

第 22 回年次大会予稿

次世代リポジトリプラットフォームに求められる機能像

～オープンソースリポジトリソフトウェアの比較調査を通して～

Functional Image Required for Next Generation Repository Platform

～A Comparative Survey of Open Source Repository Software ～

林正治^{1*}, 林洋平¹, 田邊浩介², 青山俊弘³, 池田大輔⁴, 行木孝夫⁵, 山地一禎¹
Masaharu HAYASHI^{1*}, Yohei HAYASHI¹, Kosuke TANABE², Toshihiro AOYAMA³,
Daisuke IKEDA⁴, Takao NAMIKI⁵, Kazutsuna YAMAJI¹

1 国立情報学研究所

National Institute of Informatics

〒101-8430 東京都千代田区一ツ橋 2-1-2

2 物質・材料研究機構

National Institute for Materials Science

〒305-0047 茨城県つくば市千現 1-2-1

3 鈴鹿工業高等専門学校

National Institute of Technology, Suzuka College

〒510-0294 三重県鈴鹿市白子町

4 九州大学

Kyushu University

〒819-0395 福岡市西区元岡 744

5 北海道大学

Hokkaido University

〒060-0808 北海道札幌市北区北 8 条西 5 丁目

*連絡先著者 Corresponding Author

近年、高等教育機関・研究機関における機関リポジトリの普及に伴い、機関リポジトリが持つ可能性とその活用に向けた議論が盛んである。その議論の中心にある、オープンアクセスリポジトリ連合（COAR）の次世代リポジトリ WG では、世界中に分散した機関リポジトリを地球規模の学術コミュニケーション・ネットワークとして位置付け、新たな付加価値サービスの展開を想定したユースケース及び技術の検討を行っている。我々の研究グループでは、こうした状況を鑑み、主要なオープンソースリポジトリソフトウェアの技術比較を行っ

てきた。本発表では、技術比較結果の考察および技術史的な視点から、次世代リポジトリソフトウェアに求められる技術的な機能像を明らかにする。

キーワード: 次世代リポジトリプラットフォーム, 機関リポジトリ, 技術史

next generation repository platform, institutional repositories, technical history

1. はじめに

近年、高等教育機関・研究機関における機関リポジトリの普及に伴い、機関リポジトリが持つ可能性とその活用に向けた議論が盛んである。国際的な取り組みである、オープンアクセスリポジトリ連合(COAR)では、次世代リポジトリのためのワーキンググループ(NGR WG)を立ち上げ、次世代リポジトリの機能要件及び技術検討を進めている[1]。我々の研究グループでは、こうした状況を鑑み、次世代リポジトリソフトウェアに求められる技術的な機能像を明らかにすることを旨として、主要なオープンソースリポジトリソフトウェアの技術比較を行ってきた。これまでもリポジトリソフトウェアの比較研究は実施されているが[2][3][4]、抽象化された機能レベルの比較であり、システムアーキテクチャに着目したものではない。我々はシステムアーキテクチャに着目することで、次世代リポジトリプラットフォームに求められる機能像を明らかにすることを目的とする。

2. リポジトリソフトウェアの歴史

リポジトリソフトウェアを取り巻く状況は刻々と変化している(表1)。2000年初頭はEPrints[5]の開発が先行していたが、徐々にDSpace[6]やFedora[7]陣営(Islandora[8], Samvera[9], Hyku[10],

Hyrax[11])に開発の主導権が移っている。2008年頃からDataDryad.org[12]、Figshare[13]やZenodo[14]などリポジトリサービスが展開されていることも興味深い。また、2012年のJAIRO Cloud[15]以降、DSpaceDirect[16]、HykuDirect[17]、Figshare for Institutions[18]など機関リポジトリ機能を提供するSoftware as Service(SaaS)の展開が広がっていることは興味深い。

表1 リポジトリソフトウェアの歴史

年月	事項
2000/11	EPrints 1.0
2001/01	OAI-PMH 1.0
2002/02	EPrints 2.0
2002/06	OAI-PMH 2.0
2002/08	CDSWare(後のInvenio)
2002/11	DSpace 1.0
2003/05	Fedora 1.0
2003/12	Invenio 0.1.0
2005/01	Fedora 2.0
2005/05	OpenSearch 1.0
2006/10	Islandora initial version
2007/10	SWORD 1.0
2007/12	Dataverse Network 1.0
2008/01	DataDryad.org
2008/03	DSpace 1.5.0
2008/03	WEKO 1.0.0
2008/07	Fedora 3.0
2008/09	EPrints 3.1
2009/02	ActiveFedora 1.0.0
2010/01	Hydra 0.1
2010/03	DSpace 1.6.0
2010/03	EPrints 3.2
2010/12	DSpace 1.7.0
2011/01	Figshare 設立
2011/04	HUBzero 1.0

2011/06	EPrints 3.3.1
2011/10	Islandora 6.x-11.2.0
2011/11	DSpace 1.8.0
2012/03	Invenio 1.0.0
2012/03	SWORD 2.0
2012/04	JAIROCloud 開始
2012/05	WEKO 2.0.0
2012/06	Islandora 7.x-1.0
2012/11	DSpace 3.0
2013/03	Zenodo 開始
2013/09	Figshare for Institutions
2013/12	DSpace 4.0
2014/03	DSpaceDirect 開始
2014/04	ResourceSync Framework Specification (ANSI/NISO Z39.99-2014)
2014/11	Fedora 4.0
2015/01	DSpace 5.0
2015/03	Invenio 2.0.0
2015/05	Dataverse 4.0
2015/09	HUBzero 2.0.0
2015/11	Invenio v.3.0.0-a1
2016/09	ActiveFedora 11.0.0
2016/10	DSpace 6.0

2016/11	Hyku 0.1.0
2017/05	Hyrax 1.0.0
2017/06	HykuDirect パイロットサービス開始
2017/07	WEKO 2.4.0

3. システムアーキテクチャの比較

既存リポジトリソフトウェアのシステムアーキテクチャを比較調査結果から最近のリポジトリソフトウェアには2つの傾向が観察できる(表2)。1つ目は全文検索エンジンの利用である。従来のデータベース検索と比較して、速度面での性能向上が得られるからと考えられる。2つ目は、マイクロサービスアーキテクチャの採用である。スケーラビリティや機能拡張を考慮する上で有利な点があるからと考えられる。Islandoraの次期バージョン[19]もマイクロサービスアーキテクチャを採用することが宣言されており、次世代リポジトリの重要な機能に位置付けられるものと考えられる。

表2 アーキテクチャ比較

Name	Base	CMS	Metadata Storage	Full Search	Text	OR Mapping	Messaging
DSpace	Cocoon	-	PostgreSQL	Solr		Hibernate	-
EPrints	-	-	MySQL	-		-	-
HUBzero	Joomla!	Joomla!	MySQL	-		-	-
Invenio v3	flask		PostgreSQL	Elasticsearch		SQLAlchemy	Rabbitmq / Celery
Hyrax/ Hyku	Samvera+ Fedora	-	Fedora	Solr		ActiveFedora	Redis / Sidekiq
Islandora	Drupal+ Fedora	Drupal	Fedora	Solr		ActiveFedora	-
Fedora	ModeShape	-	JDBC	-		-	-
Dataverse	-	-	PostgreSQL	Solr		-	-
WEKO	Maple	NetComm ons2	MariaDB	Mroonga		-	-

4. まとめ

本稿では、次世代リポジトリプラットフォームに求められる機能像を明らかにすることを目的として、リポジトリソフトウェアの歴史の変遷の観察、システムアーキテクチャの技術的比較を実施した。その結果、次世代を検討する上で、SaaS への展開、全文検索エンジンおよびマイクロサービスアーキテクチャの採用が重要である可能性が高いことがわかった。今後はこの結果を踏まえ、各リポジトリソフトウェアについてより詳細な比較検討を実施し、次世代リポジトリプラットフォームに必要な技術的な仕様の抽出を行ないたい。

参考文献

- [1] COAR Next Generation Repositories WG: Next Generation Repositories, <https://www.coar-repositories.org/files/COAR-Next-Generation-Repositories-February-7-2017.pdf>
- [2] Masrek, Mohamad & Hakimjavadi, Hesamedin. (2012). Evaluation of Three Open Source Software in Terms of Managing Repositories of Electronic Theses and Dissertations: A Comparison Study. Basic and Applied Scientific Research. 2. 10843-10852.
- [3] A Comparative Review of Various Data Repositories, <https://dataverse.org/blog/comparative-review-various-data-repositories>
- [4] Jean Gabriel Bankier & Kenneth Gleason: Institutional repository software comparison. In: UNESCO Communication and Information, vol. 33 (2014)
- [5] EPrints, <http://www.eprints.org/> (2017年10月27日参照)
- [6] DSpace, <http://www.dspace.org/> (2017年10月27日参照)
- [7] Fedora, <http://fedora-repository.org/> (2017年10月27日参照)
- [8] Islandora, <https://islandora.ca/> (2017年10月27日参照)
- [9] Samvera, <https://samvera.org/> (2017年10月27日参照)
- [10] HYDRA-IN-A-BOX, <http://hydrainabox.samvera.org/> (2017年10月27日参照)
- [11] Hyrax, <http://hyr.ax/> (2017年10月27日参照)
- [12] Dryad Digital Repository, <http://datadryad.org/> (2017年10月27日参照)
- [13] Figshare, <https://figshare.com/> (2017年10月27日参照)
- [14] Zenodo, <https://zenodo.org/> (2017年10月27日参照)
- [15] JAIRO Cloud, <https://community.repo.nii.ac.jp/> (2017年10月27日参照)
- [16] DSpace direct, <http://dspace-direct.org/> (2017年10月27日参照)
- [17] Hykudirect, <https://hykudirect.com/> (2017年10月27日参照)
- [18] Figshare for Institutions, <https://figshare.com/services/institutions> (2017年10月27日参照)
- [19] Islandora CLAW, <https://islandora.ca/CLAW> (2017年10月27日参照)