

[R2] IR実務担当者セッション

九州大学教員活動進捗・報告システムと researchmapデータ連携システムの開発

九州大学 インスティテューショナル・リサーチ室 松本 馨
九州大学 企画部企画課分析係 戸川 忠嗣

2020年11月5日



九州大学



1. はじめに

2

九州大学：研究情報データベース

- ・ 組織の研究評価：外資系論文DB
 - Scopus (Elsevier)
 - ・ 教員の業績評価：学内DBを自前構築
 - 九州大学教員活動進捗・報告システム (Q-RADeRS)
 - ・ 論文データ入力支援用にPure (Elsevier) も利用
 - 研究者情報 (研究者総覧) ページも自動生成
 - ・ researchmap (JST)
 - 学内DBとあわせて複数のDBへの登録をしている
 - 2020年2月にV2公開
 - ・ 科研費審査時に参照される
- ↓
- ・ 学内DBとresearchmapデータ連携システムを開発
 - 業績登録の手間を削減、業績データ保管場所として利用

科学技術振興機構 (JST) : researchmap

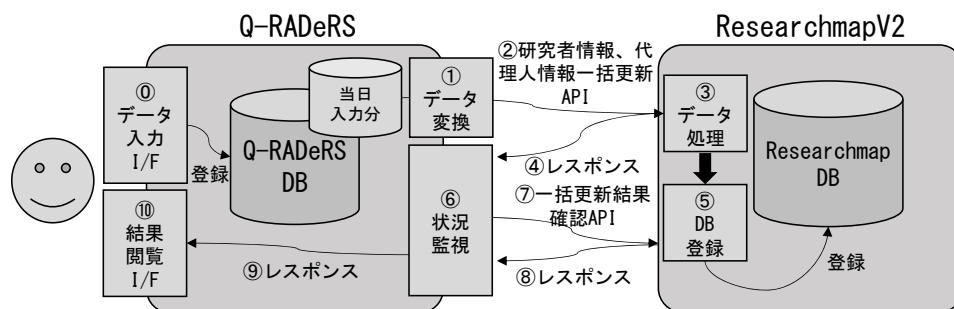
- researchmap.V1
 - XML形式のデータ
 - 参照系Web-API(読み込み)
 - データファイルダウンロード・アップロードによる全件更新
- researchmap.V2
 - JSON形式のデータ
 - 参照系Web-API(読み込み)
 - 更新系Web-API(書き込み)
 - 任意のタイミングで部分更新(データ追加)



ファイル更新から、システム間連携による常時更新

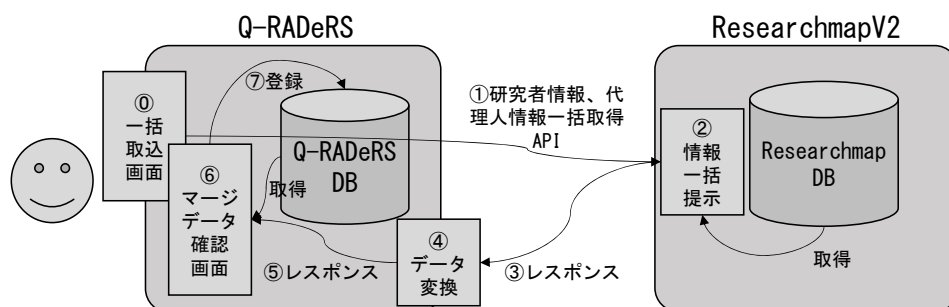
九州大学: Web-APIによるデータ更新を実装

★エクスポート(Q-RADeRS → researchmapにデータ転送)



- **即反映されない**
- 必須項目等が未入力
で登録が拒否
- **重複チェック**で登録
が拒否(**データ連携
を試さないと登録の
可否が不明**)
- データ登録、編集の
制約が多い

★インポート(researchmap → Q-RADeRSにデータ転送)



- (公開準備中)
- **即反映される**
- 必須項目等が未入力
で登録が拒否
- データを取得した後
はこちらで独自の処
理ができる

Q-RADeRS→researchmapエクスポート画面例

Web-APIによる更新

連携依頼

データの最新更新日時	Researchmapへの最新連携依頼日時	連携結果	連携エラー内容	Researchmapへ反映 ON/OFF
2020-08-05 16:01:56	2020-08-05 16:03:46	完了		<input type="checkbox"/>
2020-06-22 09:38:49	2020-06-22 09:39:20	完了		<input type="checkbox"/>
2020-05-19 10:56:55	未連携			<input type="checkbox"/>
2019-10-29 14:41:02	2020-03-11 14:00:03	失敗	本人相当が登録した類似の業績が存在したため、更新しませんでした[id=12103304].	<input type="checkbox"/>
2019-10-29 14:40:28	未連携			<input type="checkbox"/>

2. 学内DBとresearchmapのデータ定義

データ定義の思想が違う

- researchmapには学生指導に関する項目がない (授業科目しか定義していない)
- Q-RADeRSには個人の経歴に関する項目がない (自組織での業績しか対象にしていない)
- 定義が一对一の対応をしていない
連携時にどの区分に入れるか判断が必要

共通定義のみデータ連携し、それ以外は各DBにデータを保持
→学内DBを持つことが必須

当日投影

定義の違い: 必須 / 任意項目

- 例: 論文の「出版年月」が必須 / 任意
- Q-RADeRSでは任意: 出版は確定、しかし何時出版されるかが分からないものを登録したい
 - 許容した結果(2019年10月時点):
216,678件中2,908件(約1.3%)が未登録(1,235件が現職)
- researchmapでは必須のため、そのままでは登録拒否



連携時にエラーで止めて、**ユーザーに修正してもらう**

定義の違い: 選択肢 / 数値 / 文字列

- 例: 授業科目の開講時期
 - 前期、後期、春学期、夏学期、秋学期、冬学期、通年
 - researchmapは開始年月～終了年月
 - 論理的には正しいが、年月でやっている大学はある?
 - 年月を入れないと、1年に複数回開講している授業は、同じ年度に同じ名前のもが羅列されてしまう



- 推測して年月を設定?
 - 前期: 4月～9月、後期: 10月～3月
 - 春: 4～6月、夏: 7～9月、秋: 10～12月、冬: 1～3月
- 科目名に学期を()書きで付与する?



未対応とした

選択肢のマッピング: ほぼ対応出来た例

Q-RADeRS 教育対象	researchmap @type:teaching_experience/teaching_experience_type
NULL	"":未設定
基幹教育	"undergraduate_liberal_arts":学部教養科目
学部教育	"undergraduate_special_subjects":学部専門科目
大学院基幹教育	"graduate_school_liberal_arts_course":大学院教養科目
大学院教育	"postgraduate_courses":大学院専門科目
	"others":その他

双方向の連携をしても、問題がほとんど無い

選択肢のマッピング: 部分対応になった例

Q-RADeRS 執筆形態(単著等の区分)	researchmap @type:books_etc/book_owner_role
NULL	"":未設定
単著	"single_work":単著
共著	"joint_work":共著
	"single_translation":単訳
	"joint_translation":共訳
編集	"editor":編者(編著者)
	"joint_editor":共編者(共編著者)
監修	"supervisor":監修
	"contributor":分担執筆
	"translation_editing":編訳
	"compilation":編纂
	"others":その他

* Q-RADeRS→researchmap
情報の劣化は起きない

* researchmap→Q-RADeRS
未対応の項目がある
必須項目ではないのでデータ入力をして
ない(NULLにすること)で問題回避

Q-RADeRS→researchmap
は問題無いが
researchmap→Q-RADeRS
は問題になる

選択肢のマッピング:階層構造が違った例

Q-RADeRS 学会大会・会議・シンポジウム等における役割	researchmap @type:academic_contribution/academic_contribution_type
	"":未設定
	"academic_society_etc":学会・研究会等
→	"competition_etc":大会・シンポジウム等
	"exhibition":展覧会
	"review":審査・学術的助言
	"academic_research":学術調査
	"peer_review_etc":査読等
	"cultural_property_protection":文化財保護
	"others":その他

* Q-RADeRSには「学会大会・会議・シンポジウム等における役割」という業績区分があるが、researchmapでは「学術貢献活動」の「学術貢献活動種別」の選択肢に「"competition_etc":大会・シンポジウム等」がある。

→学術貢献活動の1つとして登録し、選択肢で「大会・シンポジウム等」を選択する。

選択肢化されていない:自由記述欄を利用した例

Q-RADeRS 科学研究費補助金(文部科学省、日本学術振興会)の採択状況	researchmap 共同研究・競争的資金等の研究課題 @type:research_projects
「日本学術振興会」を入れる	提供機関(日本語)
「科学研究費助成事業」を入れる	制度名(日本語)
特別推進研究/新学術領域研究/基盤研究(S)/基盤研究(A)/基盤研究(B)/基盤研究(C)/挑戦的萌芽研究/挑戦的研究(開拓)/挑戦的研究(萌芽)/若手研究(A,B)/若手研究/研究活動スタート支援/奨励研究/特別研究促進費/研究成果公開促進費/特別研究員奨励費/国際共同研究強化(A)/国際共同研究強化(B)/国際活動支援班/帰国発展研究/学術創成研究費/萌芽研究/特別研究/がん特別研究/重点領域研究/総合研究(A)/総合研究(B)/一般研究(A)/一般研究(B)/一般研究(C)/奨励研究(A)/奨励研究(B)/試験研究(A)/試験研究(B)/国際学術研究/特定奨励費/創成的基礎研究費/COE形成基礎研究費/地域連携推進研究費/特定領域研究/特定領域研究(A)/特定領域研究(B)/特定領域研究(C)/萌芽的研究/若手研究(S)/若手研究(A)/若手研究(B)/若手研究(スタートアップ)/新学術領域研究(研究領域提案型)のいずれかを入れる	研究種目(日本語)

* Q-RADeRSには「科学研究費補助金(文部科学省、日本学術振興会)の採択状況」という業績区分があるが、researchmapでは「共同研究・競争的資金等の研究課題」にまとめられている。「提供機関」「制度名」「研究種目」が自由記述欄になっており、選択肢化されていない。しかし、研究種目によって取得の難易度が大きく異なり、評価のために重要な情報である。

→エクスポートでは、文字列として登録する

(インポートでは、選択肢を提示してユーザーに選んでもらう)

どうやって二重登録を避けるか

- 管理しないと連携する度に同一業績が登録
- researchmapにはAI機能が実装
 - 業績データをWeb-APIで投げ込めば自動で重複判定を行い、処理してくれる
 - 学内DBもこれに併せて整理することを検討



- AI機能の自前構築は困難
 - データ登録時にresearchmapから返却される情報を利用
 - 登録成功時に返却されるachievement_id
 - 連携した業績データと対応
 - 登録失敗時に返却されるmessage_descriptionにあるid
 - 重複判定された業績データと対応

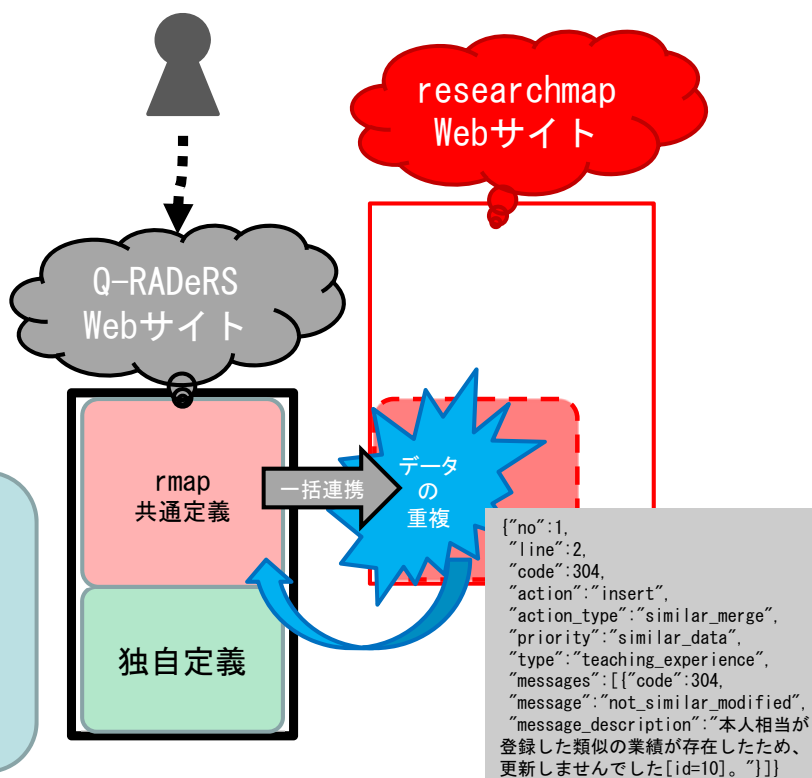
achievement_idの紐付け

ID	Q-RADeRS	ID	researchmap
1	業績A	?	業績A
2	業績B		
		?	業績C

ID	Q-RADeRS	ID	researchmap
1	業績A	10	業績A
2	業績B	12	業績B
		?	業績C

- 登録成功時にIDを取得
- 重複判定時のJSONエラーメッセージよりIDを取得 > 本人相当が登録した既存データが存在したため、更新しませんでした[id=10]

業績データエクスポート



achievement_idの紐付け

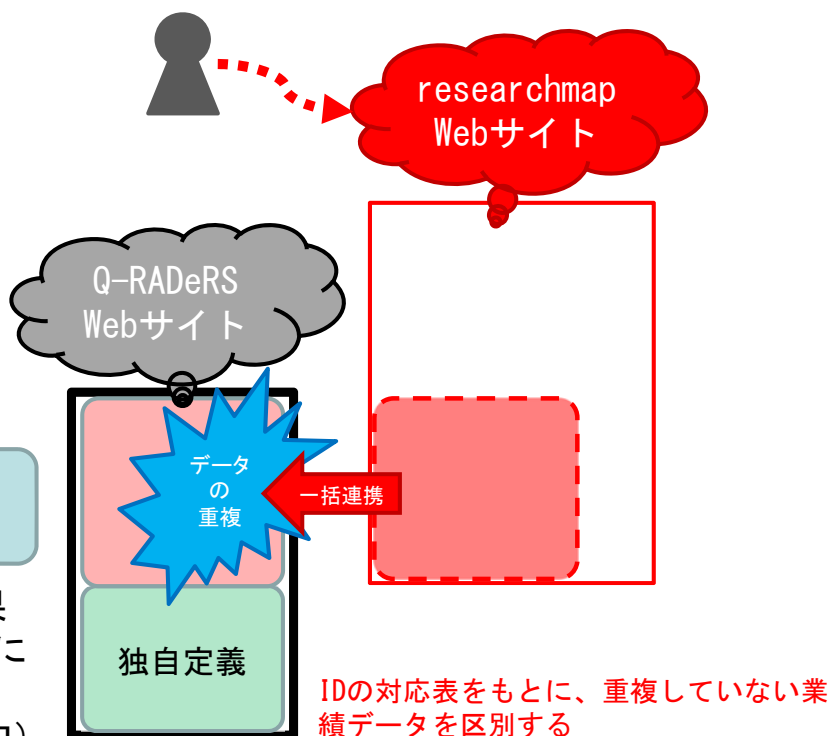
ID	Q-RADeRS	ID	researchmap
1	業績A	10	業績A
2	業績B	12	業績B
		?	業績C

ID	Q-RADeRS	ID	researchmap
1	業績A	10	業績A
2	業績B	12	業績B
3	業績C	11	業績C

Q-RADeRSに存在しないIDの業績だけを受け取る

researchmapによる重複判定結果を学内DBでも利用し、データ整理に活かす(インポート操作は手動。連携済の業績に○印を付けて案内)

業績データインポート

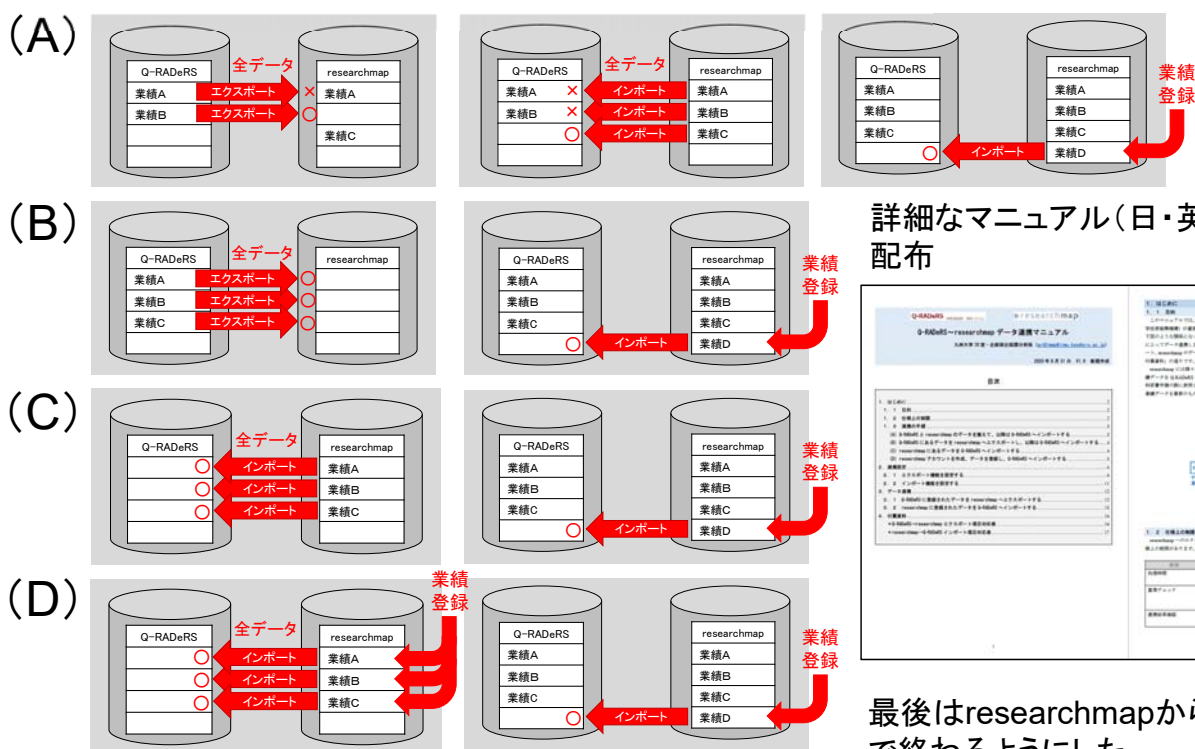


マニュアル化

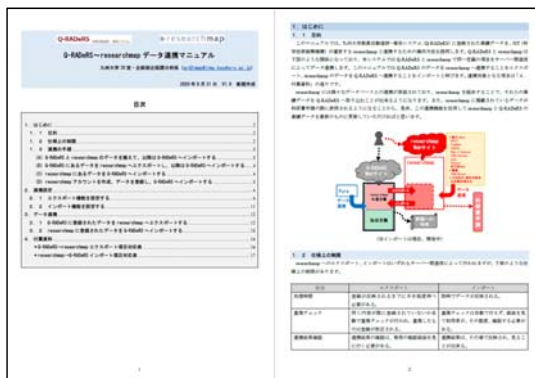
4通りのケースを想定して、全業績データが紐付けされた状態になるように手順を定めた

利用手順	Q-RADeRS	researchmap	状況
(A)	データ有り	データ有り	Q-RADeRSとresearchmapの両方に業績データを入力している
(B)	データ有り	データ無し	Q-RADeRSに業績データを入力しているが、researchmapは使っていなかった(または、researchmapアカウントを作成したが、ほとんど使わずに放置していた)
(C)	データ無し	データ有り	researchmapへ業績データを登録しており、これからQ-RADeRSへ業績データを登録する
(D)	データ無し	データ無し	Q-RADeRS、researchmapのどちらも初めて使う

マニュアル化: 4つの利用手順



詳細なマニュアル(日・英)を作成して配布



最後はresearchmapからのインポートで終わるようにした
→以後はresearchmapへの入力を主にしたい

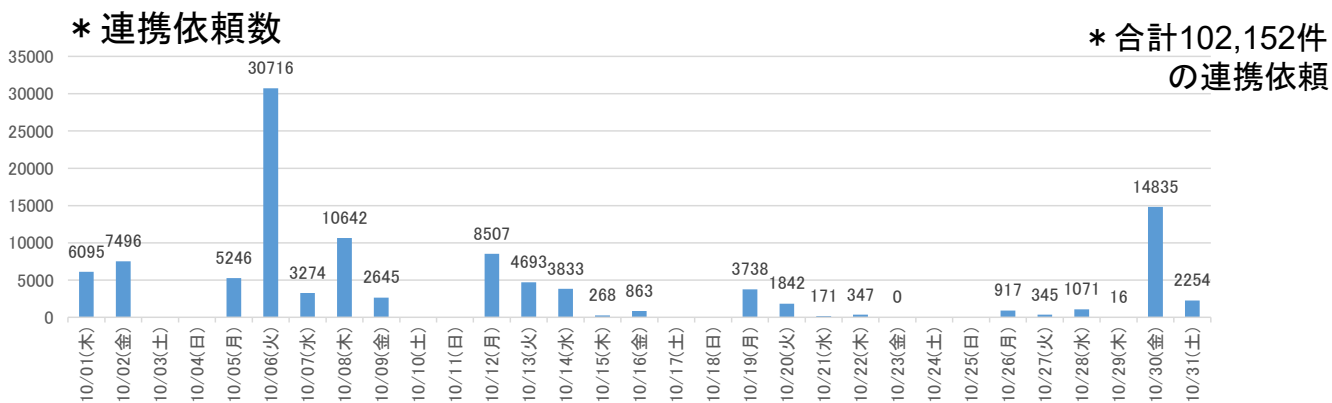
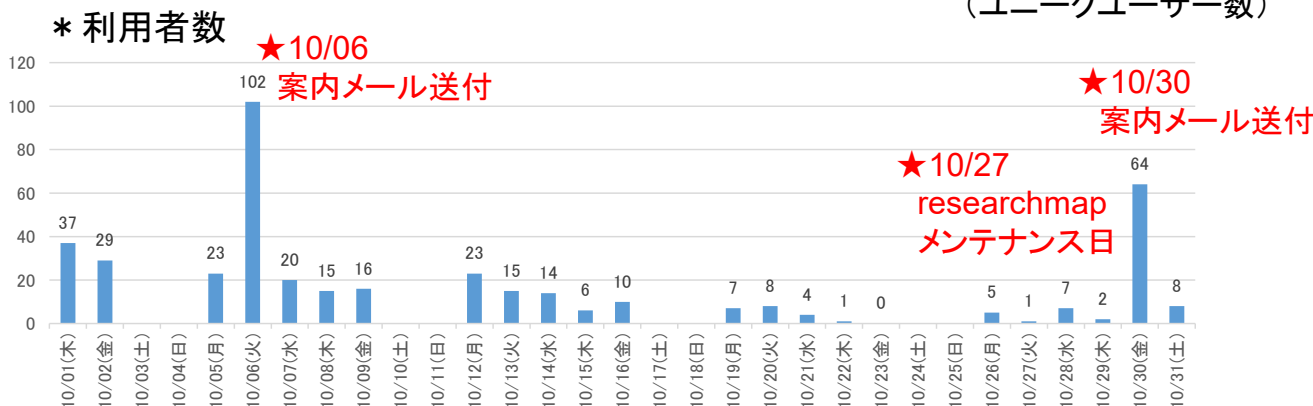
運用のスケジュール

● スケジュール

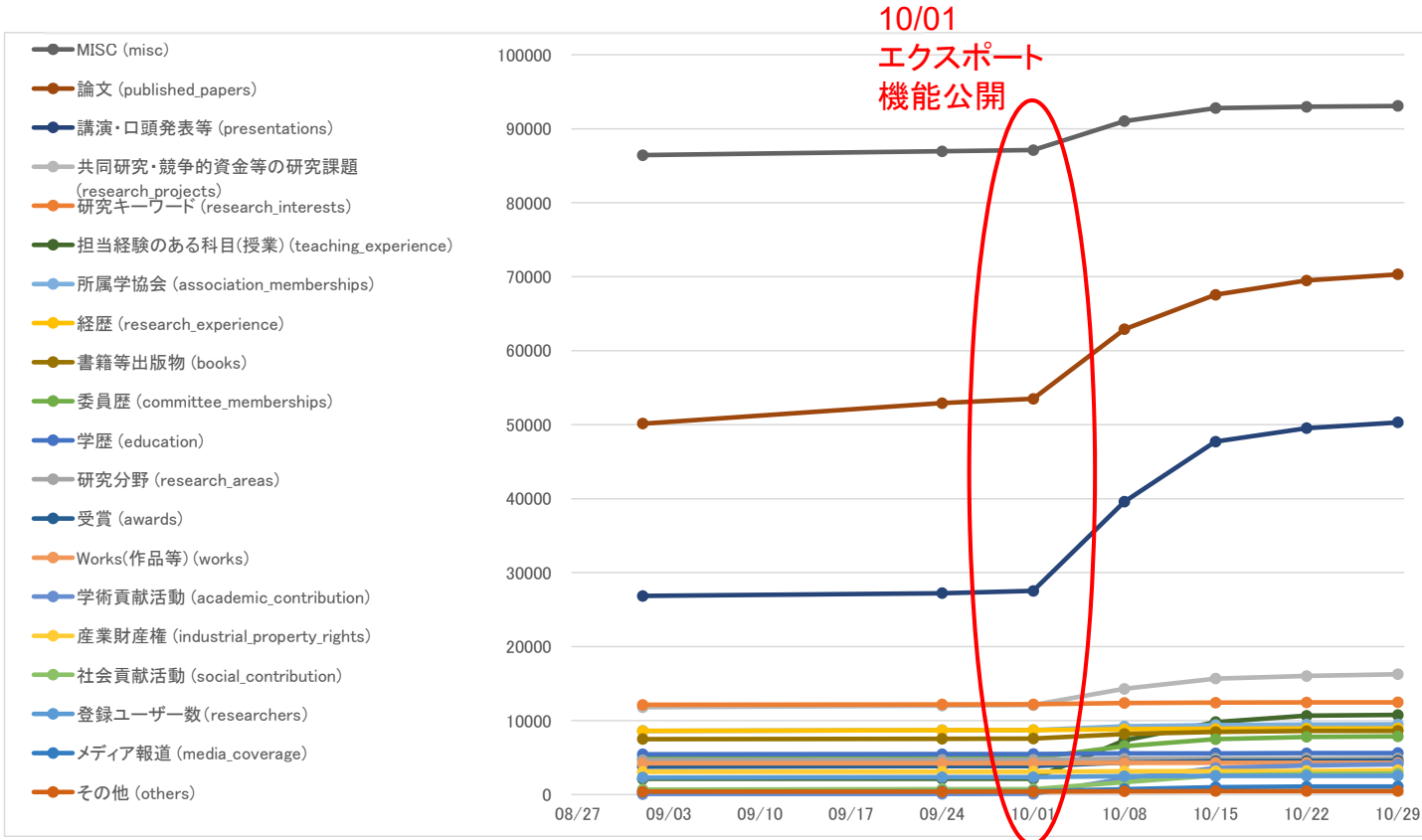
- 06月中旬: エクスポート機能運用テスト
- 08月31日: 連携機能公開の事前予告
- 09月01日: (科研費公募開始)
- 09月下旬: インポート機能運用テスト
- 10月01日: エクスポート機能公開
- 11月05日: (本日)
- 11月中旬: インポート機能公開予定

researchmapエクスポート数推移

* 335名(2,741名中)が利用
(ユニークユーザー数)



researchmap登録業績数推移



researchmap登録業績数推移

連携対象項目：
213,404 → 284,180
(+70,776レコード)

エクスポート
機能公開

	09/01	09/24	10/01	10/08	10/15	10/22	10/29	*10/01 との差異	*10/01 との比率
登録ユーザー数 (researchers)	2317	2357	2366	2481	2518	2534	2544	+178	1.08
研究キーワード (research_interests)	12129	12179	12206	12387	12437	12468	12476	+270	1.02
研究分野 (research_areas)	4646	4700	4714	4837	4868	4888	4906	+192	1.04
経歴 (research_experience)	8603	8696	8712	8832	8890	8937	8948	+236	1.03
学歴 (education)	5448	5470	5483	5568	5586	5601	5610	+127	1.02
委員歴 (committee_memberships)	4890	4942	4959	6563	7497	7806	7866	+2907	1.59
受賞 (awards)	3777	3827	3840	4333	4518	4543	4586	+746	1.19
論文 (published_papers)	50146	52905	53509	62897	67570	69498	70322	+16813	1.31
MISC (misc)	86436	86961	87124	91026	92785	92965	93090	+5966	1.07
書籍等出版物 (books)	7504	7546	7570	8196	8470	8602	8628	+1058	1.14
講演・口頭発表等 (presentations)	26852	27226	27532	39595	47713	49522	50321	+22789	1.83
担当経験のある科目(授業) (teaching_experience)	2125	2148	2153	7340	9805	10658	10762	+8609	5.00
所属学協会 (association_memberships)	8624	8712	8725	9219	9399	9465	9494	+769	1.09
Works(作品等) (works)	4221	4223	4223	4311	4297	4303	4305	+82	1.02
共同研究・競争的資金等の研究課題 (research_projects)	11802	12029	12102	14287	15685	16033	16275	+4173	1.34
産業財産権 (industrial_property_rights)	3076	3086	3086	3142	3182	3195	3283	+197	1.06
学術貢献活動 (academic_contribution)	72	93	94	2178	3579	3945	4112	+4018	43.74
社会貢献活動 (social_contribution)	671	732	731	1677	2594	2757	2850	+2119	3.90
メディア報道 (media_coverage)	361	443	445	736	999	1087	1092	+647	2.45
その他 (others)	393	395	397	452	470	474	477	+80	1.20

増加数最大

増加率最大
(シンポジウム司
会、査読業績等)

※赤字は連携対象外

researchmap登録エラー状況(10/01~10/31)

① Check	件数	割合(%)
成功	91,447	89.5
ローカルチェック(失敗)	8,750	8.6
システムエラー(失敗)	1,955	1.9
合計	102,152	100.0

※「ローカルチェック」は、データ送信前から正誤判断できる内容のエラー(必須項目がない、期間日付の始終が逆など)

② Update	件数	割合(%)
成功	85,756	94.0
researchmapチェック(失敗)	5,420	5.9
システムエラー(失敗)	15	0.0
合計	91,191	100.0

※「researchmapチェック」は、データ送信後でないと正誤判断できない内容のエラー(パーマリンクの間違い、機関担当者による編集権限を付与していないなど)

③ Complete	件数	割合(%)
成功	65,933	86.8
類似チェック、もしくは更新データなし	10,008	13.2
システムエラー(更新後処理)	42	0.1
合計	75,983	100.0

※「類似チェック、もしくは更新データなし」は、researchmapによる類似チェックまたは、既に最新のデータがあるため更新しなかったケース

researchmap登録エラー状況(10/01~10/31)

④ ReUpdate	件数	割合(%)
成功	17,984	99.8
システムエラー(再更新等)	36	0.2
合計	18,020	100.0

※通信エラーに起因するシステムエラー

⑤ ReComplete	件数	割合(%)
成功	8,631	48.0
類似チェック、もしくは更新データなし	9,346	52.0
システムエラー(再更新後処理等)	7	0.0
合計	17,984	100.0

- 7.3%(14,170/193,343)が「バリデーションエラー」(入力内容や記述内容が要件を満たしていない)
- ~~20.6%(19,354/93,967)が「類似チェック、更新データなし」で登録拒否~~
- 13.2%(10,008/75,983)が「類似チェック、更新データなし」で登録拒否



以後はインポート機能を公開、安定稼働を目指す

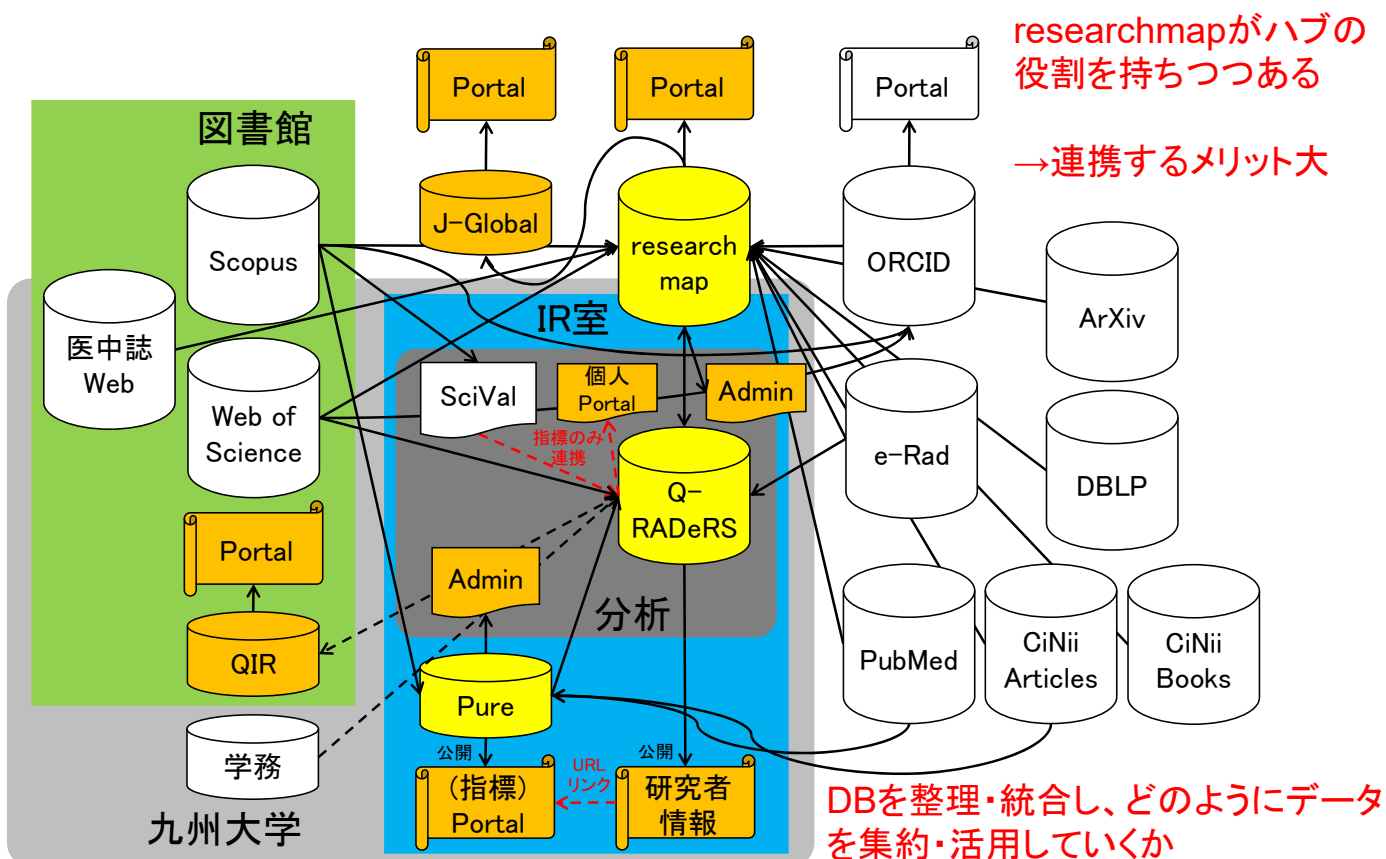
議論: researchmapのメリット/デメリット

- メリット
 - 無料で使える
 - データを投げ込めば勝手に整理してくれる
 - 評価等で同じ形式によるエビデンスが使える
 - 国の施策と連動してバリデーションが更新される?
 - 人社系の論文DBとして使える?
- デメリット
 - 研究評価で用いられる指標(Citations, FWCI, IF等)がない
 - データ定義の自己管理が出来ない
 - 例: 組織としては必須項目にしたいのに任意になっている等
 - 外国人研究者への対応
 - システム運用側との関係
 - 突然メンテナンス日が設定。事前連絡なく、急に仕様が変更

議論：データの持ち主、利用主体は誰か

- どのデータをマスタとすべきか
 - 組織としては学内DBをマスタとしたい
 - 組織内に閉じた(再利用できない)DBに真剣にデータを入れてくれるだろうか
 - 個人保有の公開されているDBならば、きちんとデータを入れる?
 - 勝手にresearchmapページを作られていたという話も散見
 - 組織でアカウントを作成したが、本人に伝わっていない
- 誰が主体となってデータを管理すべきか
 - 外部のDBでは、データ取り扱いの主導権を握られる懸念
 - 自学主導でのデータ定義が困難
 - 独立した学内DBを持つ必要がある
 - 外資系の高額なDBに頼るよりは・・・
 - 現在は、業界標準になっており使うしかない

議論：データをどこにまとめるか(九大のDB状況)



ご静聴ありがとうございました