi-) automatic provisioning Helmholtz KIP attributes
- Phase 1: Initial Showcase
 - Collect PID Record information via Zenodo DOI
 - Harvest Zenodo to build FDO collection
- Phase 2: Growth and Benefit
 - Improve automatism, support for other repositories
 - Prove benefits with growing FDO collection

Additional Helmholtz KIP Attributes	Comment
digitalObjectLocation-AccessProtocol	Access information for digitalObjectLocation, e.g., protocol, protocol version, and client
underEmbargoUntil	Access restrictions probably apply before
license	Extracted from digitalObjectPolicy
checksum	Renamed from ,etag' to be more specific
signature	Cryptographic signature of PID record
topic	Topic term from vocabulary for additional context
locationPreview	Optional preview for digitalObjectLocation
contact	Contact information, e.g., ORCID or ROR
hasMetadata	PID pointing to a related FDO containing metadata
isMetadataFor	Inversion for hasMetadata
wasGeneratedBy	W3C PROV-DM element to refer to tool/agent used for generating the digital object
provenanceGraph	Optional PID of full provenance graph

Additional attributes of Helmholtz KIP extending the RDA Draft KIP



Architecture draft for Helmholtz KIP Demonstrator