

研究成果情報管理業務のシステム化による業務プロセスの改善

○林秀和^{A)}、山田一成^{A)}、田島嘉則^{A)}

^{A)} 情報通信技術支援室 情報基盤技術グループ

概要

2022年9月現在まで、名古屋大学 情報基盤センターはスーパーコンピュータ「不老」を計算資源として利用者へ提供することにより、社会貢献や成果創出を支援して来た。計算需要支援による研究力強化や、人材育成、社会連携のために、情報基盤センターはスーパーコンピュータ「不老」利用者の研究成果情報を管理・分析する必要がある。しかし、利用者が回答する研究成果情報アンケートのツールを含め、業務プロセスには改善の余地が大きかった。

本稿では、研究成果情報アンケートのツールとして Excel ファイルを用いた業務プロセスの改善方法を述べる。具体的には、利用者は Excel 形式の研究成果情報アンケートに回答し、それを情報基盤センターへ提出する。そして、提出された Excel ファイルの中から必要セルを抽出し、管理用 Excel ファイルへ半自動的に転記するシステムを構築した。

1 はじめに

我が国において人口減少及び少子高齢化が深刻化しており、労働力は中長期的に減少傾向である。限られた人的資源において労働生産性を向上させるために、業務へ ICT を積極的に導入する必要がある。すなわち、ICT の利活用によって、業務の省力化や効率化(より簡単に、より早く、より確実に)の実現に加え、既存サービスの高付加価値化や新規サービスの展開が求められる。しかし、総務省発行の令和3年版 情報通信白書によると、我が国では業務やデータの標準化が遅れており、業務効率化やデータ活用が進んでいないと述べられている^[1]。

ICT を推進するためには、その人材の確保と育成が必要である。しかし、2019年度に実施された独立行政法人情報処理推進機構の調査によると、IT 人材の量について「大幅に不足している」又は「やや不足している」という回答の合計は、89.0%に達している。また、IT 人材の質(セキュリティなどの高度な ICT スキル、デザイン思考、アジャイル開発のスキルの習熟度合い等)についても「大幅に不足している」又は「やや不足している」という回答の合計は、90.5%に達している。この要因は、情報システムに対するコスト削減圧力や、システムの構築・運用を外部企業へ依存する傾向から組織内にノウハウやスキルが蓄積されないためと述べられている。以上のことから、組織内で情報システムを構築・運用することも重要であるといえる。その取り組みを通じ、組織は IT 人材を確保・育成でき、持続的な業務効率化や付加価値創出を実現する。

名古屋大学 情報基盤センターでは、スーパーコンピュータ「不老」の利用者より研究成果情報を集計し、管理・分析する業務がある^[2]。しかし、既存の業務プロセスを分析すると、利用者が回答する研究成果情報アンケートのツールを始めとし、各プロセスは質・量の観点から改善の余地が大きい。本稿では、研究成果情報を管理するためのツールとして Excel ファイルを用いた業務プロセスを述べる。具体的には、利用者は Excel 形式の研究成果情報アンケートに回答し、それを情報基盤センターへ提出する。そして、提出された Excel ファイルの中から必要セルを抽出し、管理用 Excel ファイルへ半自動的に転記するシステムを構築した。

本業務プロセスの改善によって、既存業務の効率化に加えて、付加価値創出に向けた研究成果情報分析における基盤づくりを目指す。

2章では研究成果情報管理業務を登場人物を含め体系的に説明すると同時に、既存業務プロセスの課題を整理する。また、他大学における同業務への取り組みを紹介する。3章ではシステムをリリースするにあたり検討した事項として要件定義や、設計について述べる。また、システムの仕様に加え、システムを業務で使用することを想定した業務プロセスも述べる。4章ではシステムの業務への適用事例を述べる。ここでは、特に既存業務プロセスと比較した場合における改善事項について述べる。最後に5章でまとめる。

以降では、スーパーコンピュータ「不老」を「不老」、名古屋大学 情報基盤センターを「センター」、「不老」利用者を「利用者」と表記する。

2 研究成果情報管理業務に関して

2.1 業務の概要

2020年7月1日より、名古屋大学 情報基盤センターでは「不老」を運用している。「不老」による計算需要支援による研究力強化や、人材育成、社会連携を実現するために、センターは利用者の研究成果情報を管理・分析する必要がある。

研究成果情報管理業務は、「不老」を通じて挙げた1年間の研究成果情報のアンケートを利用者へ実施することから始まる。ここでいう研究成果情報とは、表1に示す4種類：論文誌、国際会議、研究会報告、著書の実績としている。すなわち、利用者は年度内に投稿した論文誌や、参加した国際会議等の各種情報を研究成果情報としてセンターへ回答する。なお、実際には表1に定める回答項目に加え、利用者の傾向分析のために利用者の年齢層、性別、各種成果物の査読有無も回答項目として用意している。

表1. 研究成果情報の定義

No	種類	内容	回答項目
1	論文誌	和文雑誌、英文雑誌等	著者名、論文名、雑誌名、 巻数、号数、ページ番号、発表年
2	国際会議	国際会議プロシーディング等	著者名、論文名、プロシーディング名、 編者名、版元、ページ番号、発表年
3	研究会報告	研究会報告、ポスター発表等	著者名、論文名、研究会名、巻数、号数、 ページ番号、発表年
4	著書	単行本	著者名、タイトル、版元、発表年

センターは回答内容を集計し、各種情報を体系的に管理する。また、その内容を確認することで「どの研究分野の利用者が多いか?」、「昨年度と比較して利用者の傾向はどのように変化したか?」といった分析を行う。この分析を通じ、中長期的なハードウェア・ソフトウェアの整備、講習会開催提案等の計算需要支援を行うための判断材料とする。

2.2 既存業務プロセスとその課題

本稿で述べるシステムを開発する以前は、利用者が回答する研究成果情報アンケートのツールとして、Office365が提供する「Microsoft Forms」を用いていた。以降では、Microsoft Formsを「Forms」と表記する。既存の業務プロセスは以下のように整理できる。

- (1) 利用者は、期間内にFormsへアンケート項目を入力することで回答し、センターへ提出する。
- (2) センターは、Forms管理画面よりアンケート結果をExcel形式ファイル(.xlsx)でダウンロードする。

- (3) センターは、ファイルを Excel 形式から csv 形式(.csv)へ変換する。
- (4) センターは、スクリプトにより csv ファイルを整形する(不要行列の削除等)。
- (5) センターは、csv ファイルを Excel ファイルへ変換した後、手動調整して管理表を作成する。
- (6) センターは、必要に応じて管理表を確認する。

ここで、既存業務プロセスを観察すると、いくつか改善点が挙げられる。プロセス(1)に着目すると、研究成果情報を Forms へ登録することは利用者にとって負荷がかかることが分かった。例えば、利用者が研究成果情報の1つとして論文誌を入力することを考える。利用者は、表1に定める回答項目に従い、著者名、論文名、雑誌名、巻数、号数、ページ番号、発表年等をフォームへ入力する必要がある。Forms の仕様上、画面レイアウトは縦方向に長いために、項目を繰り返して入力することは負担がかかることに加え、登録内容を俯瞰しにくい。また、Forms は回答済みのアンケートを確認し、修正することができない。そのため、利用者は研究成果情報を追加登録したい場合は最初から登録し直す必要がある。プロセス(3)~(5)に着目すると、ファイル形式を変換する工程が複数回発生している。また、プロセス(5)に着目すると、人的作業工程があるために作業ミスが発生する可能性がある。

以上のことから、研究成果情報管理業務プロセスを改善するにあたり、利用者目線では「回答しやすさ」、「確認しやすさ」、「修正しやすさ」といった要素が求められる。その一方でセンター目線では、効率的(より簡単に、より早く、より確実に)な集計を可能とする仕組みが求められる。

2.3 他大学の取り組み

研究成果情報管理業務は他大学でも行われており、その実例を紹介する。

東京大学、九州大学ではスーパーコンピュータ利用者を対象とした研究成果情報登録用 Web ページを公開している^{[3][4]}。利用者はフォームに必要情報を入力することで、登録可能である。

Web ページは、フロントエンド側のプログラム言語の豊富さ(例えば、HTML、CSS、JavaScript 等)から拡張性が高いために、見やすさや入力しやすさといったユーザビリティを考慮した柔軟な構築が可能である。また、近年では WordPress 等のコンテンツ管理システムや Bootstrap 等のフロントエンド Web アプリケーションフレームワークが普及したために、Web ページを構築しやすくなったといえる。

その一方で Web ページを構築・運用するためには専門的なスキル、コストが要求される。具体例を挙げると、Web ページを公開するまでには、要件定義、設計、開発のスキル、人的コスト、サーバ等のリソースが必要である。コンテンツ管理システムによるテンプレート機能を用いて Web ページを簡易的に構築した場合においても、公開されているソースやプラグインに動作保証がない場合もあるためにセキュリティの観点から脆弱性を秘めているといえる。Web ページ公開後においても、利用者向けの使用マニュアルの作成と更新、使用方法に関する QA 対応、セキュリティ対策(SSL 証明書更新、不正アクセス対策)、運用サーバの管理(ソフトウェアの定期的なメジャー/マイナーバージョンアップ、LDAP、稼働状況確認)といった業務が別途発生する。

Web ページの構築・運用を外部ベンダへ委託した場合においても、金銭的成本が発生することに加え、外部企業にはノウハウが蓄積されるのに対して、組織内では IT 人材の確保や育成ができない。

3 システムリリース方針

3.1 要件定義

利用者の研究成果情報を集計し、管理するシステムを構築するにあたり、達成すべき項目を5つ挙げる。

また、5項目の中でも利用者目線の項目を特に優先するものとする。

- ・ 利用者は、アンケート回答時に負担を感じにくく、自身が入力した内容を俯瞰できる。
- ・ 利用者は、一度提出済みのアンケートに対して、内容を修正して再提出しやすい。
- ・ 利用者が提出したアンケートは、他者により盗聴・改竄されない。
- ・ センターは、提出されたアンケートを効率的に集計し、管理表を作成できる。
- ・ センターは、管理表を確認して分析しやすい。

研究成果情報アンケートツールの候補として、以下に示す大きく3つの方法がある：(1)Forms等のアンケートツール、(2)Web ページ、(3)Excel等のファイル形式。その中で、我々は(3)Excelによるファイル形式を採用することとした。

(1)のアンケートツールは、2.2節で述べた通り質問項目が多いと画面レイアウトが縦長になり、回答・俯瞰しにくい。また、再提出に手間がかかるため、候補より除外した。(2)の Web ページは、2.3節で述べた通り、利用者にとって少なからず操作方法を理解し、必要に応じて質問する手間が発生する。センター側にとっても恒久的な人的コストと専門的スキルが要求される。(3)の Excel は、アカデミック・民間問わず多くの利用者にとってセル入力操作に馴染みがあるといえる。Excel を始めとする Office はどの現場においても導入しやすい。また、利用者グループ内で Excel ファイルを共有することで、他のメンバーもアンケートに回答しやすい。

以上のことから、利用者が Excel ファイルをセンターへ提出し、センターはそれを集計・分析しやすいシステムを構築するものとする。

3.2 基本設計

本システムは、大きく分けると3つの要素から構成される。

利用者用入力 Excel ファイル：利用者用入力 Excel ファイルを図1に示す。本ファイルは研究成果情報アンケートの役割がある。利用者は本ファイルのセルを埋めることでアンケートに回答し、センターへ提出する。

② 国際会議 用 入 力 シ ー ト														
(注意) 1. Excelファイル(以下、ファイルと称する)の名称は、ご自身の登録番号に併せて変更してください。(例) 2022年度スーパーコンピュータ「不老」研究成果登録書_z49999a.xlsx 2. 内容入力後は、ファイルをNUSへアップロードしてください。 3. 受付期間内であれば、ファイルは何回でもアップロード可能です。受付期間内に複数回のアップロードが行われた場合、最後にアップロードされたファイルを受理します。 4. ファイルを正しく上書き管理するために、 同じファイル名称でアップロードして下さい。														
参考情報欄														
No	著者名	論文名	プロシーディング名	編者名	ページ番号	版元	発表年	査読	受賞	種別	謝辞	DOI	インパクトファクター	参考URL
例	Taro Meikai, Hanako Meikai	Performance evaluation of OO	Proc. OO	-	pp.xxxx-xxxx	ACM	2022	-	-	-	-	-	-	-
1														
2														
3														
4														
5														

図1. 利用者用入力 Excel ファイル

管理者用統計 Excel ファイル：管理者用統計 Excel ファイルを図2に示す。本ファイルは利用者の研究成果情報を管理し、分析する管理表の役割がある。本ファイルの各セルは、利用者用入力 Excel ファイルにおける入力セルを抽出したものである。

	A	B	C	D	E	F	G	H	I	J	K	L
1											分類『論文誌』場合のみ	分類『国際会議』場合のみ
2	分類	査読	受賞	種別	謝辞	インパ	DOI	著者名	タイトル(論文名または、著書)	雑誌名	プロシーディング名	研究会
3	論文誌	有	有	-	-	-	-	名大太郎, 名大花子	〇〇の性能評価	情報処理学会論文誌	-	-
4	国際会議	有	-	HPCI	-	-	-	Taro Meidai, Hanako Meidai	Performance evaluation of 〇〇	-	Proc. 〇〇	-
5	研究会報告	-	-	名大HPC	有	-	-	名大太郎, 名大花子	〇〇の性能評価	-	-	情報処
6	著書	-	-	-	-	-	-	名大太郎	〇〇	-	-	-
7												
8												
9												
10												
11												

図2. 管理者用統計 Excel ファイル

統計自動化スクリプト：本スクリプトは、利用者用入力 Excel ファイルから管理者用統計 Excel ファイルへの転記を自動化する役割がある。提出された利用者用入力 Excel ファイル群に対して、各ファイルのセルをコピーし、管理者用統計 Excel ファイルの指定セルへペーストする。

なお、Excel ファイルを主軸とするために、利用者、センターが用いる端末 OS は Windows であることを前提としている。

本システムをベースとした業務プロセスを下記に示す。

- (1) 利用者は、利用者用入力 Excel ファイルを回答し、センターへ提出する。
- (2) センターは、提出されたファイル群をダウンロードする。
- (3) センターは、統計自動化スクリプトを実行することで、管理者用統計 Excel ファイルを更新する。
- (4) センターは、必要に応じて管理者用統計 Excel ファイルを確認する。

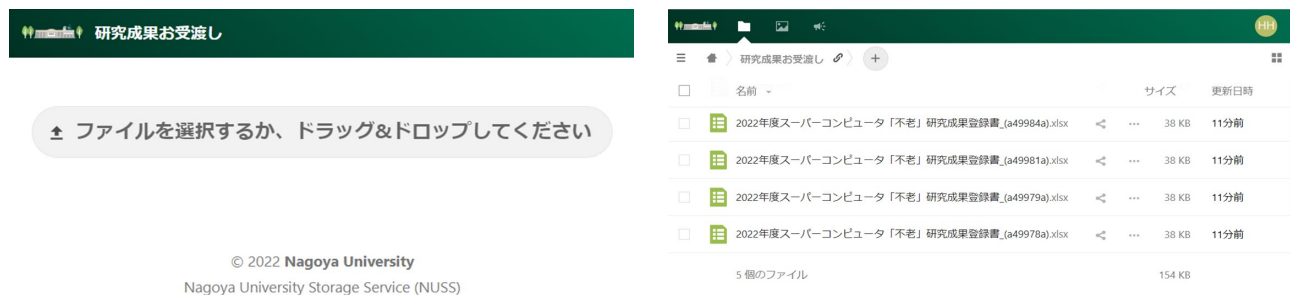
3.3 詳細設計

利用者用入力 Excel ファイル

研究成果情報アンケートに関する利用者への告知は名古屋大学 情報連携推進本部のホームページにて行っている^[5]。利用者用入力 Excel ファイルは本 Web ページよりダウンロード可能であり、利用者は各種回答する。本 Excel ファイルは 5 つのシートから構成されており、表 1 に定める「論文誌」、「国際会議」、「研究会報告」、「著書」の 4 つのシートに加え、利用者アンケート用のシートから成る。前述の 4 つのシートには、表 1 に定める回答項目に加えて、参考情報として査読有無、受賞有無、謝辞、DOI 等を回答頂く。

利用者は回答が完了した後、NUSS(Nagoya University Storage Service)教育研究ファイルサービス(以下、NUSS)へ Excel ファイルをアップロードする。NUSS とは名古屋大学が運用している、WebDAV をベースとしたファイル共有サービスである。図 3(a)に示すように、利用者は Web ブラウザを用いて、アップロード専用 URL より Excel ファイルをアップロードすることを想定する。

なお、アップロード専用 URL は「不老」アカウントでログイン可能な Web ページ「HPC ポータル」に記載されている。これにより、提出ファイルに対する盗聴・改竄等の攻撃は行われにくいと想定している。



(a) 利用者用画面(アップロードのみ可能)

(b) センター用画面(ダウンロードが可能)

図 3. NUSS(Nagoya University Storage Service)教育研究ファイルサービス

統計自動化スクリプト

アンケート回答締め切り後、図 3(b)に示すように、センターは NUSS へアップロードされたファイル群を Zip ファイル(.zip)としてダウンロードし、統計自動化スクリプトの作業対象とする。

ここで、統計自動化スクリプトは Python によって構築した。Python は他のプログラム言語と比較して、コード記法が容易であり、ドキュメントが豊富であることから初学者でも学習しやすいといえる。また、Python はライブラリが豊富であるために、システムの拡張性が高い。例えば、本スクリプトでは、Excel ファイルを扱うライブラリ「OpenPyXL」をベースに実装している^[5]。本ライブラリは Excel ファイルを開くことや、シートの各セルに対する Read/Write を可能とする。また、システムを中長期的に拡張することを検討すると、スクリプトベースのグラフ描画を可能とするライブラリ「matplotlib」や、データ分析・傾向分析を可能とするライブラリ「pandas」等も候補として考えられる。

本スクリプトの実行フローチャートを図 4 に示す。

図 4 の Step2 に着目する。NUSS の仕様では同名のファイルが複数回アップロードされた場合、ファイルは上書き更新されず、改版数が付与された別ファイルとして保存される。例えば、1 回目のアップロードにてファイル名称を「(ファイル名).xlsx」とした場合、2 回目も同名ファイルをアップロードしても「(ファイル名)(2).xlsx」として保存される。Step2 では、改版数が最新のファイルを処理対象として識別する。

図 4 の Step3 に着目する。利用者が複数回アップロードし、その都度ファイル名称を変更するケースも考えられる。例えば、1 回目のアップロードにてファイル名称を「(ファイル名).xlsx」とし、2 回目は異なるファイル名称「【修正版】(ファイル名).xlsx」とした場合には、改版数は更新されない。また、利用者は誤操作によって利用者用入力 Excel ファイル以外のものを NUSS へアップロードする可能性もある。このことから、下記の運用ルールを定義し、利用者へ提示した。

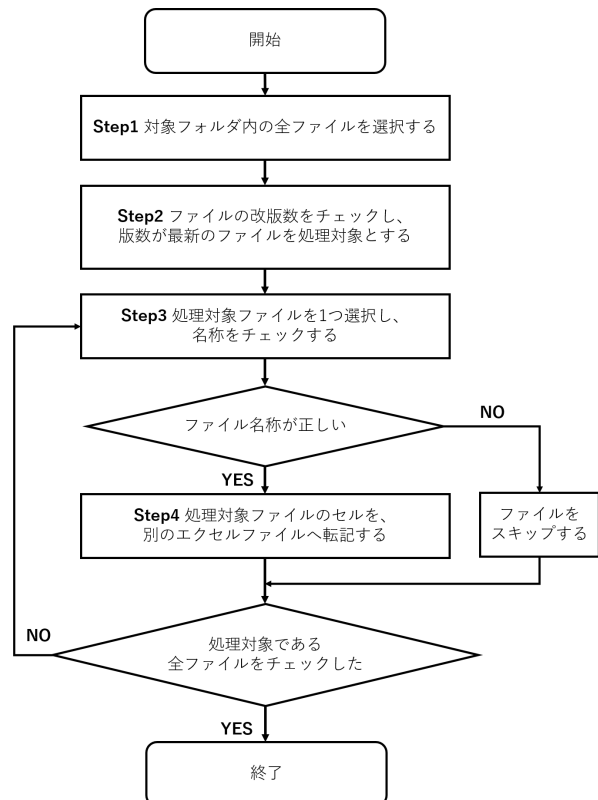


図 4. スクリプトの実行フローチャート

- アップロードする Excel ファイル名称は、”flow_seika2022_USERID.xlsx”として下さい。
(例) flow_seika2022_z49999a.xlsx
- 受付期間内であれば、ファイルは何回でもアップロード可能です。複数回のアップロードが行われた場合、最後にアップロードされたファイルを受理します。
- ファイルを正しく管理するために、同じファイル名称でアップロードして下さい。
(正しい例) flow_seika2022_z49999a.xlsx
(誤った例) 修正版 flow_seika2022_z49999a.xlsx

管理者用統計 Excel ファイル

本 Excel ファイルは、センターが検索操作をしやすい設計を目指した。具体的には、査読有無、受賞有無、プロジェクト種別といった項目列を表中の左側へ配置することで、目的に合わせてフィルター検索しやすい構成とした。一方で比較的検索頻度が少ないと思われる項目列(巻数、号数、ページ番号、版元等)は表中の右側へ配置しつつ、グループ化することで折りたたむ構成とした。センターは必要に応じて表示と非表示を切り替え可能とすることで、可読性の向上を目指した。

4 システムの業務への適用事例

本プロセスを業務へ適用した場合における、既存プロセス(2.2 節参照)からの改善事項を述べる。

利用者目線で着目する。研究成果情報アンケートの既存ツールであった Forms と比較して、Excel は表形式でかつ、画面縮尺変更を可能とするために、入力しやすく俯瞰しやすくなったといえる。また、Excel は内容を修正しやすい点、Excel をグループ内のメンバーに共有することで回答可能な点からアンケートにより回答しやすくなったといえる。

センター目線で着目する。提出後のアンケートを集計し、管理表を作成するまでの作業工程は1回あたり1時間程度減ることが見込まれる。これは既存プロセス(5)で述べた管理表レイアウトに対する手動調整工程が減ったためである。既存プロセスにおいて利用者が Forms で複数回登録した際、登録内容が重複するケース、重複しないケースの両者が発生する。前者の例は、利用者が Forms で全項目を最初から登録し直したケースである。後者の例は、利用者が Forms で変更分のみを登録するケースである。このことから既存プロセスにおいて、センターは目視で登録内容が重複しているかを判断し、手作業で削除する必要があった。本プロセスではスクリプト、及び運用ルールによって、その手作業工程を減らすことができた。それに付随して、作業ミスを減らすことができた。

5 まとめ

本稿では、「不老」の利用者における研究成果情報管理業務プロセスと、そのツールを見直した。利用者にとって馴染み深い Excel を研究成果情報アンケートとして用いることで、ユーザビリティを意識した設計を目指した。また、その集計をスクリプト化することで業務の効率化に加え、分析のための基盤づくりを目指した。

今後の課題として、2022年9月より公開している利用者用入力 Excel ファイルに対して、利用者視点のフィードバックを受けることが挙げられる。また、スクリプトに対して Python ライブラリを導入することで、分析・可視化の自動化が挙げられる。

参考文献

- [1] 総務省 - 令和3年版情報通信白書“第1部序章我が国におけるデジタル化の歩み”
<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r03/pdf/01honpen.pdf>(accessed 2023-03-23).
- [2] 名古屋大学情報連携推進本部“スーパーコンピュータシステム” <https://icts.nagoya-u.ac.jp/ja/sc/> (accessed 2023-03-23).
- [3] 東京大学情報基盤センタースーパーコンピューティング部門“研究成果登録”
<https://regist.cc.u-tokyo.ac.jp/ar/> (accessed 2023-03-23).
- [4] 九州大学情報基盤研究開発センター“成果報告システム” <https://hpc-report.cc.kyushu-u.ac.jp/> (accessed 2023-03-23).
- [5] PyPI“openpyxl” <https://pypi.org/project/openpyxl/> (accessed 2023-03-23).