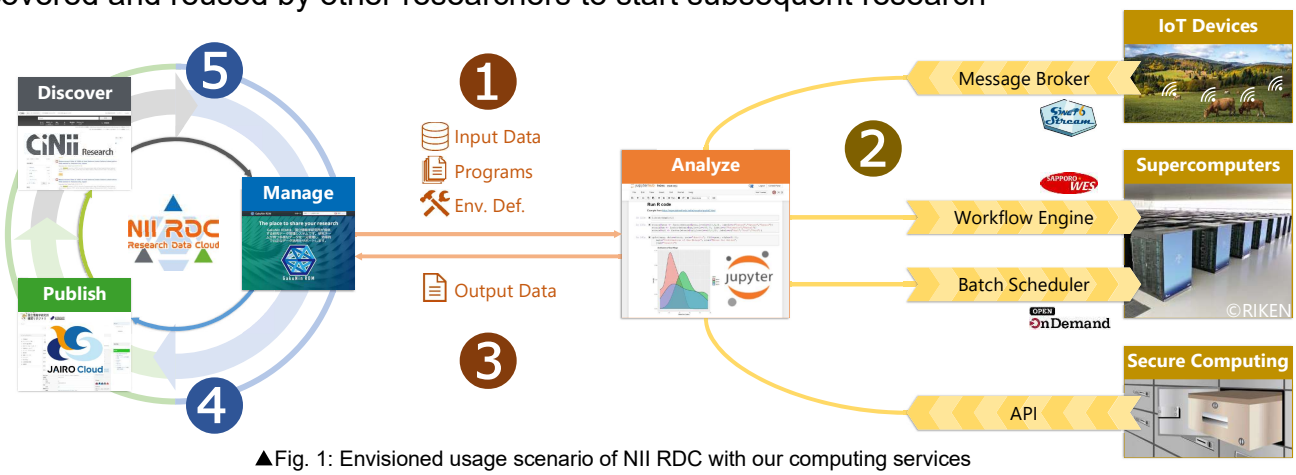
Research Data Analysis Platform for Reproducible Data-driven Science

Background

- To promote data-driven science, a high degree of integration between research data and computational resources is desired.
- NII is developing the national research data platform called NII Research Data Cloud (RDC)[1]. It consists of GakuNin RDM[2], an OSF[3]-based research data management service, JAIRO Cloud, an Invenio[4]-based repository service, and CiNii Research, a discovery service.
- CCRD is developing a variety of computing services that cooperate with NII RDC to facilitate the use of research data by diverse researchers.

Envisioned Scenario

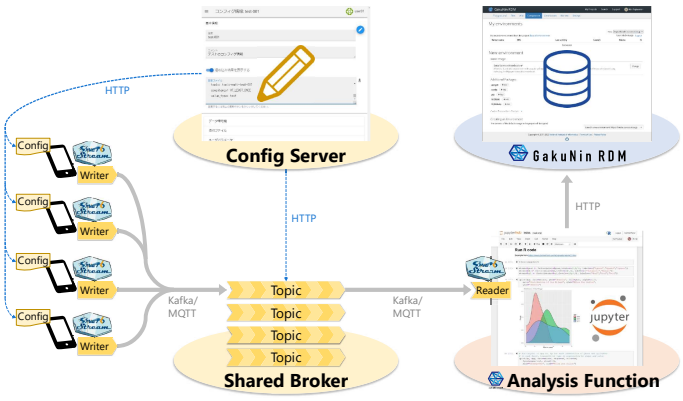
1. Auto-build a JupyterHub environment on the NII Cloud from within GakuNin RDM
2. Cooperate with a variety of data sources and computational resources
3. Share the analysis results with the project members in GakuNin RDM
4. Publish the project as a "computational reproduction package"
5. Discovered and reused by other researchers to start subsequent research



▲Fig. 1: Envisioned usage scenario of NII RDC with our computing services

Systems Development

- The SINETStream Trial Service provides a framework for researchers to collect and analyze streaming data without having to deploy a broker. [Fig. 2]
- The GakuNin RDM Data Analysis Function provides researchers with a Jupyter environment on NII Cloud. Researchers can run Python, R, and MATLAB programs to analyze their data stored in GakuNin RDM.
- The Reproducible Package Function archives a GakuNin RDM project in an RO-Crate format and makes it available for reuse by other researchers through JAIRO Cloud and CiNii Research.
- The GakuNin RDM App for Open OnDemand mounts a GakuNin RDM's storage from within a supercomputer. It is currently supported by Fugaku at RIKEN[6].



▲Fig. 2: Architecture of the SINETStream Trial Service

References

[1] https://rcos.nii.ac.jp/en/service/
[2] https://support.rdm.nii.ac.jp/en/
[3] https://www.cos.io/products/osf
[4] https://inveniosoftware.org/
[5] https://support.nii.ac.jp/en/cir/manual_outline
[6] https://www.r-ccs.riken.jp/outreach/topics/20230726-1/

National Institute of Informatics
https://ccrd.nii.ac.jp/sc24/