令和4年度経済産業省デジタルプラットフォーム構築事業 (Gビズインフォの保守・運用) 最終報告書

2023年2月28日株式会社日立社会情報サービス

目次

本書サマリ	3
1. 事業の概要	4
1.1. 本事業の概要	4
1.2. 事業計画	6
2. Gビズインフォにおけるデータ収集・追加及びユーザ支援に関する運用・調査	7
2.1. Gビズインフォへのデータ収集・追加に関する運用・調査	7
2.2. ユーザへの支援等	10
2.3. 「営業品目」の改修について	11
2.4. 補助金情報における出典元情報の明確化について	16
2.5. その他サービス改善のための機能追加等	17
3. 稼働環境構築に関する整備	19
4. Gビズインフォの保守・運用	21
4.1.1. 保守・運用の実施	21
4.2. 保守・運用作業準備	25
4.3. 掲載情報の管理	26
5. 事業実施上の知見	27

本書サマリ

本書には各章各項において以下にあげる事項を記載している。

- (1) Gビズインフォにおけるデータの追加・ユーザ支援に関する運用・調査
 - ・ データ追加について、Gビズインフォヘデータを登録し運用・調査を実施しデータ拡充を行った。本事業で実施した内容を本書 2.1 章にまとめた。
 - ・ ユーザ支援について、システム管理者、データ提供者、一般利用者に関して支援を行った。本事業で実施した内容を本書 2.2 章にまとめた。
 - ・ 「営業品目」にかかる改修を実施した。これらを実施するにあたり課題の背景と実施した概要を本書 2.3 章にまとめた。
 - ・ 「補助金情報における出典元情報の明確化」にかかる改修を実施した。これらを実施するにあたり課題の背景と実施した概要を本書 2.4 章にまとめた。
 - ・ 「その他サービス改善のための機能追加等」として当事業で可能な範囲でいくつかの施策を実施した。これらの実施概要を本書 2.5 章にまとめた。

(2) 稼働環境構築に関する整備

・ Gビズインフォの稼働環境構築に関する整備として、稼働環境ならびに利用状況に応じた最適な環境を検討した結果と実施結果を本書3章にまとめた。

(3) Gビズインフォの保守・運用

- ・ 保守・運用の実施内容として、稼働監視状況と障害対応、セキュリティ監視に関する事項を実施 し、実施内容について本書 4.1 章にまとめた。
- ・ 保守・運用作業準備について、滞りなく実施して本事業を開始しており実施内容について本書 4.2 章にまとめた。
- ・ 掲載情報の管理として、当該サイトに掲載する情報について管理を行った。本事業で実施した内容を本書 4.3 章にまとめた。

(4) 事業実施上の知見

- ・ 当年度に「営業品目」(業種コード)の取り扱いを変更したため、これに伴うGビズインフォに おける留意点を記載した。
- ・ 当年度に「BI ツールの活用」の対応を行い、利用環境の構築を実施した。これに伴うGビズインフォにおける留意点を記載した。
- ・ 当年度においてgBizINFOの将来構想を討議した。これらの結果を知見として記載した。
- ・ 運用における成果指標に関する考え方を記載した。本指標に関しては次年度以降も実績を元に 目標を設定し計測することが有効であると考える。

1. 事業の概要

1.1. 本事業の概要

(1) 事業目的

我が国は、「行政手続における特定の個人を識別するための番号の利用等に関する法律」(平成二十五年五月三十一日法律第二十七号)に基づき、平成27年10月、全ての国内の法人に対して法人番号を付番した。

政府は、予算執行、届出、表彰等で多数の企業に接触・関与し、様々な法人活動情報(法人の名称等に紐付いて管理されている情報)を取得しているが、これらを法人番号と紐付けて機械判読可能な形式に整理し、各情報を多面的に活用・分析することが可能となれば、効率的・効果的な政策立案・執行が可能である。また、民間事業者等においても、組織内で保有する法人活動情報とともにこれら政府の公開情報を活用することにより、新たなサービスの創出やサービスの高度化に資すると考えられる。

こうした考えの下、経済産業省では、平成28年度から、経済産業省版法人ポータルサイトベータ版の試行運用の後、平成29年1月に全府省庁が保有する法人情報を一括検索、閲覧、取得が可能なサイトである「Gビズインフォ」(https://info.gbiz.go.jp)(令和2年3月に法人インフォメーションを改称)を公開し、法人情報の拡充や追加機能の検証を図りながら、法人情報の利活用に対する要望の収集を行ってきた。

本事業では、Gビズインフォの保守運用を行うほか、令和3年度経済産業省デジタルプラットフォーム構築事業 (Gビズインフォの保守運用・機能追加)(以下、「前年度事業」という)の調査報告書等も踏まえ、より迅速・正確かつ効率的なデータ収集・追加方法に関する調査研究や運用評価を実施する。

(2) Gビズインフォの概要

Gビズインフォでは、政府機関が保有する法人活動情報を法人番号で紐付けし、オープンデータとして公開している。利用者は、法人番号、法人名その他詳細情報をもって該当法人を検索し、各府省庁が保有する補助金、表彰、届出・認定、調達、特許、財務情報といった法人活動情報の検索・閲覧に加え、ファイル(PDF、CSV 及び XML)ダウンロード及び API(SPARQL API 及び情報提供 REST API)を利用した機械可読な形式での構造化されたデータ取得が可能となっている。また、これらのデータをGビズインフォでは、独立行政法人情報処理推進機構が作成した「共通語彙基盤コア語彙 V 2(V e r 2.4)」を参考に、法人情報を表現するための「法人活動情報語彙 V 2」を定義の上、法人情報についてオープンデータとして利活用しやすいように構造化されたデータ整備を行っている。

(3) 事業概要

Gビズインフォの保守運用を通じて、より迅速・正確かつ効率的なデータ収集・追加等を行うとともに、データ量増加や政府調達制度に応じたシームレスなシステム開発・運用に対応する。事業概要は次のとおりである。これらを実現するため本書2章~4章の取組を実施している。

- (ア)前年度事業で構築・運用開始したシステムの引継ぎと保守運用ならびに新たな法人活動情報の データ収集・追加を行う。
- (イ) G ビズインフォの稼働環境を提供するとともに、利用状況に応じた最適な環境を構築する。
- (ウ)法人活動情報のGビズインフォのデータ収集・追加作業や法人番号付与(提供されるデータに法 人番号を付与する)作業を通じて、既存環境やツール等の検証、評価を行い、経済産業省と協議 の上、必要に応じて機能追加や改修等を検討・実施する。
- (エ)事業期間における本システムの適切な運用を行うとともに、アクセス状況分析からのユーザ層特定や問合せ項目等からのニーズ把握など、運用を通じて得られる知見について調査を行い、本システムについての評価・改善事項の検討を行う。

1.2. 事業計画

本事業は2022年4月(令和4年4月)より開始され、事業は下記に従い実施した。

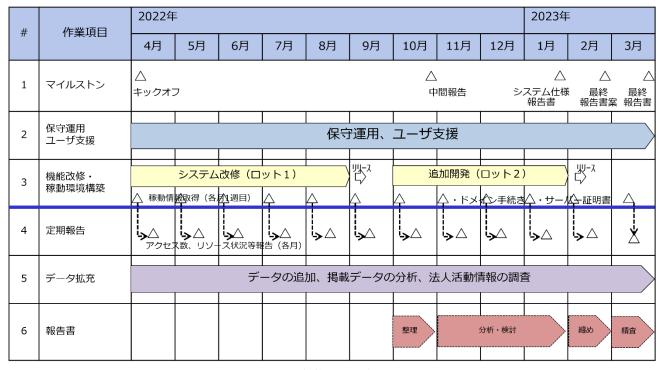


図 1.2-1 全体スケジュール

(1) 活動状況

個々の活動状況は次のとおり

- (ア)保守・運用、ユーザ支援 計画どおり実施。
- (イ)機能改修・稼働環境構築 計画どおり実施。
- (ウ)定期報告 計画どおり実施。
- (エ)データ拡充計画どおり実施。
- (オ)報告書 計画どおり実施。

- 2. Gビズインフォにおけるデータ収集・追加及びユーザ支援に関する運用・調査
- 2.1. Gビズインフォへのデータ収集・追加に関する運用・調査

以下(ア)及び(イ)のデータについて、本システムに追加すること。その際、法人番号が付与されていないデータについては、国税庁法人番号公表サイトの法人情報と、(ア)及び(イ)のデータの「法人名」、「所在地」をキーとして突合するなど機械的な名寄せ、本システムへのデータの追加に必要なフォーマットの変換を行い、データ収集・追加に関してより迅速・正確かつ効果的・効率的な方法等について調査する。

またこれら作業を通じて、既存環境やツール等の検証、評価を行い、経済産業省と協議の上、必要に応じて機能追加や改修等を検討・実施する。

(ア)経済産業省より提供するデータの追加登録

経済産業省が他府省庁等へ提供を依頼し収集する以下の法人活動情報データ区分について本システムに追加すること。データ提供は、区分1については、6月を想定、区分2については、上半期と下半期に1度ずつ、各府省庁保有分の情報提供に加え、毎月100件程度の独立行政法人保有分及び国立印刷局保有の官報決算公告情報の提供を想定している。

<データ区分>

- ①法人活動情報(特許情報) 過去10年分(約400万件想定)
- ②法人活動情報(特許以外) 10万件以上(上半期、下半期合計)

<作業実績>

データ区分①法人活動情報(特許情報)について登録を実施。約430万件を登録した。

データ区分②法人活動情報(特許以外)については2月までに10万件(うち8万件は前年度データ)の登録を行っているが、さらに当年3月に約5万件を追加登録する予定。

※表中の数字は gBizINFO に収録されている全数。カッコ内の数字は当年度の作業実績を示す。

財務情報 特許情報 届出・認定 補助金情報 調達情報 表彰情報 職場情報 行政処分情報 中計 4,553 4,370,781 204,871 576,958 338,750 63,720 20,995 (+4,370,781)(+26,338)(+1,824)(+16,229)(+62,372)5,580,628(+4,459,713) 合計 (特許情報: +4,370,781)

(特許以外: +88,932)

表 2.1-1 提供データの登録実績(令和4年2月8日時点)

▶ 特許情報:法人番号が特定できたデータのうち、過去10ヵ年分(特許出願日が2012年~2021年)のものを新たに再収録。約430万件を登録した。特許は長期スパンで申請・審査中データが失効や取消等になるため、Gビズインフォでは年1回の全件洗い替えを行なうこととしている。

▶ 特許以外:国立印刷局保有の官報決算公告情報は(イ)に記載。

(イ) 政府機関のウェブサイトから取得したデータの追加

次のデータを各ウェブサイトから API 機能等を通じて取得し、日次で本システムにデータを追加した。

- ◆ 法人基本3情報(法人番号、法人名称、所在地):法人番号公表サイト(国税庁)
- ◆ 法人基本情報 (代表者名、資本金、従業員数、設立年月日、営業品目など): 統一資格審査申 請・調達情報検索サイト (デジタル庁)、しょくばらぼ (厚生労働省)、EDINET (金融庁)
- ◆ 勤務実態情報:しょくばらぼ(厚生労働省)
- ◆ 財務情報:EDINET (金融庁)
- ◆ 法人活動情報 (認定情報): スマートSMEサポーター (経済産業省) また、次のデータについてはGビズコネクトを経由して取得し、適切なタイミングで本システム にデータを追加すること。
- ◆ 補助金情報(デジタル庁)
- ◆ 官報決算公告情報(国立印刷局)

なお、事業途中において、上記について各ウェブサイトでの仕様変更等相手サイトの要因により データの取得、追加ができなくなった場合は対象から除外する。

<作業実績>

表 2.1-2 の網掛け箇所が当年度において更新・追加が生じた情報を表す。

表 2.1-2 ウェブサイトからの登録実績(令和4年2月8日時点)

登録法人件数: 5,272,099件 職場情報 財務情報 特許情報 届出・認定 補助金情報 調達情報 表彰情報 行政処分情報 中計 4,553 4,370,781 204,871 576,958 338,750 63,720 20,995 合計 5,580,628 決算情報:5,874

- ▶ 法人基本3情報:法人番号公表サイト(国税庁)より登録法人件数として約520万法人を収録。当年度において約14万法人が増加。
- > 法人基本情報:統一資格審査申請・調達情報検索サイト(デジタル庁)より約6万法人分を 日次で洗い替え収録。その他サイトの各情報より基本情報の補完している。
- ▶ 勤務実態情報:しょくばらぼ(厚生労働省)より、職場情報を約2万レコード収録。
- ▶ 財務情報: EDINET (金融庁)より、大企業約4500法人の決算報告を収録。当年度において約4000法人の決算情報を更新。
- ➤ 法人活動情報(認定情報):統一資格審査申請・調達情報検索サイト(デジタル庁)より約6万法人分を収録。当年度に新たに収録された法人は約13500法人。スマートSMEサポーター(経済産業省)より、届出認定情報を約1300法人分収録。当年度に新たに認定された法人は約300法人。しょくばらぼ(厚生労働省)より、ユースエール認定を登録。

- ▶ 補助金情報 (デジタル庁): 3月に jGrants システムとの連携によりデータ収録予定。
- ➤ 官報決算公告情報 (国立印刷局):中小企業の決算情報として約6000件 (1法人につき 最大5件)を収録。

なお、各サイトとの情報連携においては対象となる各ウェブサイトでは不定期にシステムメンテナンスが発生する。このため日々ウェブサイトの告知を監視しシステムメンテナンスが発生する都度、システム影響がないか(正常にデータが取得できているか)検証を行うなどの対応をとっている。当年度の発生状況は表 2.1-3 のとおり。(計 2 5 回)

表 2.1-3 サイトメンテナンス発生一覧

#	サイト	回次	メンテナンス発生日	影響確認日	備考
1	法人番号公表サイト(国税庁)	1	22/09/23~22/09/24	22/09/26	
2		2	22/10/28~22/10/30	22/10/31	
3		3	22/11/11	22/11/14	
4	統一資格審査申請・	1	22/04/08~22/04/10	22/04/11	
5	調達情報検索サイト(総務省)	2	22/04/28~22/05/02	22/05/03	
6		3	22/06/15	22/06/16	
7		4	22/07/22~22/07/25	22/07/26	
8		5	22/08/05~22/08/08	22/08/09	
9		6	22/09/09~22/09/12	22/09/13	
10		7	22/10/27	22/10/28	
11		8	22/11/04~22/11/07	22/11/08	
12		9	22/12/02~22/12/05	22/12/06	
13		10	22/12/09~22/12/12	22/12/13	
14		11	22/12/28~23/01/04	23/01/05	
15	しょくばらぼ (厚生労働省)	1	22/06/30		インターフェースに影響する修正が
				22/07/01	あり、システム連携を一時的に中断
					(22/07/01~22/07/13)
16		2	23/01/23	23/01/24	
17	EDINET (金融庁)	1	22/04/16~22/04/18	22/04/19	
18		2	22/04/29~22/04/30	22/05/01	
19		3	22/05/14~22/05/16	22/05/17	
20		4	22/09/17~22/09/20	22/09/21	
21		5	22/11/26~22/11/28	22/11/29	
22		6	22/12/31~23/01/04		インターフェースに影響する修正が
				23/01/05	あり、システム連携を一時的に中断
					(23/01/01~23/01/25)
23		7	23/01/07~23/01/10	22/01/11	

24	SMART SME(経済産業省)	1	22/07/20	22/07/21	
25		2	23/01/04	23/01/05	

2.2. ユーザへの支援等

本システムのユーザ (システム管理者、データ提供者、一般利用者) を支援するため、技術面の問い合わせや利用者の利便性向上のためのシステム変更対応等を実施した。当年度は2月8日までに60件程度(技術支援:約40件、システム変更:約20件)の対応を行っている。

2.3. 「営業品目」の改修について

法人プロフィール画面での「営業品目」表記について、資格項目のコード表を適切なものに置換するとともに、情報提供元との整合がとれているかプログラム仕様の確認を行い、不具合があった場合には 改善のための提案を行い実施すること。

(1) 背景

Gビズインフォはオープンデータサイトとして、データを標準化して提供することに取り組んでおり様々なサイトから収集したデータを当プロジェクトで定めたデータ形式やコード体系に集約して公開している。ただし一部の例外があり、そのうちの1つに業種コードの取り扱いがある。

これまでGビズインフォで取り扱う業種コードは日本標準産業分類をマスターコードとして採用することとしてきた。一方で業種コードを収集しているのはデジタル庁の運営する全省庁統一資格記録の「営業品目コード」に限られており、この「営業品目コード」分類や体系はマスターコードのそれとは異なる考えであるためマスターコードのものとは1対1に対応付かず、マスターコードに対応するものだけを扱うとすると採用できる情報量として著しい低下を招くものとなっていた。

よって、業種コードの収録がデジタル庁の運営する全省庁統一資格記録に限られているうちは、業種コードに営業品目コードをそのまま収録し、日本標準産業分類への対応付けを公開することで日本標準産業分類に置き換えて表示する対応をおこなってきた。

このため「営業品目」から無理に置き換えられた日本標準産業分類は実態に合わない表示となるケースがあり、利用者の利便性低下や誤解を招くことになっている。

これまで長らくデジタル庁・全省庁統一資格記録からのみの収集となっていることから、利用者の利便性を鑑みて、業種コードの収集を同サイトの「営業品目」に限定すると共に、Gビズインフォにおける公開方法を見直すこととした。

(2) 実施概要

下記 (r) ~ (r) をロット1において実施した。(9月リリース)

(ア)プロフィール画面、プロフィール PDF

日本標準産業分類で表示を行っていたが、全省庁統一資格記録の営業品目コードに対応する名称を表示するように変更。

営業品目コードに対する資格等級も付記するように修正。資格等級はこれまでは届出認定情報の 資格レコードに独立していたが、法人基本情報(その他)として営業品目と並んで表記するように 改めた。



図 2.3-1 プロフィール画面



図 2.3-2 プロフィール PDF

(イ) REST-API

検索 API (/v1/hojin) を除く情報取得 API で取得可能な営業品目「business_items」の取得内容を日本標準産業分類の名称表示から営業品目のコード値に変更。コード値に対応する名称は利用者側がデジタル庁側を参照するものとする。

検索 API (/v1/hojin) については営業品目・資格等級をセットで指定して法人を絞込み検索可能となるサービスを追加した。

(ウ)詳細検索機能

プルダウンから日本標準産業分類の名称を選択し検索していたものを、営業品目の名称を選択して資格等級とも組み合わせてながら検索できるように変更した。

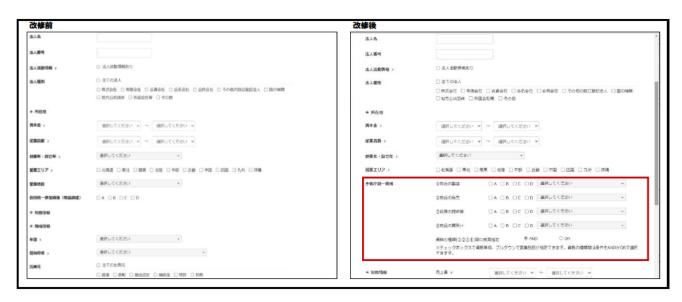


図 2.3-3 詳細検索機能

(3) 開発手法

本事業の改修全般において開発手法としてアジャイル開発を用いる。各スプリントを短く設定し、機能 実装に対して柔軟な対応ができるようにすることを目的として(ア)~(ウ)のように進めた。

(ア)進め方

本事業におけるアジャイル開発は、手法としてスクラムを利用する。開発の進め方を図 2.3-4 に示す。

- ✓ 1スプリントの期間を最大4週間をとし、数スプリントで開発(スプリント回数は開発ロット毎に調整して定める)
- ✓ スプリント毎にスプリントレビューを実施し、動作可能なプロダクトのデモを実施
- ✓ デモで得た意見を踏まえて次回スプリントの開発内容、要件を再度見直し
- ✓ スプリント毎に振り返りを実施、開発の問題点を洗い出し、次スプリントへインプットする



図 2.3-4 開発の進め方

(イ)管理方法

スプリント実施における管理方法を図 2.3-5 に示す。

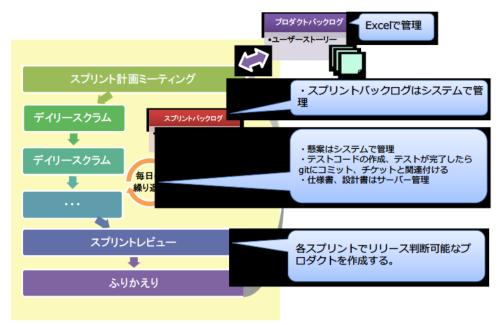


図 2.3-5 スプリント管理方法

(ウ)ロールと会議

スプリント実施におけるロール (役割) を表 2.3-1 とし、会議は表 2.3-2 にて実施した。

表 2.3-1:ロール(役割)

#	ロール	役割	参加会議と実施内容
1	プロダクト	・プロダクトの結果責任者	■スプリント計画ミーティング
	オーナー	・プロダクトバックログの管理者。優先順位の最終	・プロダクトバックログの優先順位の決定
		決定権限。	■スプリントレビュー
		・リリース計画を定める	・プロダクトバックログに追加すべき項目の検討
		・プロダクトが要求にあっているか検査する	・プロジェクトを進めるうえの懸案についての議論
		・開発チームに相談できるが、干渉はできない。	・現在の進捗を踏まえて、リリース日や完了日の決定
2	スクラムマスター	・スクラムがうまくいくよう外部からチームを守る	■スプリント計画ミーティング
		・スプリント内での計画変更、仕様変更等の防止	・プロダクトバックログ作成
		・教育、ファシリテート、コーチ、推進役	・スプリントバックログの作成
		・プロダクトバックログの良い管理方法を検討する	・ベロシティ(見積)の作成
		・プロダクトオーナーや開発チームにアジャイル開	■デイリースクラム
		発や Scrum について説明して理解してもらう。	・日々の作業確認
		・スプリント計画ミーティングやふりかえりなどの	■スプリントレビュー
		会議の進行	・デモの実施
		・生産性向上のための変化を促す	・完了できなかったプロダクトバックログについて説明
			・プロジェクトを進めるうえの懸案についての議論
			・リリース日や完了日の検討
			■ふりかえり
			・うまくいったこと、改善点の整理
			・ベロシティの確認(見積)
3	開発チーム	プロダクトの開発	■デイリースクラム
1			・日々の作業内容の更新(タスクボード・バーンダウン
			チャートの更新)
			■ふりかえり
			・うまくいったこと、改善点の列挙

表 2.3-2:会議体

#	会議	タイミングと時間	参加者	実施内容	資料
1	スプリント	スプリントの	・プロダクトオーナー	・プロダクトバックログを再度分析・評価し、そ	[インプット]
	計画ミーテ	開始前	・スクラムマスター	のスプリントで開発するプロダクトバックログ	・プロダクトバックログ
	ィング		・開発チーム	の項目を選択する。	[アウトプット]
				・選択した項目を個々の作業(タスク)に分割す	・スプリントバックログ

2	デイリー スクラム	毎日決まった時間、決まった場所で 30 分程度	・スクラムマスター ・開発チーム	る。 ・スプリントレビューのゴールイメージを明確にする。 朝会 ・毎日同じ時間・同じ場所で 1 人,数分で実施・前日やったこと、今日やること、懸案等の確認(対策検討は必要なメンバーだけで別途実施)・残作業の追跡・懸案の詳細は別途実施	[インプット] ・スプリントバックログ ・バーンダウンチャート
3	スプリントレビュー	スプリントの 終了時	・プロダクトオーナー ・スクラムマスター ・開発チーム (各チームリーダ)	・プロダクトオーナーへのデモを実施 ・完了できなかったプロダクトバックログについて説明 ・プロダクトバックログに追加すべき項目検討・プロジェクトを進めるうえで問題となる事項について関係者で議論・現在の進捗を踏まえて、リリース日や完了日を検討する。	「インプット」 ・動作するプロダクト 「アウトプット] ・プロダクトバックログ の追加項目
4	ふりかえり	スプリントの終了 後、1 時間程度。	・スクラムマスター ・開発チーム	・プロセスやツールなどの観点で実施したスプリントの検査。・うまくいったこと、今後改善すべき点を整理する。・今後のアクションプランをつくる・ベロシティの確認(見積)	КРТА

2.4. 補助金情報における出典元情報の明確化について

法人プロフィール画面で表示される補助金情報について、データソースを明確にする方法を検討する とともに、当該検討を踏まえ、実施すること。

(1) 背景

Gビズインフォで取り扱う補助金情報については、各府省から情報収集し人手を介して登録している補助金情報のほか、システム連携(デジタル庁・jGrants)によって収録している情報がある。後者は連携サイクルも異なり、また前者に比べて情報精度やデータ品質の面でも優れることから、将来的には同システムを介した情報収録としていくことが望ましい。jGrantsが出典元であるデータについてはそれであると判別できる方が利便性が高いため、これを明確にするためその方法を検討し実装する。

(2) 実施概要

(ア)プロフィール画面、PDF への対応

ユーザーがプロフィール画面上で補助金情報を閲覧する際、jGrants より収録したデータだと判別できるよう「jGrants」アイコンを付加する対応を行った。これにより jGrants 出典のデータであることを視覚的に把握できるようにする。(図 2.4-1)

画面表示の変更に伴い、プロフィール画面から PDF をダウンロードした際には当該データに対して「出典: jGrants」を印字する。(図 2.4-2)



図 2.4-1 プロフィール画面

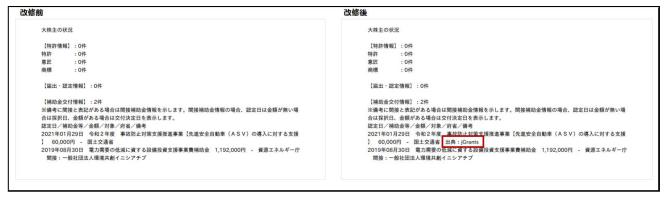


図 2.4-2 プロフィール PDF

2.5. その他サービス改善のための機能追加等

Gビズインフォのサービス改善に資する機能追加等について、当年度の事業範囲内でいくつかの対応を実施した。その内容は以下のとおり

(1) プロフィール画面表示(出典元リンク対応)

プロフィール画面の法人基本情報(その他)には複数の出典元サイトからの情報が入り乱れるため、情報の末尾に出典元を付記して表示している。これらデータの可読性を向上するため、出典元にアンカーを付して出典元サイトへ飛べる(必ず外部ウィンドウで開く)ように変更した。これにより次のような効果が期待される。

- ・ データ末尾が、データそのものでなく出典元を表していることを直感的に認識できる。
- 出典元からより詳細なデータを得たい利用者の操作を補助する。

なお、出典元の飛び先の変動を考慮しておく必要があるため、外部パラメータによって出典元サイト の追加・変更に対して柔軟に対応できるように設計している。

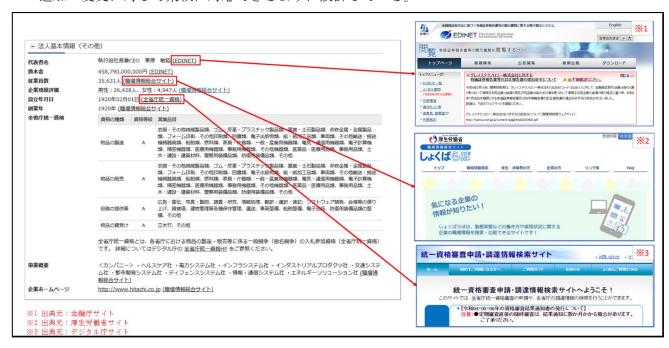


図 2.5-1 出典元リンク対応

(2) BI ツールの活用

運営管理者が API の利用状況や、データ登録状況の分析しやすくするため BI ツールを利用できる環境構築を図った。具体的には、必要な情報を AWS の S3 上に配置し、解析用のデータセットとして用意する。解析データセットについては、クエリサービスである Athena に対応した形式となるよう新規バッチにてデータ作成を行なう改修をした。今回構築した環境下では次の情報が解析用データセットして配置される。

- ・REST-API のログ
- ・gBizINFO 公開データの状況 (法人活動情報)

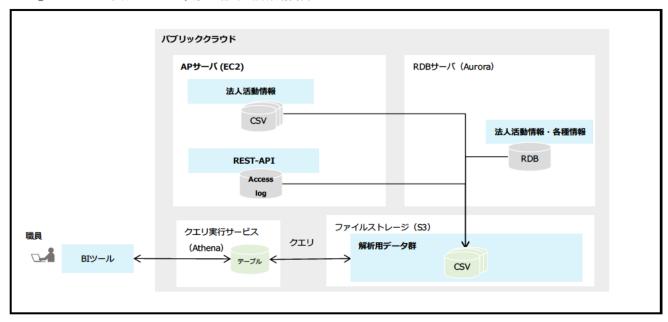


図 2.5-2 BI ツール活用

(3) gBizINFO の将来構想

gBizINFO は法人情報サイトとして法人活動データのオープンデータサイトとしてシステムを運営してきた。利用者からのアンケートには好意的な反応がある一方で、このようなサイトが存在することの認知度が低いといった話や、サイト運営側にとってはオープンデータサイトにおける特有の課題となるデータ利活用の状況が把握しづらいといった背景から、今後のシステムの在り方を討論し、利用者へ積極的にアプローチするサービスへの転換を図る必要があるのではといった議論を推し進めた。これらに関する詳細は5章の知見に記載する。

3. 稼働環境構築に関する整備

稼働環境については、原則として現行システムを引き継ぎ、22年4月時点で現在提供しているサービスを途切れなく提供した。提供している環境ならびに最適な稼働環境に関する見解を以下に示す。

(1) 稼動環境

当年度稼働環境のシステム構成は図 3-1 の通り。表 3-1 に各サーバの稼働環境のハードウェア仕様詳細を示す。

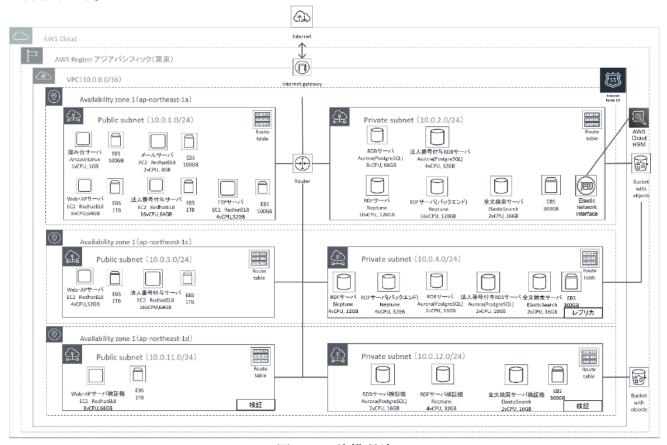


図 3-1 稼働環境

#	サーバ名	サービス名	CPU	メモリ	ディスク
			コア数	(GB)	(GB)
1	踏み台サーバ	EC2	1	1	100
2	FTP サーバ	EC2	4	32	500
3	Web・AP サーバ 1 号機	EC2	8	64	1000
4	Web・AP サーバ 2 号機	EC2	4	32	1000
5	メールサーバ	EC2	2	8	100
6	法人番号付与サーバ 1 号機	EC2	16	64	1000

表 3-1 サーバ構成 (2022 年度)

7	法人番号付与サーバ 2 号機	EC2	16	64	1000
8	RDB サーバ	Aurora	8	64	ı
9		Aurora	2	16	-
10	RDF サーバ	Neptune	16	128	Ī
11		Neptune	4	32	ı
12	RDF サーバ(バックエンド)	Neptune	16	128	Ī
13		Neptune	4	32	-
14	全文検索サーバ	ElasticSearch	2	16	300
15		ElasticSearch	2	16	300
16	法人番号付与 RDB サーバ	Aurora	4	32	-
17		Aurora	2	16	-
18	Web・AP サーバ検証機	EC2	8	64	1000
19	RDB サーバ検証機	Aurora	2	16	-
20	RDF サーバ検証機	Neptune	4	32	-
21	全文検索サーバ検証機	ElasticSearch	2	16	300
22	Bucket with objects	S3	=	=	=
	合計		127	873	6600

(2) 利用状況に応じた最適な稼働環境の検討

Web・AP サーバを 2 台構成としてロードバランサによる処理振り分けや、データベースサーバ (RDB サーバ、RDF サーバ)を運用系と待機系とするなど各サーバの冗長化構成を行いサービス提供しており、当年度の環境として問題なく稼働できている。ただし、Gビズインフォの利用量 (特に API アクセス数) は年々増加しており、当年度は円安の影響によりランニングコストにおいても増加している状況にあった。ランニングコストについては利用状況の変化や為替の変動をウォッチしつつ見直しを行っていく必要がある。当年度においては超過するコストを低減するため検証機の縮退運転等を実施した。

なお、EC2 インスタンスはコストを抑えるために AWS リザーブドインスタンスとなっているため、インスタンスの拡張・縮退については長期的な変更計画となる。また、法人番号付与サーバは利用状況に応じた拡張・縮退の運用が可能となっている。

- 4. Gビズインフォの保守・運用
- 4.1.1. 保守・運用の実施

(1) 稼動監視

(ア)Web 画面へのアクセス状況

画面アクセスにおけるサービス状況を運用監視し、アクセス状況については下段に記載する訪問者、訪問数、ページとしてそれぞれの推移を月毎にまとめた。

当年度の傾向としてアクセス数が多く、前年度のアクセス状況を上回っている状況である。

特異なアクセス傾向のあった11月については、新規ホストから約5百万アクセス (ページ)もの 大量アクセスがあったことを確認しており翌月には他ホストと同レベルに推移したため同月のみ の一過性のものだと判断している。

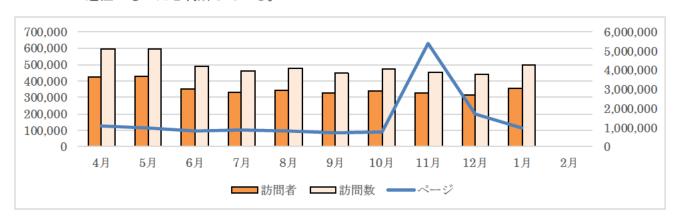


図 4.1-1 画面アクセス状況

図表における左右目盛りは、棒グラフは左目盛り、折れ線グラフは右目盛りを示す。

集計単位・用語は解析ツールで採用している AWStats のものに同じ。

- ✓ 訪問者・・・ユーザ数に相当。IP アドレス数でカウントしたもの。
- ✓ 訪問数・・・セッション数に相当。

同じIPアドレスであっても一定時間の間隔があったものは複数カウントする。

✓ ページ・・・PV 数に相当。

(イ)API へのアクセス状況

API アクセスにおけるサービス状況を運用監視し、アクセス状況については下段に記載する訪問者、訪問数、ページとしてそれぞれの推移を月毎にまとめた。

当年度の傾向として、API アクセス数 (ページ) が前年度を大幅に上回っており、大量アクセスの約8割を特定ユーザが占めている。このため特定ユーザに対しては利用状況のヒアリングを実施し今後の利用方法について議論をしている。利用方法の変更がなければ、次年度には年間で20億件を超えるアクセス量となる見通しとなり、リソース面でも影響が発生し得る。

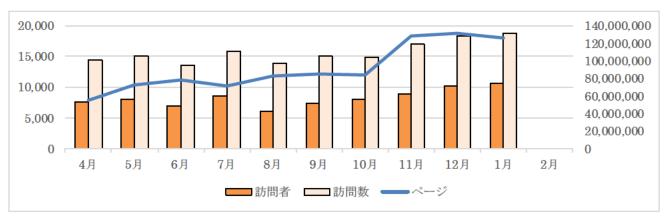


図 4.1-2 API アクセス状況

図表の見方については(ア)の図4.1-1の説明を参照のこと

(2) 障害対応

保守運用期間における障害発生は当年度 0 件。ここでの障害発生とはシステムトラブルによるサービス停止を指す。計画停止や経済産業省(運営事務局)との協議に基づく意図的なサービス停止は含まない。

表 4.1-1 発生した障害

	発生日	事象	影響のあったサービス	備考
1	なし	なし	なし	

表 4.1-2 重要度の定義

重要度	サービス停止	利用者	利用者への告知	特別対応
Α	停止	不特定多数	実施	
В	部分停止		必要に応じて実施	
С		限定される	原則実施しない	(*)

※API サービス停止時については API 利用者にメール通知する。

(3) セキュリティ監視

(ア) 脆弱性情報の収集

Gビズインフォで使用しているソフトウェアに関する脆弱性情報を脆弱性公表サイトから日々取得して予防保守を実施した。

(イ)侵入検知

fail2ban を使用して SMTP サーバへの不正アクセスを検知し計1,658件(当年度+651件)の IP アドレス遮断を実施した。(2月8日現在)

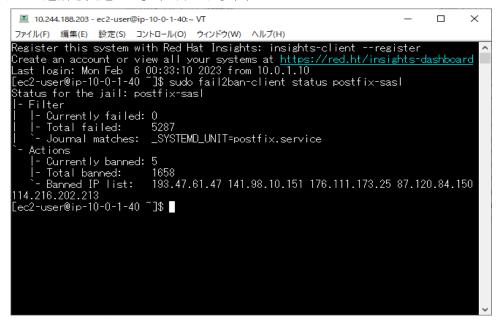


図 4.1-3 IP アドレス遮断状況

(4) 予防保守・定期保守

予防保守としてセキュリティ情報・アップデート情報を日々確認し、約2,500件(2月8日現在)の確認を実施している。予防保守によって検知した本システムに影響のあるセキュリティ情報・アップデート情報については、プロジェクトで定めた基準に基づいて対策を実施している。

基本的にソフトウェアのアップデートは定期保守にて実施するが、緊急性が高い事象については、回避や低減等といったアップデート以外の対応も含めて協議して実施する。当年度は緊急性の高い事象として、11月にAPサーバへの脅威1件を緊急対応した。

定期保守については四半期ごとに実施し、予防保守で検知した脆弱性について4件の対策を実施した。(2月8日現在)

4.2. 保守·運用作業準備

(1) 保守・運用作業計画書の作成と更新

保守・運用作業計画書を作成し、令和4年4月に承認を受けて本事業を開始した。

(2) 障害対応

障害管理手順を定め、令和4年4月に承認を受けて本事業を開始した。

(3) システム管理者マニュアルの更新

当年度の改修内容に応じて適宜修正を実施。最終版は納品物として納める。

(4) 利用者マニュアルの更新

当年度の改修内容に応じて適宜修正を実施。最終版は納品物として納めるほか、Gビズインフォサイトにて公開中。

- https://info.gbiz.go.jp/about/document/HowTo.pdf
- https://info.gbiz.go.jp/common/data/resourceinfo.pdf
- https://info.gbiz.go.jp/api/

(5) 定期的な報告

運用・保守の状況、稼働統計情報等を毎月実施している。

(6) 引継ぎ

前年度より引き続き令和4年4月からシステム保守を実施。

(7) 引き渡し

次年度保守・運用調達に際して、システムの引き渡しがスムーズにできるように各種ドキュメントを 作成・整理し、成果物として納品する(令和5年3月予定)。

4.3. 掲載情報の管理

当該サイトに掲載する情報について管理を実施した。具体的には以下の通り。

(1) サイトの活用事例の管理

下記のとおり実施(トップ>活用事例集)

https://info.gbiz.go.jp/example/

運営事務局において収集され提供、公開の許諾の得られたコンテンツを公開することとしている。当年度においては追加・変更・削除は生じていない。

(2) 「お知らせ」追記

下記のとおり実施(トップ)お知らせ)

https://info.gbiz.go.jp/information/

Gビズインフォの利用変更が発生するイベント (サービスの改善や変更等) や運営に関する告知が発生した場合に、利用ユーザー (不特定かつ一般者が対象) に対して運営事務局指示のもと告知を行うこととしている。当年度においては計6回の告知を行った。

(3) 「法人活動情報追加のお知らせ」追記

下記のとおり実施 (トップ >法人活動情報追加のお知らせ)

https://info.gbiz.go.jp/data_additional_information/

Gビズインフォでは 2.1 章の取組により法人活動情報追加が行われており、特に手作業による情報追加は不定期で追加更新となるため、それらを実施したタイミングで告知を行うこととしている。当年度においては計33回の告知を行った。

(4) 本システムに関する記事情報の管理

下記のとおり実施(トップ)関連掲載記事)

https://info.gbiz.go.jp/related_article/

本事業と関係あるコンテンツとして、当年度において新たな関連掲載記事の紹介は0件となっている。また、定期的に以下の対応を実施している。

- ・ 公開先で非公開となった場合のリンク閉塞対応
- ・ 関連掲載から1年以上を経過した記事に対するトップページ表示からの除外

5. 事業実施上の知見

(1) 「営業品目の改修」に関する事項(その1)

営業品目コードの管掌はデジタル庁であり、コード体系に変更が発生した場合でもGビズインフォの影響を最小限に留めるため、Gビズインフォで定義する業種コード(収録は営業品目コード)はデジタル庁のものへのリンクとしているが、画面表示(ならびに画面 PDF 取得)についてはコード値から名称に変換するため内部にテーブルを保有している。営業品目コードに変更が生じた場合、このテーブルを修正する保守作業が必要となる。

(2) 「営業品目の改修」に関する事項(その2)

今回改修を行った REST-API においては、既存のユーザーへの影響を考慮し旧態のサービスを維持しているが、REST-API の大がかりなシステム更改を行う際には旧態のサービスは停止する必要がある。また、既存ユーザーへの配慮としてサービス継続できない旨は1年前から告知をしておく等のユーザー影響への考慮が必要である。

(3) 「BI ツールの活用」に関する事項

今回、解析可能とした情報は REST-API のアクセス状況、法人活動情報に関する情報だが、次のような利用状況への拡張も可能である。項目の拡張に関しては、gBizINFO で定めたセキュリティ要件に従って取り扱う項目を決定し必要に応じて加工・編集する必要がある。

- ・画面、SPARQL-API のアクセス状況
- ・その他、RDB にのみ保有されており、一般公開されていない情報

また、今後はバッチ機能の拡張により他の情報も今回構築したスキームを利用して閲覧できるが BI ツールの利用コストは別途発生する。現在、適用している QuickSight であればユーザーID の従量課金や SPICE 利用(解析を高速に実施できるインメモリ)の使用サイズにより利用コストは増加する。このため利用の拡大に合わせ、これらの費用を事業費に折り込んでいく必要がある。

(4) 「gBizINFO の将来構想」に関する事項

法人活動データのオープンデータサービスの基盤を実現することからスタートした本事業では、今日までの取組みがデータ品質の確保や運用を主体としたシステム中心の運営をしてきており、生じている課題として、法人活動データの利活用が活性化しない、利活用状況が目に見えない、官民からのデータ収集が捗らないという課題が見えてきている。

これらの課題に着目しgBizINFOの将来の方向性について議論を推し進めた。これらをシステム変革するためには具体的な施策としては以下が候補としてあげられた。政策との整合性や潜在的なニーズに符合していると考えられる。

- ・法人情報を利用する他システムと連携したワンスオンリーのスモールスタート
- ・データ利活用を活性化させるオンラインコミュニティの整備・構築
- ・民間からのデータ収集し、より活性化するための基盤やインセンティブ設計の取組み

(5) 保守・運用における成果指標

本案件の保守・運用作業を実施するにあたり、以下の考え方をもとに成果指標を選定している。

- a) いつでも計測可能なものであること
- b) 測定した数値によって実施効果が推量できるものであること
- c) 経年のベンチマークによって改善効果が分析できるものであること
- d) 定義づけが不動なものであって PDCA サイクルを確立でき、次年度以降にも組み込めるものであること

各成果指標について目標を設定して実行しており、これら成果指標に対する主な実績について以下 に記載する。

2022 年度の成果指標 (実績)

① 稼働維持

➤ HP稼働時間 24 時間 365 日 (※1)

▶ サービス稼働率

・画面サービス 100.0% (※1)
 ・REST-API サービス 100.0% (※1)
 ・SPARQL-API サービス 100.0% (※1)

- ▶ サービス停止除外時間
 - ・計画停止時間および、DDoS/DoS 攻撃等によるインターネット遮断時間
 - ・政府または政府機関の行為、法律・規則・命令の遵守、戦争、反乱、革命、 暴動などの不可抗力の事態によりサービス提供ができなくなった場合
 - ・その他、貴省が停止を必要と判断した場合 (※1) サービス停止除外時間を含まない。
- ② データ拡充 (2023年2月8日現在)
 - ▶ 登録レコード数(件) # Gビズインフォが保有する法人活動情報の登録件数 5,581,237件
 - ・特許データ 4,370,781件
 - ・特許以外 1,210,456 件
 - ▶ 追加レコード数(件) # 2022 年度に SE 作業でGビズインフォに登録した件数 89,097 件件(特許以外)
 - ・特許データ 10 ヵ年分(出願日:2012年~2021年)
 - 特許以外 89,097 件
 - ▶ データ品質
 - (a) 2022 年度受領したデータにおいて法人番号が特定されたもの
 - ・法人番号付番率 97.0% (Gビズインフォ側で付番したものを含む)
 - ・元付番率 96.8% (Gビズインフォ側で付番したものを除く)

(b) 2022 年度受領したデータにおいて正常登録できたもの

・登録率 91.2% (Gビズインフォ側で補正したものを含む)

・適合率 88.9% (Gビズインフォ側で補正したものを除く)

※22 年度の各府省データはこれからデータ登録を実施するため、データ品質の数字は大きく変動する見込み。

③ 利用促進(2023年2月8日時点)

▶ 画面訪問数 40 万以上/月

▶ API 訪問数 1.3 万以上/月

利用促進の成果指標にアクセス数を採用していない背景として、例年の傾向より訪問数(セッション数に相当)の件数に比べて月毎の偏差が極端になる傾向があり、月毎に評価する成果指標として訪問数を採用。アクセス数については参考値として年間統計を記録するに留める。当年度(2022年度)のアクセス数に関しては次紙の<参考>欄を参照のこと

<参考>2022 年度のGビズインフォ実績

(ア)データ登録件数(法人活動情報)2023年2月8日現在

	2021 年度	2022 年度
	登録レコード数 (件)	登録レコード数(件)
財務情報	4, 432	4, 553
特許情報	4, 106, 513	4, 370, 781
届出・認定情報	218, 793	204, 871
補助金情報	494, 428	576, 958
調達情報	288, 644	338, 750
表彰情報	61, 896	63, 720
職場情報	15, 026	20, 995
決算情報	30	5, 676
合計	5, 189, 732	5, 580, 628

(イ)アクセス数

2022 年度	連	面	AI	PI
	訪問数	PV 数	訪問数	アクセス数
2022 年 04 月	594, 579	1, 082, 612	14, 366	54, 929, 595
2022年05月	596, 818	953, 000	15, 018	72, 462, 514
2022 年 06 月	490, 102	802, 970	13, 664	77, 610, 560
2022年07月	463, 109	874, 030	15, 956	71, 595, 378
2022年08月	478, 229	824, 828	13, 892	82, 234, 610
2022年09月	446, 841	719, 575	15, 017	84, 450, 146
2022年10月	470, 675	785, 395	14, 972	83, 578, 746
2022年11月	457, 141	5, 388, 086	16, 964	128, 511, 580
2022年12月	441, 074	1,711,380	18, 327	131, 282, 521
2023年01月	496, 455	950, 398	18, 809	126, 379, 300
2023年02月	_	_	_	_
2023年03月	_		-	-