
● Gaia-X Meets FAIR Data: FAIR Data Spaces

Development of a Cloud Based Data Space for Sharing
Data Among Science and Industry

Dr. Marius Politze
IT Center, RWTH Aachen University

Project FAIR Data Spaces: Overview

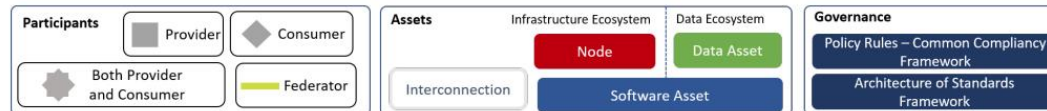
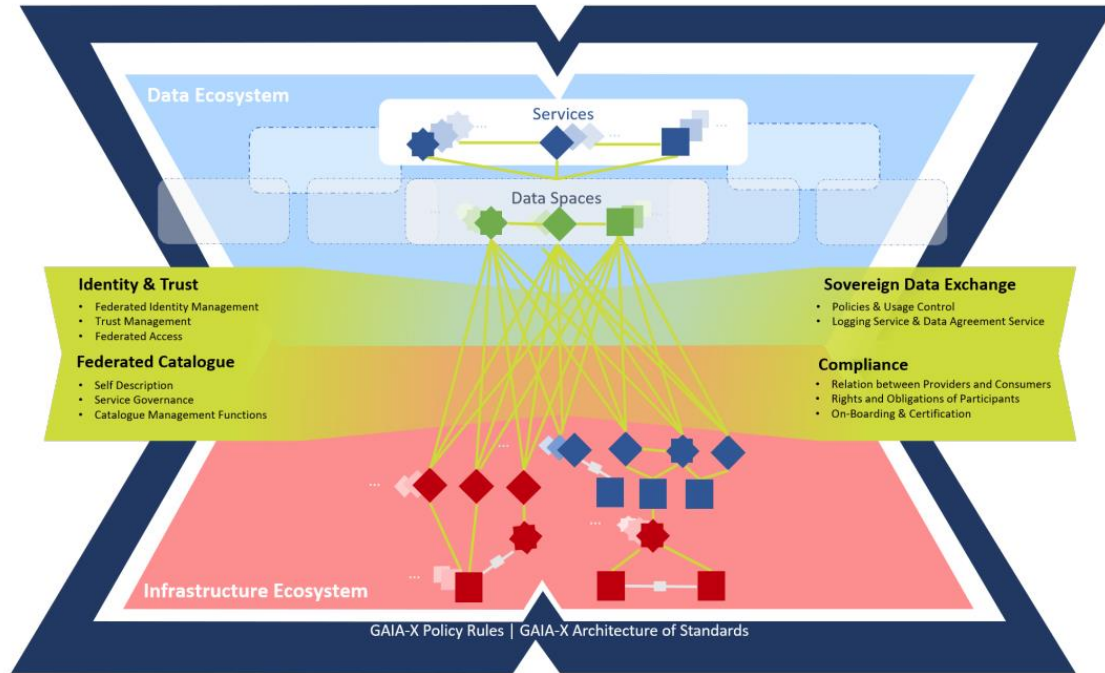
- Vision: Development of a Cloud based Data Space for sharing data among science and industry
 - Based on FAIR Principles

- Mission: Find and make use of synergies of the existing initiatives

	Deutschland	Europa	Welt
Wirtschaft	International Data Spaces (IDS): Grundlage für souveränen Datenaustausch in Gaia-X; insbesondere in Deutschland schon „reifere“ Initiative	Gaia-X : föderierte, sichere Dateninfrastruktur	
Wissenschaft	Nationale Forschungsdateninfrastruktur (NFDI)	European Open Science Cloud EOSC	Global Open Science Cloud GOSC

Gaia-X Ecosystem

From: Gaia-X AISBL: Gaia-X Architecture Document



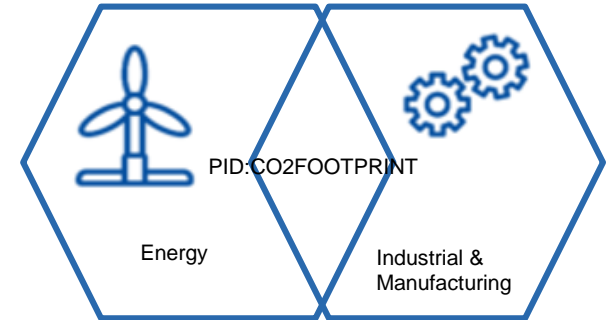
Gaia-X Shared European Data Spaces

- Part of European Data Strategy
- Based on FAIR Principles (?)
- It is not a fixed set



A Data Space is not an Island

- Data Space Interoperability
- Common PIDs on Data Level
- Ontologies and Vocabularies

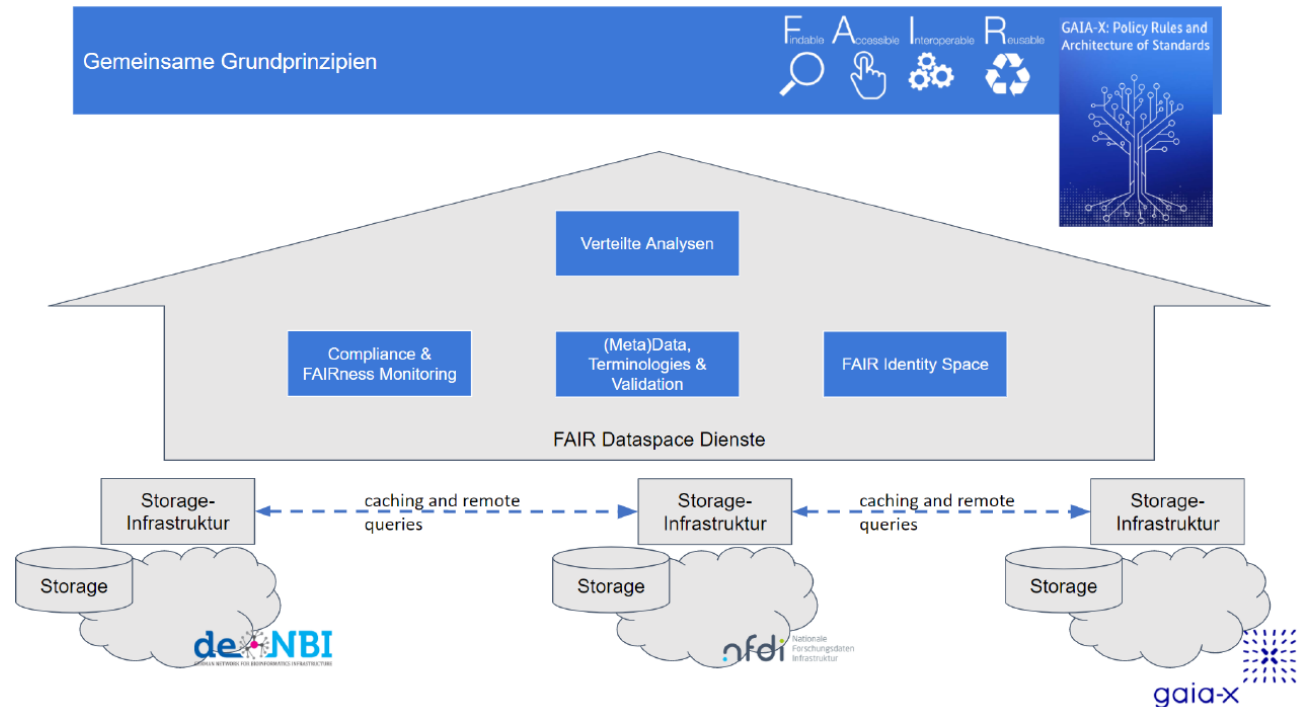


FAIR Data Spaces: Work Plan and First Achievements ✓

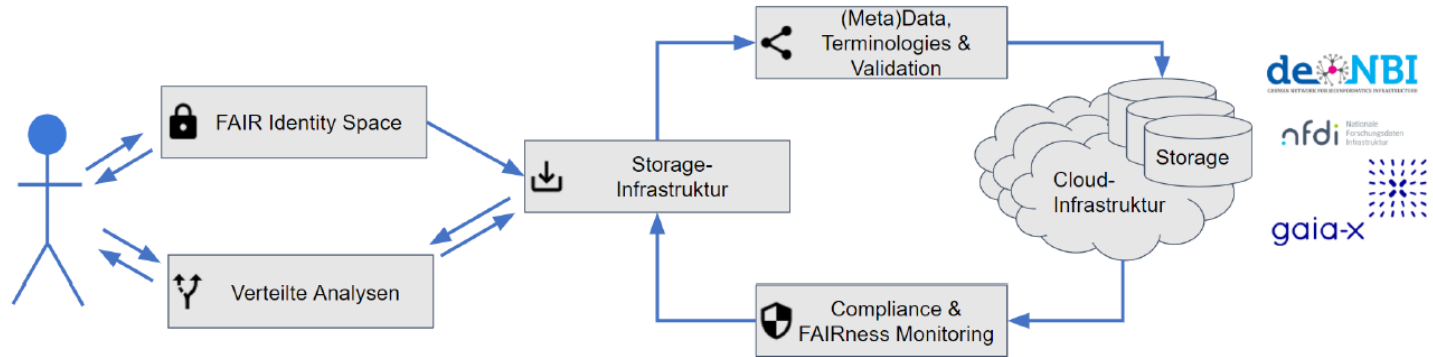
- 1. Roadmapping and Community:** ✓ Map of topics and actors, Application Scenarios und (✓) Requirements, Workshops and Conferences
- 2. Ethical and legal framework:** ✓ regular exchange with AP-Technological Basis, (✓) Consent Templates, legal compatibility with other data spaces (✓ e.g. Mobility Data Space)
- 3. Technological Basis**
 - Cloud-native Architecture, Mapping NFDI-Services to Gaia-X Federation Services (✓)
 - Storage, Monitoring of Compliance (and FAIRness), Identity, Metadata and Terminologies, Distributed Analysis
- 4. Demonstrators** – (✓) first version End 2021: Gaia-X-compatible description
 1. Biodiversity (NFDI4Biodiversity)
 2. Research Data Quality (NFDI4Ing)
 3. Cross Plattform Data Analysis (NFDI4Health and GO FAIR Personal Health Train)
 4. Future extension with more demonstrators in 2022 and 2023/24: (✓) Market research, (✓) Discussions
 5. ELSA-Schulungen für Data Scientists

Challenges for all WPs: The Gaia-X specification is still a moving target (Policy Rules, Federation Services, ...)!

Structure of the “FAIR Data Space”



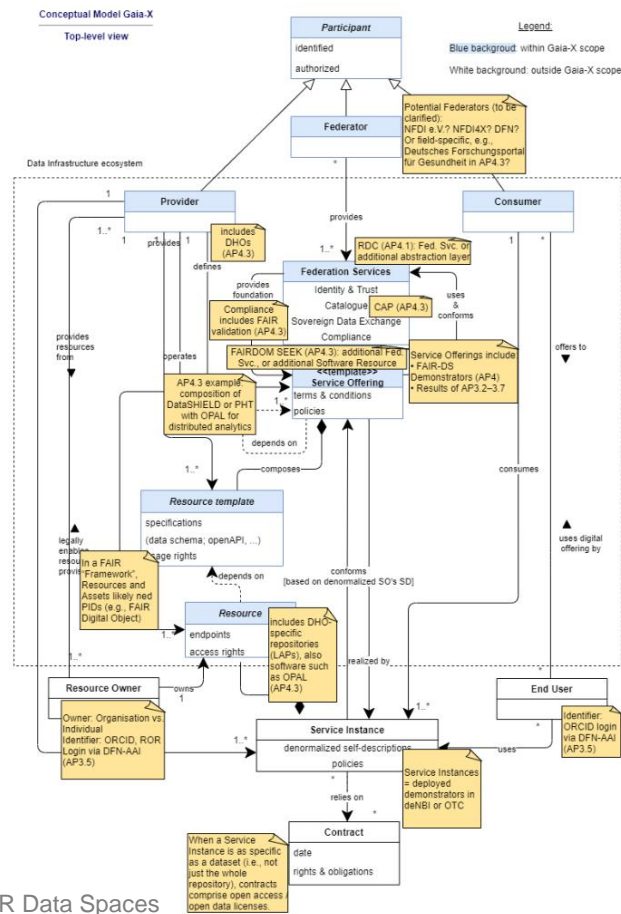
A FAIR Data Life Cycle for Gaia-X



Common Architecture for FAIR Data Spaces

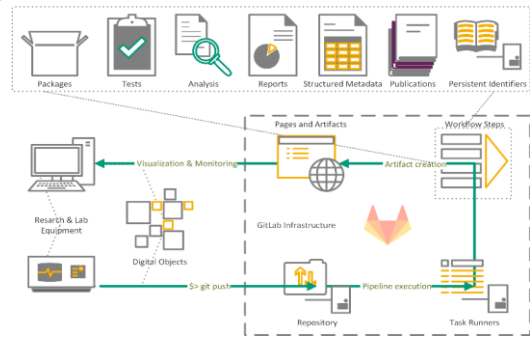
- NFDI Services are mapped to the Gaia-X conceptual Model
 - First version based on Gaia-X 21.09
 - Next iteration in March 2022

- Continuously approaching standards in
 - NFDI-Section Common Infrastructures
 - EOSC Interoperability Framework



FAIR Data Spaces Demonstrators Phase 1 (End 2021)

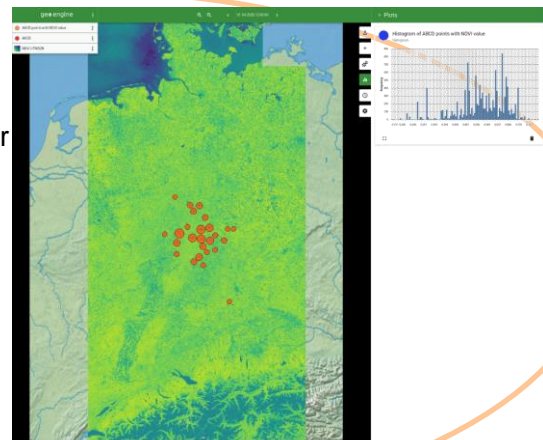
Available services within the NFDI consortia, slightly extended and described conforming to Gaia-X



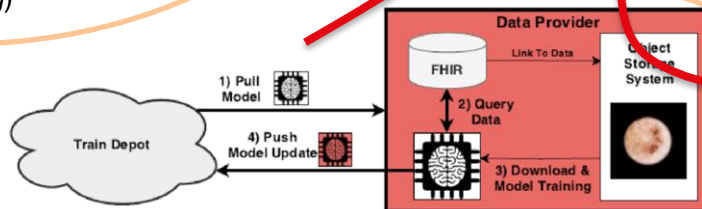
↑ FAIR Research Data Quality Assurance and Workflows (NFDI4Ing)

gaia-x

FAIR-DS Demonstrator
NFDI4Biodiversity and
Gaia-X →

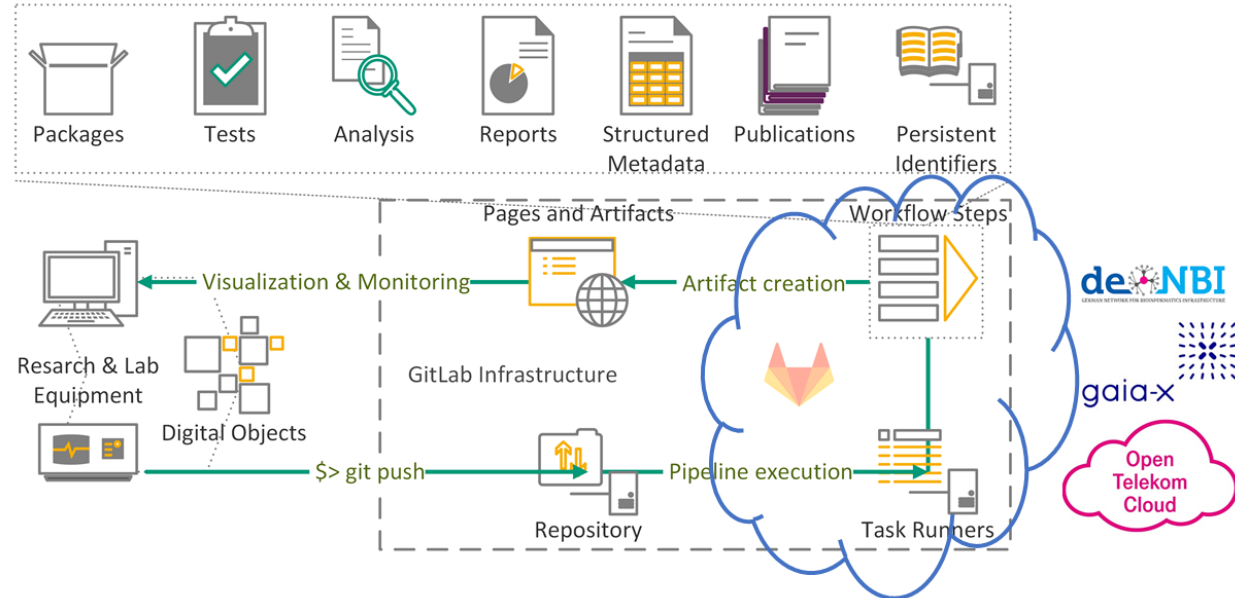


Cross-Plattform FAIR
Data Analysis
(NFDI4Health) →



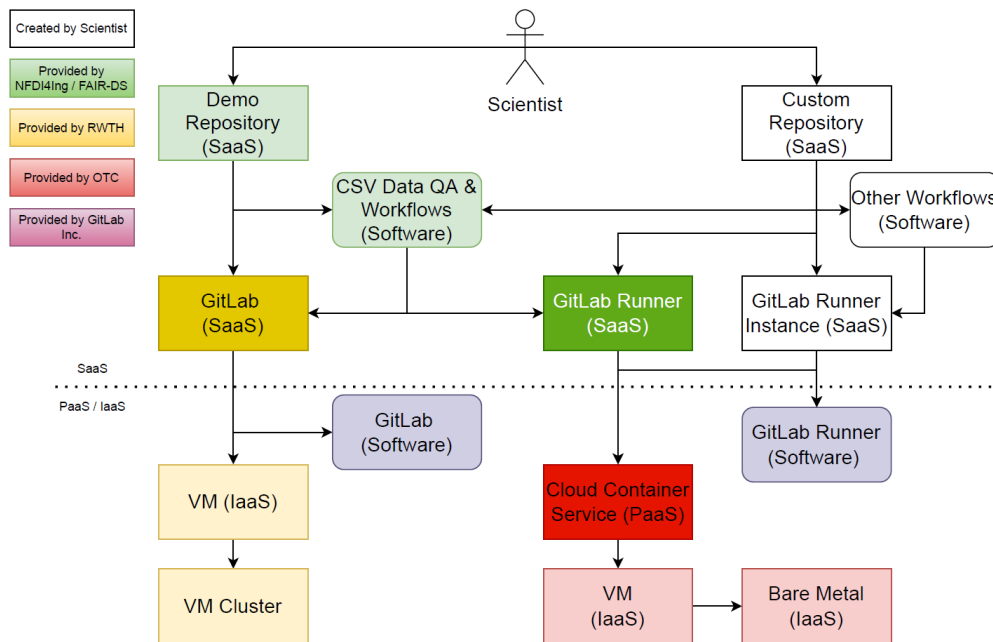
Starting May 2022: Onboarding more demonstrators connecting science and industry

Gaia-X Conformant Service by Example – The NFDI4Ing Demonstrator



Gaia-X Conformant Modelling of the Demonstrators

1. Create architectural Model



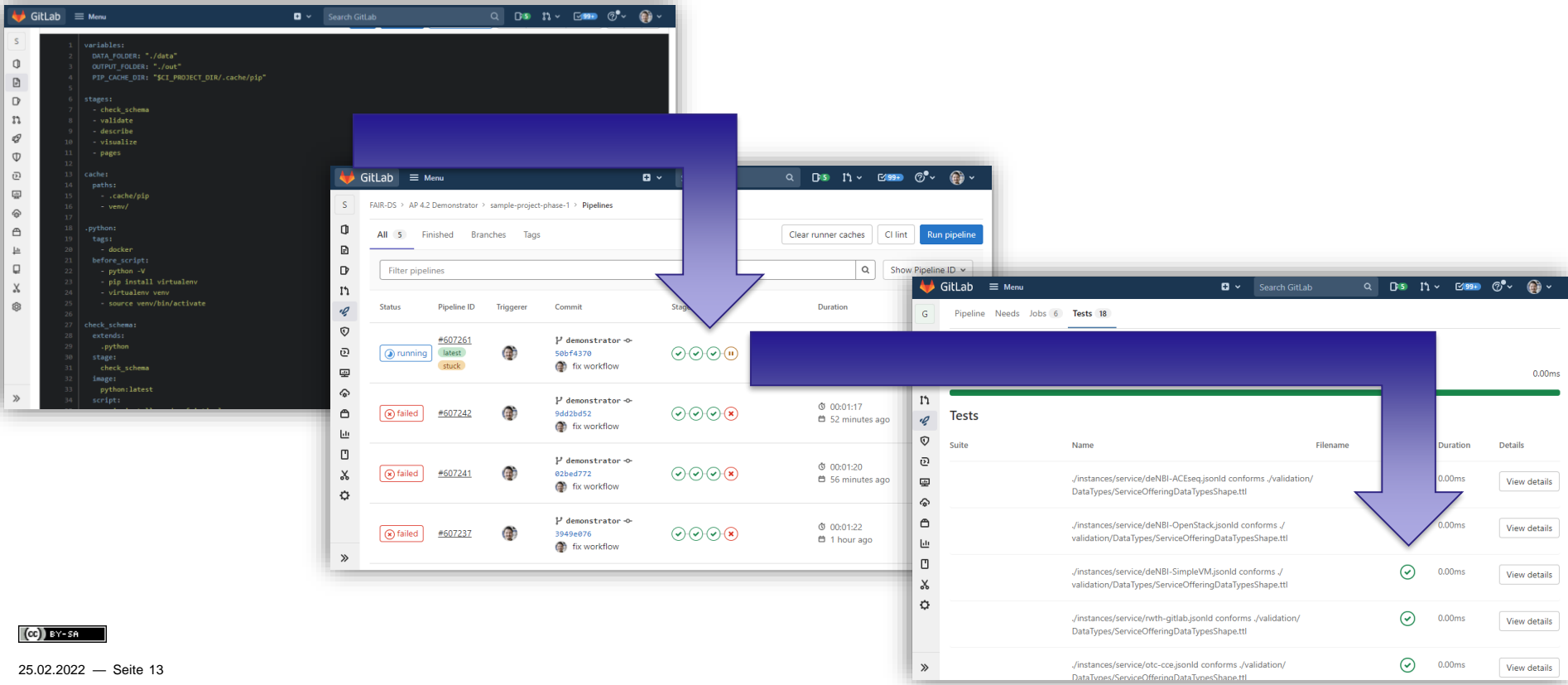
2. Create self description

```

1 {
2   ... "@context": {
3     ... "dct": "http://purl.org/dc/terms/",
4     ... "dcat": "https://www.w3.org/ns/dcat#",
5     ... "foaf": "http://xmlns.com/foaf/0.1/",
6     ... "ids": "https://w3id.org/idsa/core/",
7     ... "owl": "http://www.w3.org/2002/07/owl#",
8     ... "rdf": "http://www.w3.org/1999/02/22-rdf-syntax-ns#",
9     ... "rdfs": "http://www.w3.org/2000/01/rdf-schema#",
10    ... "schema": "http://schema.org/",
11    ... "vcard": "https://www.w3.org/2006/vcard/ns#",
12    ... "xsd": "http://www.w3.org/2001/XMLSchema#",
13    ... "gax-core": "http://w3id.org/gaia-x/core#",
14    ... "gax-participant": "http://w3id.org/gaia-x/participant#",
15    ... "gax-service": "http://w3id.org/gaia-x/service#",
16    ... "ex": "http://example.org/"
17  },
18  ... "@id": "ex:rwth-gitlab-service",
19  ... "@type": "gax-service:ServiceOffering",
20  ... "gax-service:hasServiceTitle": "GitLab of RWTH Aachen University",
21  ... "gax-service:hasVersion": "0.1",
22  ... "gax-service:providedBy": {
23    ... "@id": "ex:rwth",
24    ... "@type": "gax-participant:Provider"
25  },
26  ... "dcat:keyword": [
27    ... "Version Control",
28    ... "Source Code",
29    ... "Workflows",
30    ... "Devops"
31  ],
32  ... "gax-service:hasMarketingImage": {
33    ... "@value": "https://api.app.rwth-aachen.de/v3/sabio/app/image?path=0962a98",
34    ... "@type": "xsd:anyURI"
35  }
36 }

```

Putting Gaia-X into Practice for Data Quality Assurance



```
1 variables:
2   DATA_FOLDER: "/data"
3   OUTPUT_FOLDER: "/out"
4   PIP_CACHE_DIR: "${CI_PROJECT_DIR}/cache/PIP"
5
6 stages:
7   - check_schema
8   - validate
9   - describe
10  - visualize
11  - pages
12
13 cache:
14   paths:
15     - cache/PIP
16     - venv/
17
18 .python:
19   tags:
20     - docker
21   before_script:
22     - python -V
23   - pip install virtualenv
24   - virtualenv venv
25     - source venv/bin/activate
26
27 check_schema:
28   extends:
29     .python
30   stages:
31     - check_schema
32   image:
33     python:latest
34   script:
```

Status	Pipeline ID	Triggerer	Commit	Stages	Duration
running	#607261	latest	demonstrator 58bf4370	fix workflow	
failed	#607242	9dd2bd52	demonstrator 9dd2bd52	fix workflow	00:01:17 52 minutes ago
failed	#607241	e2be4772	demonstrator e2be4772	fix workflow	00:01:20 56 minutes ago
failed	#607237	3949e876	demonstrator 3949e876	fix workflow	00:01:22 1 hour ago

Suite	Name	Filename	Duration	Details
	./instances/service/deNBI-ACseq.jsonld conforms ./validation/DataTypes/ServiceOfferingDataTypesShape.ttl		0.00ms	View details
	./instances/service/deNBI-OpenStack.jsonld conforms ./validation/DataTypes/ServiceOfferingDataTypesShape.ttl		0.00ms	View details
	./instances/service/deNBI-SimpleVM.jsonld conforms ./validation/DataTypes/ServiceOfferingDataTypesShape.ttl		0.00ms	View details
	./instances/service/nwth-gitlab.jsonld conforms ./validation/DataTypes/ServiceOfferingDataTypesShape.ttl		0.00ms	View details
	./instances/service/otc-cce.jsonld conforms ./validation/DataTypes/ServiceOfferingDataTypesShape.ttl		0.00ms	View details

Want more details?

Register now!



Leitbild einer fairen Datenökonomie in Deutschland und Europa

Virtueller Diskurs
15. März 2022 ab 09:00 Uhr

Mit Vorträgen von

Bettina Stark-Watzinger, Bundesministerin für Bildung und Forschung,
Prof. Dr.-Ing. Reimund Neugebauer, Präsident der Fraunhofer-Gesellschaft
und vielen weiteren!



 **Fraunhofer**
IUK-TECHNOLOGIE

UNTER DER SCHIRMERSCHAFT VON
BETTINA STARK-WATZINGER MdB
 **Bundesministerium
für Bildung
und Forschung**