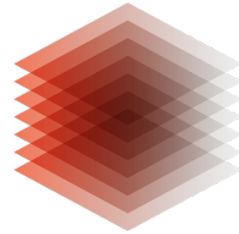

LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

Future plans for TIB AV-Portal – modularization and replacement of software components

Felix Saubier
Hamburg, November 7
Fedora User Group Meeting

Agenda

1. TIB and TIB AV-Portal
2. Technology
3. Systems Architecture
4. Future

Agenda

- 1. TIB and TIB AV-Portal**
- 2. Technology**
- 3. Systems Architecture**
- 4. Future**

Technische Informationsbibliothek (TIB)

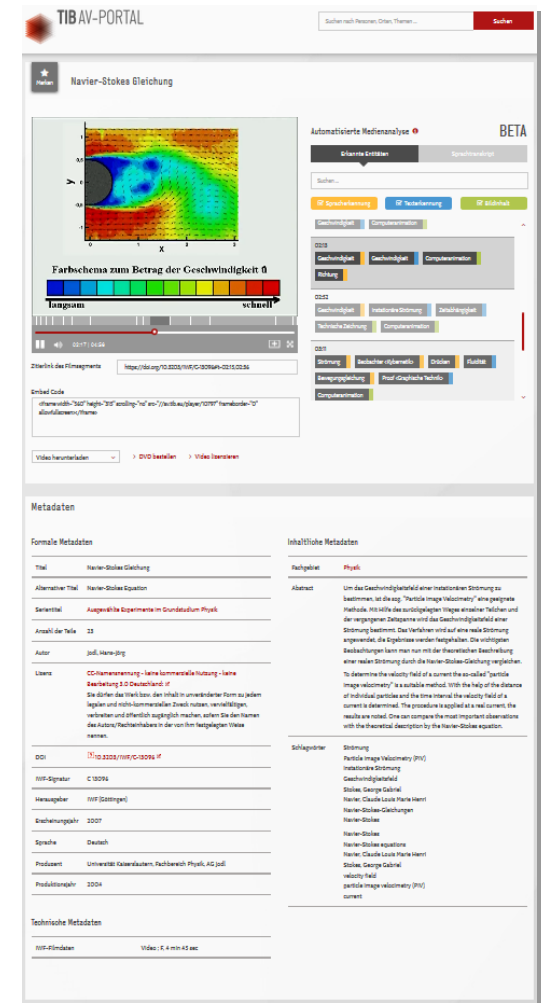


- German National Library for Science and Technology
- University Library at Hannover
- The world's largest science and technology library
- An infrastructure provider for the whole scientific work process
- TIB strategy: **Move beyond text**
- **Competence Centre for Non-Textual Materials**



TIB AV-Portal (av.tib.eu)

- Platform for **quality-tested scientific videos**
- Online since April 2014
- Developed by TIB and Hasso Plattner Institute
- Automatic metadata enrichment**, DOI/MFID, long-term preservation, semantic search
- More than 11,500 Videos (December 2017)
- Conference recordings, lectures, experiments, video abstracts, simulations, animations
- Videos predominantly under open access licences



The screenshot displays the TIB AV-Portal interface. At the top, there is a search bar and a 'Suchen' button. Below this, the video title 'Navier-Stokes Gleichung' is shown. The main content area features a video player with a color-coded velocity field visualization. To the right of the player is a control panel with buttons for 'Automatisierte Medienanalyse', 'Suchen...', 'Erneuern', 'Teilen', and 'Erstellen'. Below the player, there is a 'Zurück des Filmmaterials' section with a URL and a 'Druck Code' section with a QR code. The bottom section is titled 'Metadaten' and is divided into 'Formale Metadaten' and 'Inhaltliche Metadaten'. The 'Formale Metadaten' table includes fields for Title, Alternative Title, Summary, Anzahl der Teile, Autor, Lizenz, DOI, MFID, Herausgeber, Erscheinungsjahr, Sprache, Produzent, and Publikationsjahr. The 'Inhaltliche Metadaten' table includes fields for Fachgebiet and Physik. The 'Formale Metadaten' table is as follows:

Formale Metadaten
Titel Navier-Stokes Gleichung
Alternativer Titel Navier-Stokes-Exposition
Summary Ausgewählte Experimente im Grundvorlesung Physik
Anzahl der Teile 11
Autor juil, Hans-Jörg
Lizenz CC-BY-NC-ND/4.0 - keine kommerzielle Nutzung - keine Bearbeitung 3.0 Deutschland 0F Sie dürfen das Werk bzw. den Inhalt in unveränderter Form zu jedem legalen und nicht-kommerziellen Zweck nutzen, vorbehaltlich vorheriger und offenkundig angegebener, sofern Sie den Namen des Autors/Rechteinhabers in der von ihm festgelegten Weise nennen.
DOI 10.52020/nfz-cibosn-0211223a
MFID CIBOSN
Herausgeber INF (Stuttgart)
Erscheinungsjahr 2007
Sprache Deutsch
Produzent Universitätsbibliothek, Fachbereich Physik, AG juil
Publikationsjahr 2004

The 'Inhaltliche Metadaten' table is as follows:

Inhaltliche Metadaten
Fachgebiet Physik
Abstract Um das Geschwindigkeitsfeld einer inkompressiblen Strömung zu bestimmen, ist die sog. "Particle Image Velocimetry" eine geeignete Methode. Mit Hilfe des zurückgelegten Trages einzelner Teilchen und der erregenen Dehnung wird das Geschwindigkeitsfeld einer Strömung bestimmt. Das Verfahren wird auf eine reale Strömung angewendet, die Ergebnisse werden festgehalten. Die wichtigsten Beobachtungen kann man mit der theoretischen Beschreibung einer realen Strömung durch die Navier-Stokes-Gleichung vergleichen.
Schlüsseler Strömung Particle Image Velocimetry (PIV) Inkompressible Strömung Geschwindigkeitsfeld Navier, George Gabriel Navier, Claude Louis Marie Henri Navier-Stokes Gleichungen Navier-Stokes Navier-Stokes equations Navier, Claude Louis Marie Henri Stokes, George Gabriel velocity field particle image velocimetry (PIV) current

The 'Technische Metadaten' table is as follows:

Technische Metadaten
MFID-Findcode Video: 1, 4 min 42 sec

Agenda

1. TIB and TIB AV-Portal

2. Technology

3. Systems Architecture

4. Future

Semantic Video Analysis – Process



Scene Recognition (SBD)



Speech Recognition (ASR)



Text Recognition (OCR)



Image Recognition (VCD)

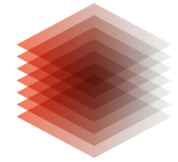


Named Entity Linking (NEL)

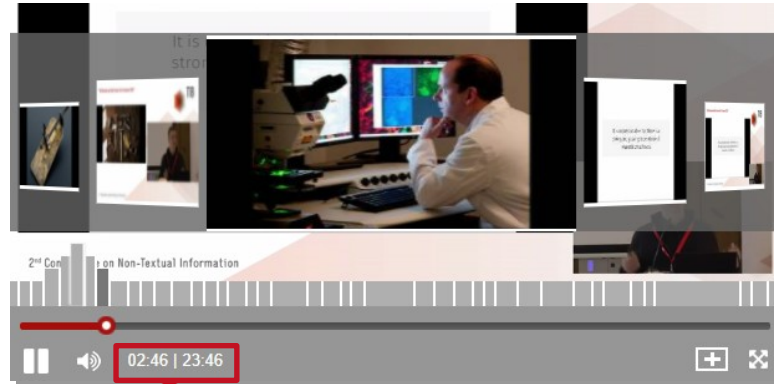


The screenshot displays the TIB AV-Portal interface. At the top, there is a search bar with the text "Search for people, places, topics ..." and a "Search" button. Below the search bar, the video title "What is good scientific practice for research software?" is shown. The video player shows a man speaking at a podium. To the right of the video player is the "Automated Media Analysis" sidebar, which is currently in "Recognized Entities" mode. The sidebar lists various entities such as "Compilation album", "Data analysis", "Dew point", "Meeting/Interview", "Maxima and minima", "Code", "Mathematical analysis", "Computer animation", "Database", "Software", "Utility software", "Estimation", "Protein", "Heat transfer", "Scripting language", "Demo (music)", "Maize", "Partition (number theory)", "Simultaneous localization and mapping", "Data structure", "Mereology", "Mass", "Species", "Sheaf (mathematics)", and "Real number". The sidebar also includes a search bar and buttons for "Speech", "Text in the video", and "Image content".

Semantic Video Analysis – Results



Video Segments



Audio Transcript

02:46
 hitting science well about what people images that in software to crimes this data right
 you need these tools there is no signs without software today and minus catching
 maybe bottflies that at some point you would use this to a generator a database of the
 butterflies again and researchers need maybe to classify all this so software is essential
 for science today the it can also

Named Entities

02:44

Resultant	Point (geometry)	Database	Service (economics)
Information	Whiteboard	Computer animation	Software
Software	Right angle	Sign (mathematics)	Game controller
Medical imaging	Function (mathematics)		


Semantic Analysis and Indexing

Named Entity Linking

- Based on Audio-/OCR-Transcripts
- Knowledge Base:
 - 63.356 GND Entities
 - TIB core subjects

Cross-lingual Mapping

- Heuristic Mapping approach
- English labels from
 - DBpedia
 - LCSH, MACS, TEMA



The screenshot displays the 'Automated Media Analysis' interface, which is currently in 'BETA' mode. The interface is divided into two main sections: 'Recognized Entities' (active) and 'Speech transcript'. A search bar is located below the tabs. Three filter buttons are visible: 'Speech' (checked), 'Text in the video' (checked), and 'Image content' (checked). The main content area shows a timeline of recognized entities from 00:00 to 02:06. The entities are represented as colored bars with labels: 'Software' (blue), 'Self-organization' (orange), 'Service (economics)' (blue), 'Information' (blue), 'Computer animation' (green), 'Electronic program guide' (orange), and 'Physicalism' (orange). The interface also includes a vertical scrollbar on the right side.

Live Demo



<https://av.tib.eu/>

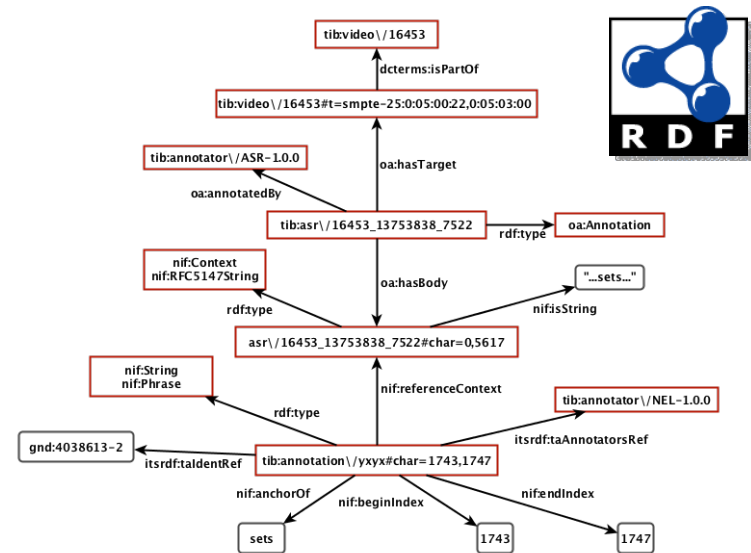
Data model

Resource Description Framework (RDF)

```
tib:asr/31031_26463533_23007#char=483,494 itsrdf:taAnnotatorsRef tib:annotator/NEL-1.0.0 ;
  itsrdf:taIdentRef <http://d-nb.info/gnd/4488209-9> ;
  nif:anchorOf "source code" .
```

Vocabularies

- Bibframe Vocabulary
- DCMI Metada Terms
- DCMI Type Vocabulary
- Friend of a Friend Vocabulary
- **Open Annotation Data Model**
- **NLP Interchange Format**
- **Internationalization Tag Set (ITS) Ontology**



<https://av.tib.eu/opendata>

Metadata Publication & Linked Open Data

- CC0 RDF-Dumps (since 2016)
- Dereferencable URIs & 303 (end of 2017)
- LDF-Server (since 2017)
- Public SPARQL-Endpoint (2018+)

DATENSÄTZE

GESAMTBESTAND


Dateiname	Format	Größe	Erzeugt am:	Version:
tib-av-portal-export-1.0.2.rdf.zip	rdf/xml	102M (unzipped 1655M)	31.08.2016	1.0.2
tib-av-portal-export-1.0.2.nt.zip	text/n-triples	101M (unzipped 1287M)	31.08.2016	1.0.2
tib-av-portal-export-1.0.2.ttl.zip	text/turtle	98M (unzipped 856M)	31.08.2016	1.0.2

TIB-FÄCHER: TECHNIK SOWIE ARCHITEKTUR, CHEMIE, INFORMATIK, MATHEMATIK UND PHYSIK
 Diese Dumps sind eine Teilmenge des Gesamtbestands. Sie enthalten die Videos zu den TIB-Fächern Technik sowie Architektur, Chemie, Informatik, Mathematik und Physik.

Dateiname	Format	Größe	Erzeugt am:	Version:
tib-av-portal-export-tib-subjects-1.0.2.rdf.zip	rdf/xml	101M (unzipped 1646M)	31.08.2016	1.0.2
tib-av-portal-export-tib-subjects-1.0.2.nt.zip	text/n-triples	100M (unzipped 1281M)	31.08.2016	1.0.2
tib-av-portal-export-tib-subjects-1.0.2.ttl.zip	text/turtle	97M (unzipped 851M)	31.08.2016	1.0.2

FILME DES HERAUSGEBERS IWF WISSEN UND MEDIEN G6MBH I.L.
 Diese Dumps sind eine Teilmenge des Gesamtbestands. Sie enthalten nur die Videos des Herausgebers IWF Wissen und Medien gGmbH i.L.

Dateiname	Format	Größe	Erzeugt am:	Version:
tib-av-portal-export-iwf-1.0.2.rdf.zip	rdf/xml	5M (unzipped 78M)	31.08.2016	1.0.2
tib-av-portal-export-iwf-1.0.2.nt.zip	text/n-triples	5M (unzipped 60M)	31.08.2016	1.0.2
tib-av-portal-export-iwf-1.0.2.ttl.zip	text/turtle	5M (unzipped 41M)	31.08.2016	1.0.2

Linked Data des TIB AV-Portals 

Powered by Triple Pattern Fragments.

Choose datasources: ⌵

Type or pick a query:

```

SELECT DISTINCT ?movie
WHERE {
  ?movie rdf:type schema:Movie .
  ?movie schema:publisher <http://av.tib.eu/resource/IWF_%28G%C3%86ttinger%29> .
}
    
```

Execute query

Query results:

Execution log:

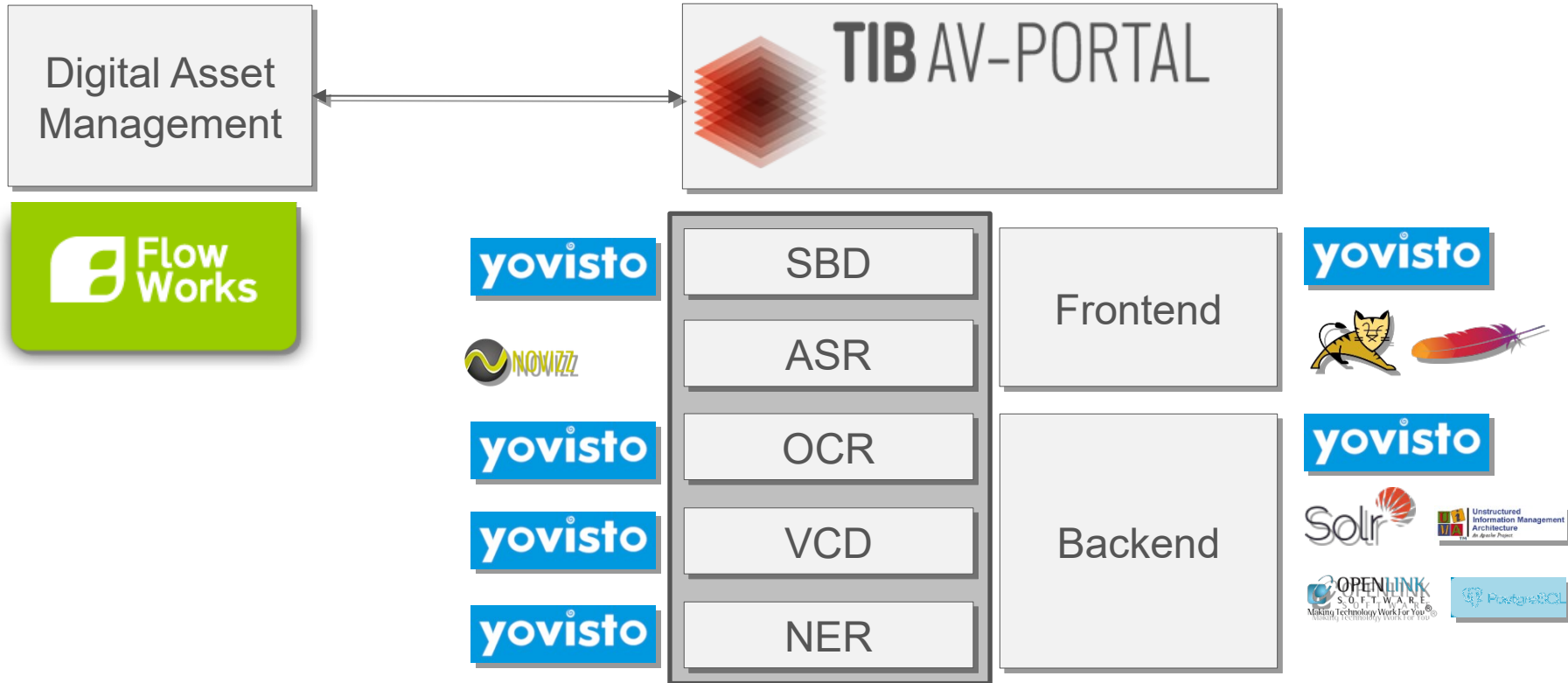
Discover how [Linked Data Fragments](#) enable Web-scale querying of Linked Data.

©2013-2016 Ghent University - Inez, Belgium. Source code.

Agenda

1. TIB and TIB AV-Portal
2. Technology
- 3. Systems Architecture**
4. Future

Systems Architecture 2017



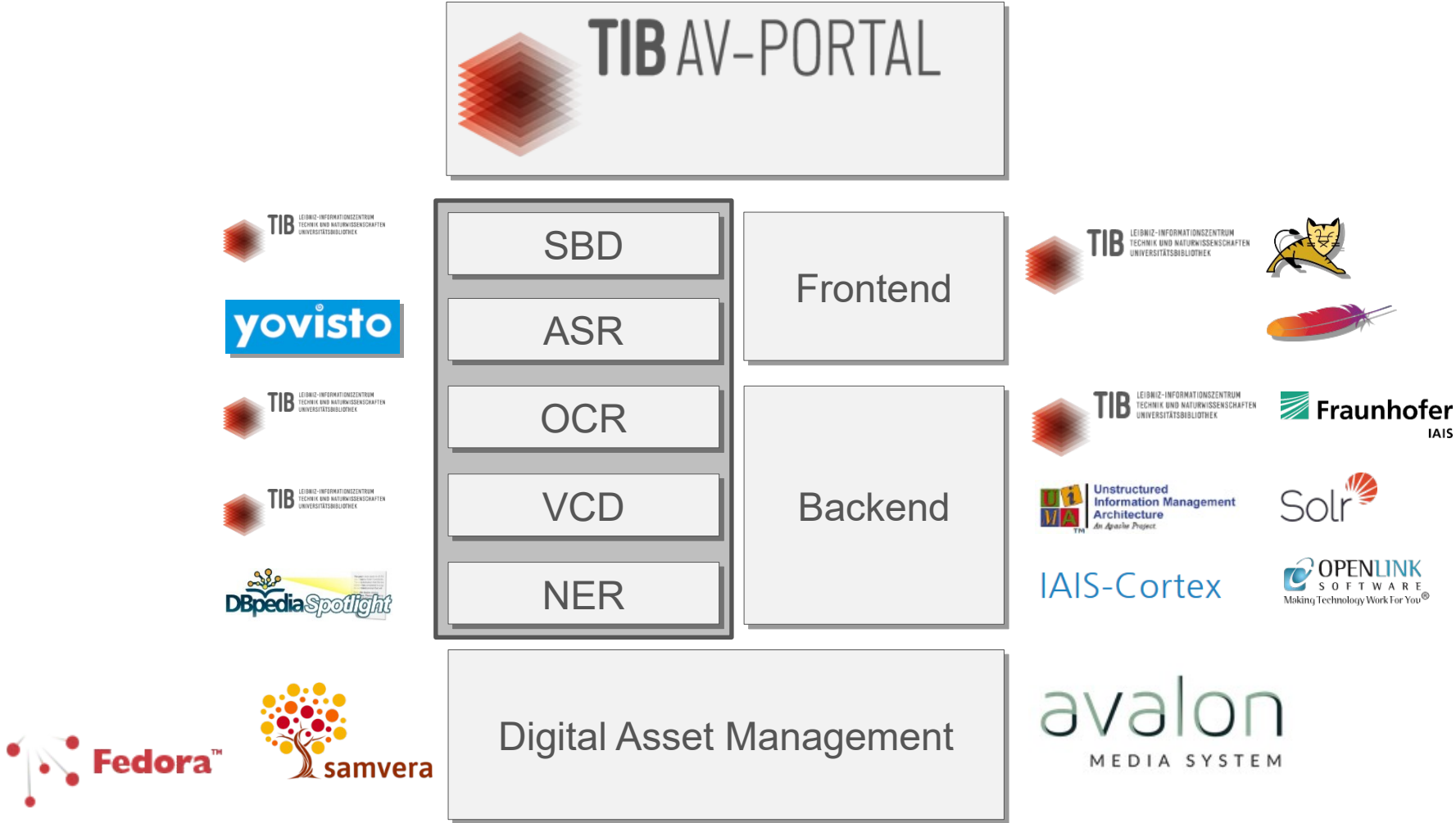
Agenda

1. TIB and TIB AV-Portal
2. Technology
3. Systems Architecture
4. Future

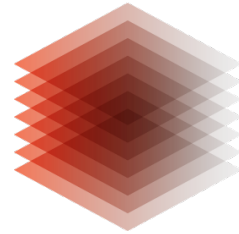
Goals 2018 – 2020

- Inhousing and Hosting at TIB
- Modernization of all System Components
- Modularization of Systems Architecture
- Technical Development at TIB
- Basis: Open Source Software

Systems Architecture 2018 – 2020



LEIBNIZ INFORMATION CENTRE
FOR SCIENCE AND TECHNOLOGY
UNIVERSITY LIBRARY



TIB

More Infos

KNM@tib.eu

av.tib.eu

Contact

Felix Saurbier

T +49 511 762-14645, felix.saurbier@tib.eu