

平成29年度補正予算 「産業データ共有促進事業」 に係る支援委員会業務委託事業 報告書

2019年4月23日

本資料について

- 本資料は、「平成29年度産業データ共有促進事業補助金」において、一般社団法人環境共創イニシアチブ（SII）より、「産業データ共有促進事業」に係る支援業務を委託した一般社団法人データ流通推進協議会（DTA）による実績報告書です。
- 支援業務を通して得られた実績と効果、その支援において整備した教材について、産業データ共有を検討されている皆様にも広く周知・ご活用いただくため、公開を行っております。

1. 産業データ共有促進事業に係る支援委員会業務委託事業の概要

- 1.1 事業の背景・目的
- 1.2 実施事項
- 1.3 実施体制
- 1.4 実施スケジュール

2. 補助事業者へのアンケートの実施

- 2.1 アンケート実施概要
- 2.2 アンケート実施結果

3. 事業者支援教材整備

- 3.1 データ標準化に関する教材の開発
- 3.2 データカタログに適用する教材の開発
- 3.3 ドメイン語彙を策定する教材の開発

4. 事業者支援

- 4.1 データ標準化の教育実施
- 4.2 アドバイス実施

5. まとめ

- 5.1 本事業の成果

【別冊1】「IT標準化支援プログラム」説明会資料

【別冊2】「IT標準化支援プログラム」勉強会資料 -データカタログ篇-

【別冊3】「IT標準化支援プログラム」勉強会資料 -共通語彙篇-

1. 産業データ共有促進事業に係る 支援委員会業務委託事業の概要

1. 産業データ共有促進事業に係る支援委員会業務委託事業の概要

1.1 事業の背景・目的

2017年10月に、機械、データ、技術、ヒト、組織など様々なものの繋がりによって新たな付加価値の創出や社会課題の解決をもたらす「Connected Industries」の推進に向け、日本の産業の未来像である「Connected Industries東京イニシアティブ2017」が、世耕経済産業大臣より発表された。

これを踏まえ、経済産業省では、重点分野※の協調領域において、複数企業者間でのデータ収集・活用を目的として、平成29年度補正予算により「産業データ共有促進事業」を推進している。

この「産業データ共有促進事業」においては、2018年2月より募集を開始し、2018年7月18日に支援の対象となる補助事業者が決定し、経済産業省では、協調領域において、事業者等が保有するデータの活用（共有・共用）が促進されるように、その取組を支援している。

以上の背景を踏まえ、本事業においては、事業者等が保有するデータの活用（共有・共用）が促進されるよう、補助事業者に、データ互換性やデータ品質の有効性を説明するとともに、補助事業者の要望（アンケート等）に基づき、データ流通に関する支援（IT 基準策定等）を実施するものである。

※重点分野：自動運転・モビリティサービス、ものづくり・ロボティクス、プラント・インフラ保安、バイオ・素材、スマートライフ

1. 産業データ共有促進事業に係る支援委員会業務委託事業の概要

1.2 実施事項

本事業では、産業データの活用が促進されるよう、経済産業省「平成29年度補正予算 産業データ共有促進事業費補助金」にて採択された25の補助事業者に対し、データ共有の実態や課題・要望を調査し、その結果を踏まえてデータ共有に関する支援を行うものである。

事業者への支援の内容としては、全体に対する説明会/勉強会の実施と、希望する補助事業者に対しては、事業内容に即したアドバイスを実施した。

本事業の実施事項と実施概要を以下に示す。

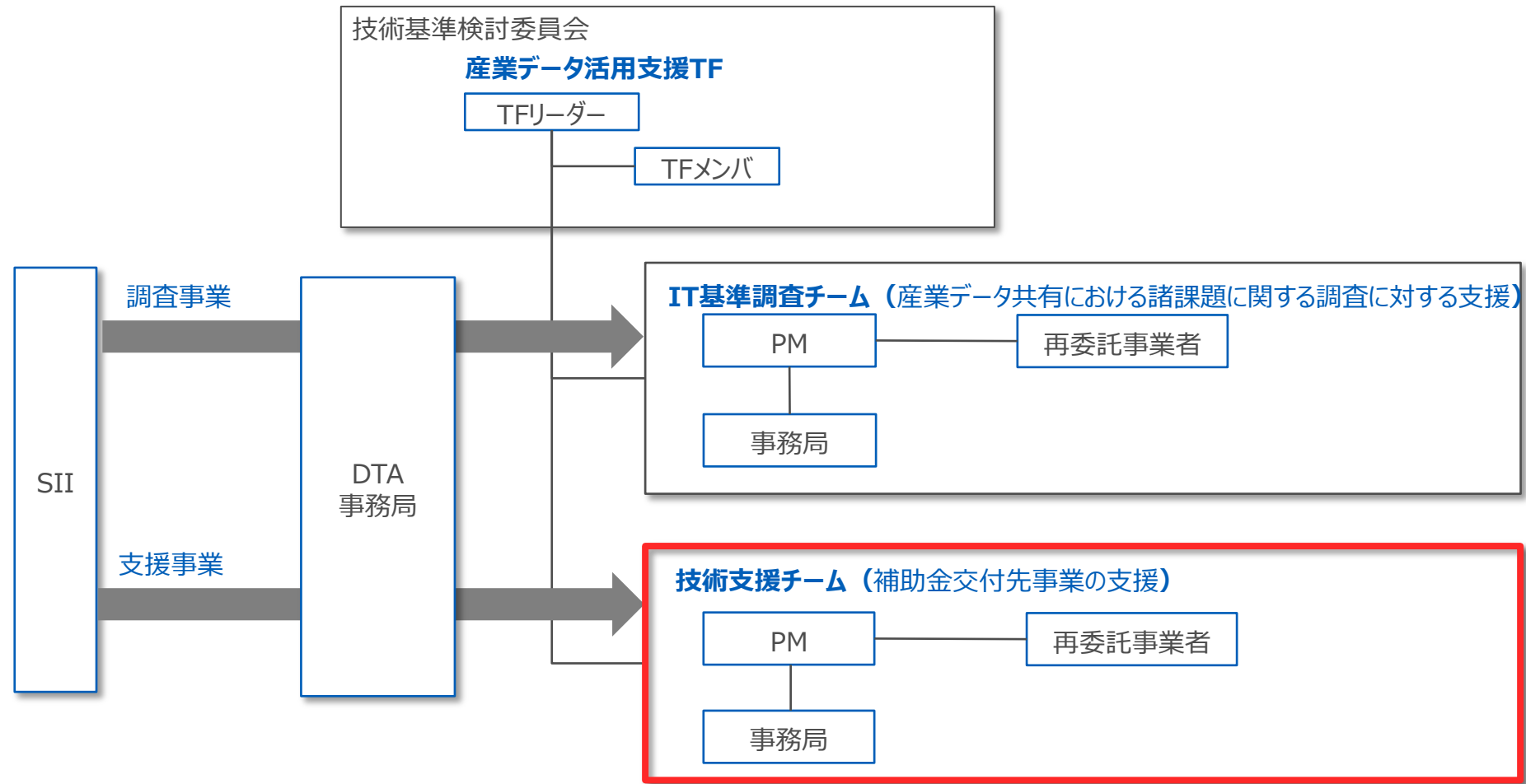
No.	実施事項		実施概要	調達仕様書の実施内容との対応	
1	補助事業者へのアンケートの実施		<ul style="list-style-type: none"> 25の補助事業者に対し、データ流通の促進のための「データカタログ策定」「共通語彙策定」やデータ品質の実態と要望等アンケート作成 アンケートの実施及び結果整理 	(1)補助事業者アンケート	
2	事業者支援教材整備	(1)データ標準化に関する教材の開発	<ul style="list-style-type: none"> データ標準化に関するIT規格の概要、標準化の利点、IT規格の適用の進め方等についての教材作成 	(2)事業者支援教材整備 (ア) データ標準化（互換性、品質）に関するIT規格の概要や、標準化の利点、IT規格の適用の進め方に関する一般教材の開発	
		(2)データカタログを適用する教材の開発	<ul style="list-style-type: none"> 補助事業者が推進する事業のデータカタログを、補助事業者自身で策定するための手順書作成 	(2)事業者支援教材整備 (イ) 「データカタログ」に適用する教材の開発	
		(3)ドメイン語彙を策定する教材の開発	<ul style="list-style-type: none"> 補助事業者が推進する事業のドメイン語彙を、補助事業者自身で策定するための手順書作成 	(2)事業者支援教材整備 (ウ) ドメイン語彙を策定する教材の開発	
3	事業者支援	(1)データ標準化の教育実施	説明会の実施	<ul style="list-style-type: none"> データ標準化に関するIT規格の概要、標準化の利点、IT規格の適用の進め方等を補助事業者に理解してもらうための説明会の実施 	(3)事業者支援 (ア) データ標準化の教育実施
			勉強会の実施	<ul style="list-style-type: none"> 「データカタログ」「ドメイン語彙」を補助事業者自身で策定できるための勉強会を実施 事業推進にあたり、個別のアドバイスを希望するか要望調査を実施 	
		(2)アドバイス実施	<ul style="list-style-type: none"> アドバイスを希望する補助事業者に対し、補助事業者の推進状況、要望に応じた、「データカタログ」「ドメイン語彙」策定に関するアドバイスを実施 	(3)事業者支援 (イ) データカタログ作成アドバイス (ウ) 分野データセットと標準語彙整理アドバイス	
4	報告書作成	作業報告書作成	<ul style="list-style-type: none"> No.1～3の実施事項における作業報告作成 本事業実施の中で課題となった事項のうち、IT規格にフィードバックすべき事項の整理 	(4)報告書作成	

1. 産業データ共有促進事業に係る支援委員会業務委託事業の概要

1.3 実施体制

本事業を実施する上で、データ流通推進協議会（以下、DTA）の技術基準検討委員会の下部組織として、産業データ共有促進事業を支援するためのTF（以下、産業データ活用支援TF）を上げた。本事業は、下図に示す産業データ活用支援TF内の、「技術支援チーム」を主体として実施した。また、本事業は、関連事業である調査事業と連携して実施している。

本事業の実施体制を以下に示す。主体制を赤枠で示す。



1. 産業データ共有促進事業に係る支援委員会業務委託事業の概要

1.4 実施スケジュール

本事業は、以下に示すスケジュールで実施した。

No.	実施事項	4月	5月	6月	7月	8月	9月	10月	11月	12月	1月	2月
0	定例				▲	▲		▲		▲		▲
1	補助事業者へのアンケートの実施				アンケート作成	実施	集計・整理					
2	事業者支援教材整備	(1)データ標準化に関する教材の開発			作成							
		(2)データカタログに適用する教材の開発			方式・構成検討	作成					フィードバック検討	
		(3)ドメイン語彙を策定する教材の開発			方式・構成検討	作成					フィードバック検討	
3	事業者支援	(1)データ標準化の教育実施	説明会の実施		▲ 第1回採択者向け説明会 (SII主催)	▲ 第2回採択者向け説明会 (SII主催)						
			勉強会の実施						勉強会	アンケート	集計	
	(2)アドバイス実施								アドバイス結果整理			
4	報告書作成									構成検討	作成	

2. 補助事業者へのアンケートの実施

2. 補助事業者へのアンケートの実施

2.1 アンケート実施概要

(1) アンケートの目的

本アンケートの目的は、IT標準化支援プログラムの実施に向け、以下事項を明確にすることである。

- ・補助事業者がデータカタログやドメイン語彙を策定するにあたり、現状抱える課題や要望
- ・データカタログやドメイン語彙策定のアドバイス等の希望有無

(2) アンケート対象者

本アンケートは、経済産業省「平成29年度補正予算 産業データ共有促進事業費補助金」にて、採択された以下の25補助事業者に対して実施した。

#	補助事業名称	幹事社名（共同申請者およびコンソーシアム事業者）
1	仮想統合解析手法を用いたビッグデータ利活用プラットフォーム構築事業	公益社団法人 関西経済連合会 (国立研究開発法人 国立循環器病研究センター、田辺三菱製薬株式会社、日本ユニシス株式会社)
2	PHR※及びNon-PHR領域における統合データ収集・活用を実現する基盤構築～異業種間情報共有と新産業・サービス創出に向けて～	国立研究開発法人 国立がん研究センター (株式会社ファインデックス、デジタルデータソリューション株式会社)
3	海外衛星データ連携調査および国内地上空間データ統合API等環境整備事業	さくらインターネット株式会社
4	中小サービス事業者への高度なデータ利活用推進プラットフォーム構築運営事業	株式会社 J T B (エブリセンスジャパン株式会社)
5	魅力あるデータを創るPHR※二次利用基盤事業	学校法人 慈恵大学 (株式会社アルム)
6	IoS(インターネットオブシプス)オープンプラットフォーム実証事業	株式会社シップデータセンター
7	集中型プラットフォームの社会実装に向けた検討事業	一般財団法人 石油エネルギー技術センター
8	履修履歴データ利活用一般化事業	株式会社大学成績センター (一般社団法人日本オープンオンライン教育推進協議会)
9	高精度 3次元地図の効率的なメンテナンス及びデータ収集/提供システム構築	ダイナミックマップ基盤株式会社 (株式会社スマートドライブ)
10	コンボリユースショナルデータの収集及び企業横断的活用事業	株式会社ちとせ研究所 (三井化学株式会社、味の素株式会社)
11	ロボットプラットフォームを核としたデータ共有・利活用の高度化	株式会社日本総合研究所 (学校法人 慶應義塾、ウォーターセル株式会社)
12	映像価値データ流通促進事業	日本電気株式会社 (中部電力株式会社、関西電力株式会社、アイホン株式会社)

※PHR：パーソナル・ヘルス・レコード

2. 補助事業者へのアンケートの実施

(つづき)

#	補助事業名称	幹事社名（共同申請者およびコンソーシアム事業者）
13	音声データを基にした製造業パイプラインのつまり予知・予兆診断事業	横河ソリューションサービス株式会社 (Hmcomm株式会社、日本ゼオン株式会社)
14	製造プラットフォームオープン連携事業	一般社団法人 インダストリアル・バリューチェーン・イニシアティブ (DMG森精機株式会社, 株式会社日立製作所, ファナック株式会社, 三菱電機株式会社)
15	マスターセンター事業	株式会社インテージテクノスフィア (株式会社ジャパン・インフォレックス、ジーエフケーマーケティングサービスジャパン株式会社、株式会社Payke)
16	運輸・物流動態データの共同利活用推進事業	SGシステム株式会社 (学校法人 国際大学)
17	水道業務データの利活用拡大に向けた検討事業	株式会社エヌ・ティ・ティ・データ
18	医療機器トレーサビリティデータバンク利活用実証事業	国立研究開発法人 国立国際医療研究センター (株式会社イノメディックス、株式会社エフエスユニマネジメント)
19	業界横断型の官民データ利活用プラットフォーム事業	一般財団法人 さっぽろ産業振興財団
20	保安高度化プラットフォーム実装促進検討	千代田化工建設株式会社 (J X T G エネルギー株式会社, コスモ石油株式会社)
21	発電設備運転データ共有事業	東京電力フュエル＆パワー株式会社 (鹿島共同火力株式会社)
22	ラストマイル自動移動サービスを核としたデータ連携・共有事業の創出	株式会社日本総合研究所 (関西電力株式会社)
23	需給最適化プラットフォーム事業	日本電気株式会社 (一般財団法人 日本気象協会)
24	ブロックチェーン活用によるヘルスケアデータ共有モデル構築事業	阪急阪神ホールディングス株式会社 (国立研究開発法人 理化学研究所、株式会社OKEIOS)
25	素材・化学分野における新事業情報共有プラットフォーム構築事業	ユニバーサルマテリアルズ インキュベーター株式会社 (有限会社化学品イー・データ開発)

(3) アンケート実施期間

平成29年8月1日～8月27日

(4) アンケート回答状況

25補助事業者/25補助事業者（回答率：100%）

Copyright©Sustainable open Innovation Initiative. All Rights Reserved.

2. 補助事業者へのアンケートの実施

(5) アンケート項目

今回実施したアンケート項目を以下に示す。

No.	区分	アンケート項目		アンケート詳細
1	データ互換性に関する取り組み状況と、課題・要望	(1)データカタログ	①データカタログに関する検討状況	データカタログ検討の有無を含む検討状況（本事業で検討/将来検討予定/検討済み/検討しない）
			②データカタログの活用範囲	データカタログにおけるデータを活用する範囲（事業メンバー内/分野内/分野横断）
			③データカタログの検討課題	データカタログを検討する上での課題
		(2)語彙	①語彙に関する検討状況	語彙検討の有無を含む検討状況（本事業で検討/将来検討予定/検討済み/検討しない）
			②語彙の活用範囲	語彙におけるデータを活用する範囲（事業メンバー内/分野内/分野横断）
			③語彙の検討課題	語彙を検討する上での課題
		(3)データ交換API	①データ交換APIに関する検討状況	データ交換API検討の有無を含む検討状況（本事業で検討/将来検討予定/検討済み/検討しない）
			②データ交換APIの活用範囲	データ交換APIにおけるデータを活用する範囲（事業メンバー内/分野内/分野横断）
			③データ交換API検討課題	データ交換APIを検討する上での課題
		(4)データ品質	①データ品質に関する検討状況	データ品質検討の有無を含む検討状況（本事業で検討/将来検討予定/検討済み/検討しない）
			②データ品質の活用範囲	データ品質におけるデータを活用する範囲（事業メンバー内/分野内/分野横断）
			③データ品質検討課題	データ品質を検討する上での課題
2	ガイドライン整備に関する要望			データ共有促進に向け整備を希望するガイドライン
3	IT標準化支援プログラムによる支援に向けた希望	(1)データカタログ	①データカタログに関する勉強会への参加希望	データカタログに関する勉強会参加希望の有無
			②データカタログに関するアドバイスの希望	データカタログ策定におけるアドバイス希望の有無
		(2)語彙	①語彙に関する勉強会への参加希望	語彙に関する勉強会参加希望の有無
			②語彙に関するアドバイスの希望	彙策に関するアドバイス希望の有無

2. 補助事業者へのアンケートの実施

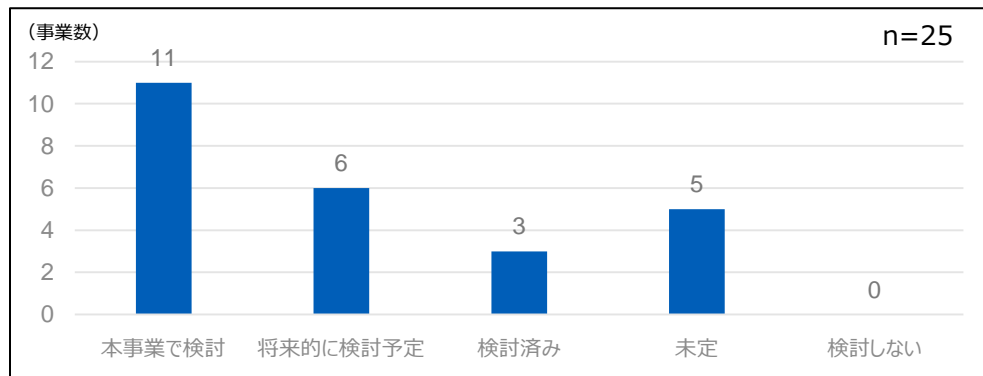
2.2 アンケート実施結果

2.2.1 データカタログ

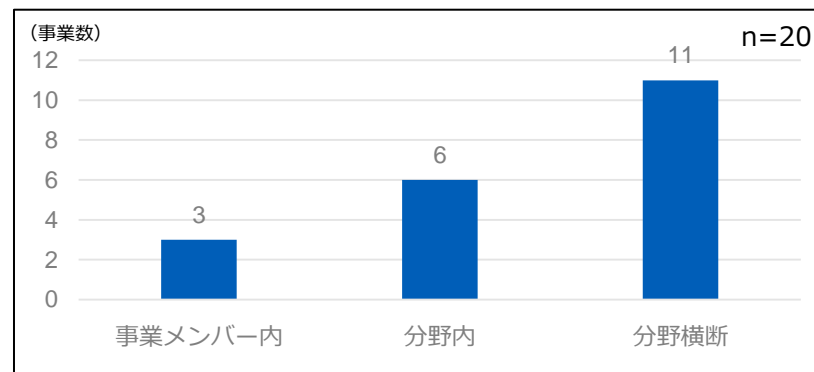
(1) 検討状況と活用範囲

データカタログの検討状況は、「検討済み」が3事業あるものの、「本事業で検討」が11事業、「将来的に検討予定」が6事業で計17事業が検討を想定している。また、活用範囲は、「分野横断：11事業」となっており、他分野との連携を検討している事業が多い。

① データカタログの検討状況



② データカタログの活用範囲



※対象、左記にて「本事業で検討」「将来的に検討予定」「検討済み」と回答した事業

(2) データカタログを策定する上での課題

データカタログの検討を行うとした17事業（「本事業で検討」「将来的に検討予定」）のうち、課題「有」は10事業あった。主な課題を以下に示す。「参照すべき基準がわからない」や「策定プロセスがわからない」といった課題を抱えている事業者が多い。

主な課題

【参照すべき基準がわからない】

- ・データを利活用するシーン、要件によって、求められるデータの粒度や構造が異なるドメインである。この背景の中で、分野横断でデータを利活用するカタログを作成する際に、どのような基準を参照すれば適切な形になるかわからない。
- ・可能な限り標準化（国際標準を含む）に則ったカタログ作成を目指しているが、何を基準にすべきか、手探りの状態になっている。

【策定プロセスがわからない】

- ・データカタログの具体的策定プロセス、記載項目、記載レベルに関する現実的な選択（落としどころ）
- ・データカタログの策定プロセスがわからない
- ・参画企業の国際競争力強化に資するカタログ内容とその共有方法の策定プロセスがわからない。

【その他】

- ・分野横断で共通化すべき項目と、独自に整理する項目の整理方法

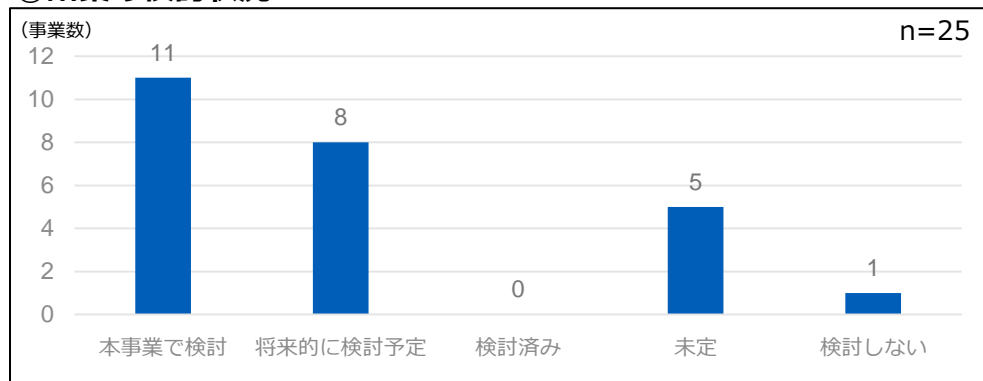
2. 補助事業者へのアンケートの実施

2.2.2 語彙

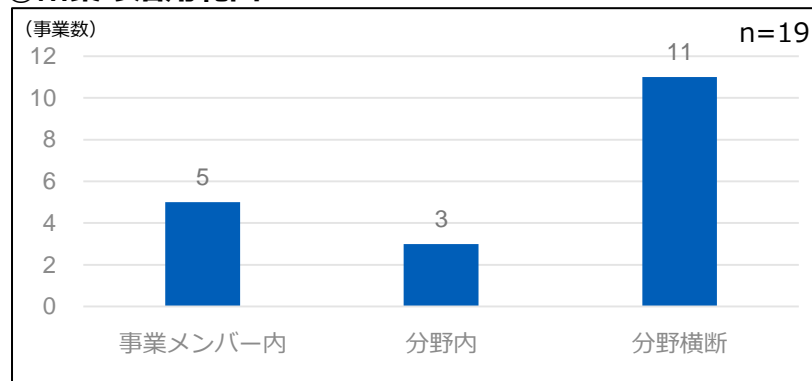
(1) 検討状況と活用範囲

語彙は、25事業中、「本事業で検討」が11事業、「将来的に検討予定」が8事業の計19事業が検討を想定している。活用範囲は、「分野横断：11事業」、「分野内：3事業」、「事業メンバー内：5事業」となっている。

① 語彙の検討状況



② 語彙の活用範囲



※対象は、左記にて「本事業で検討」、「将来的に検討予定」、「検討済み」と回答した事業

(2) 語彙を策定する上での課題

語彙の検討を行うとした19事業（「本事業で検討」「将来的に検討予定」）のうち、課題「有」は15事業あった。主な課題を以下に示す。「参照すべき基準がわからない」や「策定プロセスがわからない」といった課題を抱えている事業者が多い。

主な課題

【参考にするデータモデルの選定が困難】

- ・参考にする標準データモデルが多い中での採用するデータモデルの選定が課題

【策定プロセスがわからない】

- ・IPA共通語彙(コア語彙)を使用した語彙策定プロセスがわからない
- ・良いツールや作成した語彙のレビューをうける方法がわからない。
- ・策定プロセスがわからない

【その他】

- ・作成した語彙のレビュー、評価を委ねることのできる専門家の方針が、おそらく統一できない。

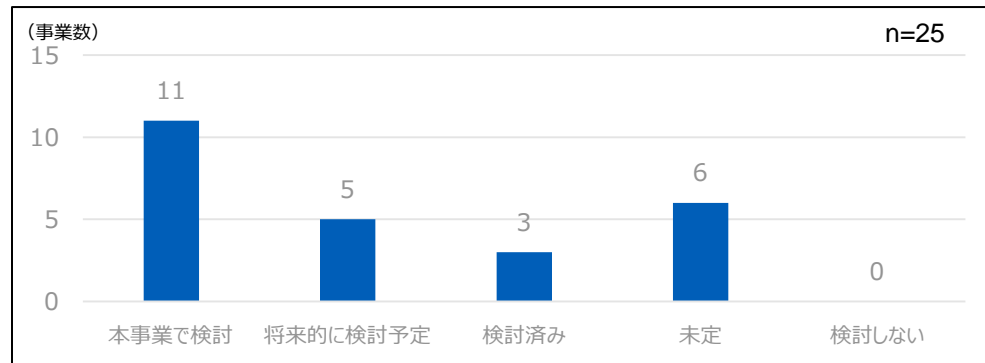
2. 補助事業者へのアンケートの実施

2.2.3 データ交換API

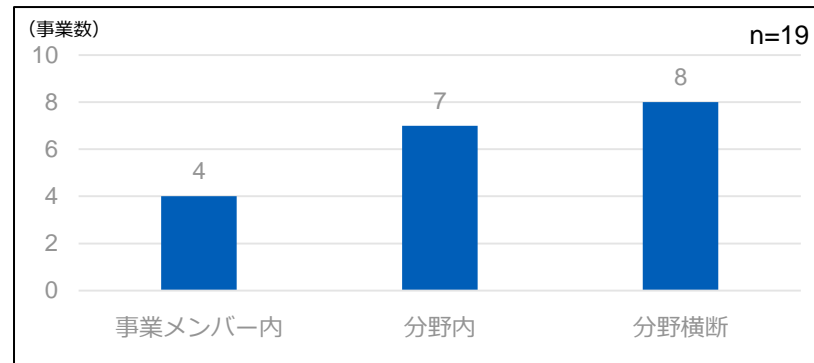
(1) 検討状況と活用範囲

データ交換APIは、25事業中、「検討済み」が3事業あるものの、「本事業で検討」が11事業、「将来的に検討予定」が5事業の計16事業が検討を想定している。活用範囲は、「分野横断：8事業」、「分野内：7事業」が多い。

① データ交換APIの検討状況



② データ交換APIの活用範囲



※対象は、左記にて「本事業で検討」、「将来的に検討予定」、「検討済み」と回答した事業

(2) データ交換APIを策定する上での課題

データ交換APIの検討を行うとした16事業（「本事業で検討」「将来的に検討予定」）のうち、課題「有」は11事業あった。主な課題を以下に示す。「国際標準がわからない」や「データ交換様式が整備されていない」といった課題を抱えている事業者が多い。

主な課題

「国際標準がわからない」

- ・データ交換API基準として、採用すべき、適切な国際標準がわからない。

「データ交換様式が整備されていない」

- ・保健指導やウェアラブル等のヘルスケアに関連するデータについてデータ交換様式が整備されていない。
- ・センサーから出力データ形式がメーカー・デバイスごとに統一されていないため、APIを作る際に事前の加工が必要である。

【その他】

- ・APIアクセスの際の認証方式やアクセス制限のかけ方をどうするか？
- ・他分野も活用できるデータ交換API基準の策定方法がわからない。

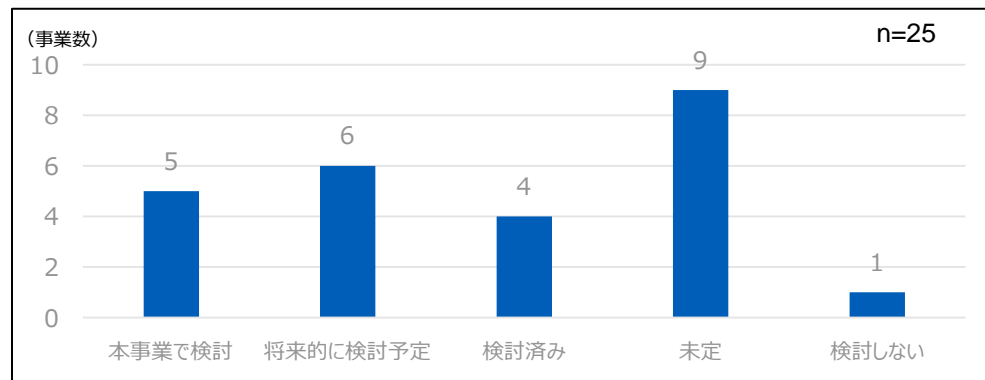
2. 補助事業者へのアンケートの実施

2.2.4 データ品質

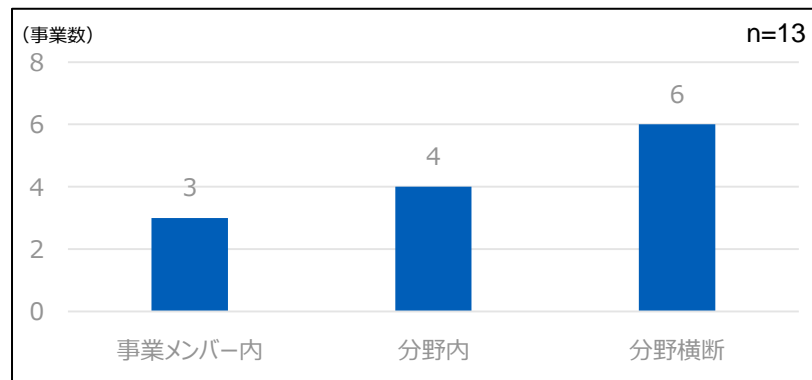
(1) 検討状況と活用範囲

データ品質は、「未定」が9事業と最も多いが、「本事業で検討」、「将来的に検討予定」が計11事業、そして、「検討済み」も4事業となっている。活用範囲は、「検討済み」を含めた中で回答があった13事業のうち、「分野横断」が6事業で、「分野内」は4事業、「事業メンバー内」は3事業となっている。

① データ品質の検討状況



② データ品質の活用範囲



※対象は、左記にて「本事業で検討」、「将来的に検討予定」、「検討済み」と回答した事業のうち未回答を除く13事業

(2) データ品質を策定する上での課題

データ品質の検討を行うとした11事業（「本事業で検討」「将来的に検討予定」）のうち、課題「有」は7事業あった。主な課題を以下に示す。「品質評価基準と評価方法がわからない」や「策定プロセスがわからない」といった課題を抱えている事業者が多い。

主な課題
「品質評価基準と評価方法がわからない」 <ul style="list-style-type: none">データに対する品質評価基準と評価方法がわからない。
「策定プロセスがわからない」 <ul style="list-style-type: none">データ品質基準の策定プロセスがわからない。データ品質向上の取り組みを国際競争力向上に資するものにするためにどのようなプロセスを踏めばよいか。
【その他】 <ul style="list-style-type: none">同じデータであっても、業界/分野によっては、期待される品質の精度が異なる。

2. 補助事業者へのアンケートの実施

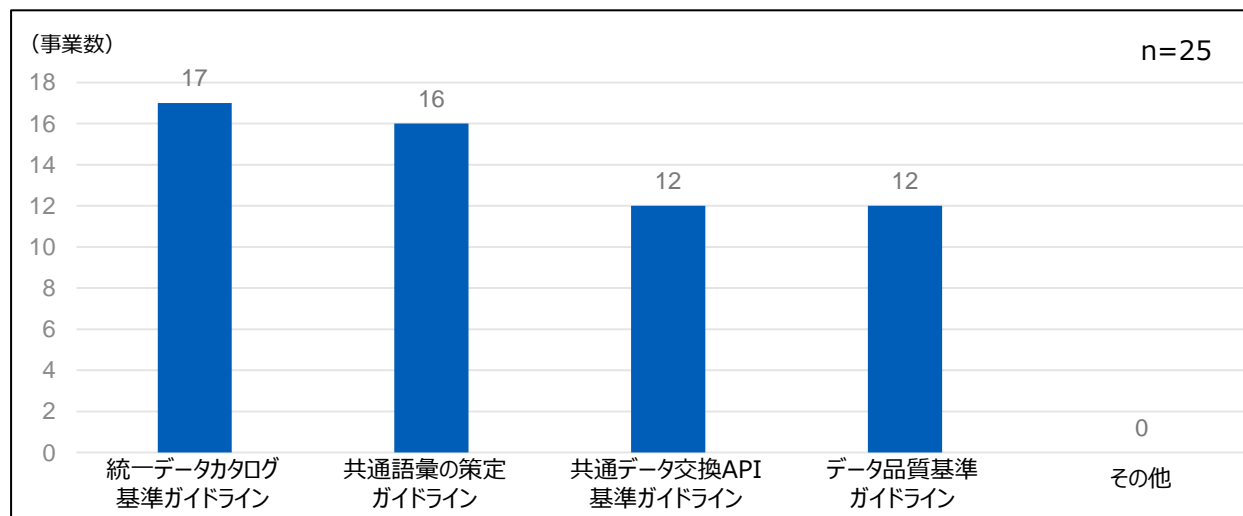
2.2.5 ガイドライン整備に対する要望

データ共有促進に向けた以下4つのガイドラインについて、整備を希望するかどうかを調査(複数回答可)した。

- 「統一データカタログ基準ガイドライン」
 - ・分野横断でデータを探しやすくするために、分野横断で流通可能なデータカタログ形式の策定方法を定めたガイドライン
- 「共通語彙の策定ガイドライン」
 - ・分野横断でデータ共有を容易にするために、他団体でもデータ項目の意味を把握しやすくする共通語彙の策定方法を定めたガイドライン
- 「共通データ交換API基準ガイドライン」
 - ・データ共有を容易化するために、通信基準、認証、プロトコルなどのデータ交換API基準を定めたガイドライン
- 「データ品質基準ガイドライン」
 - ・データ利用者が他団体のデータを安心して利用できるようにデータ品質基準を定めたガイドライン

調査結果を下図に示す。

「統一データカタログ基準ガイドライン」、「共通語彙の策定ガイドライン」の整備への希望が、それぞれ17、16事業と多く、「共通データ交換API基準ガイドライン」、「データ品質基準ガイドライン」についても半数近くが希望している。データ共有促進に向けたガイドライン整備に対する要望は、データカタログ、共通語彙を中心に、高い要望が存在している。



2. 補助事業者へのアンケートの実施

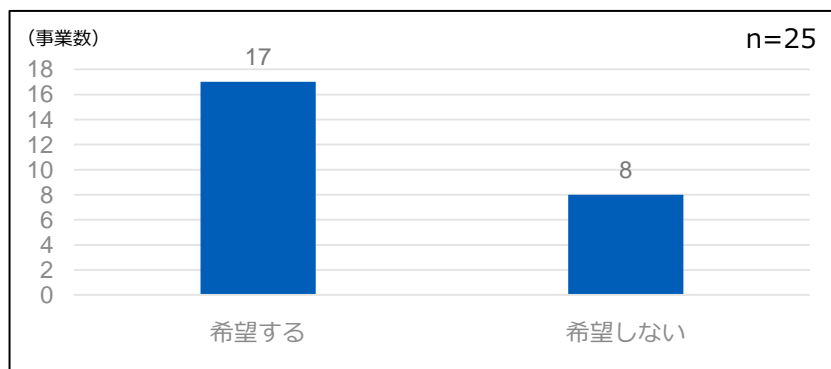
2.2.6 IT標準化支援プログラムによる支援に向けた希望

(1) データカタログ策定

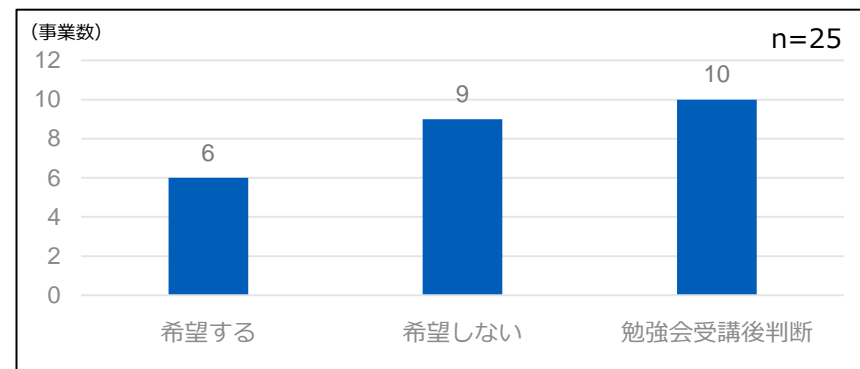
データカタログ策定に関する勉強会希望有無のアンケート結果は以下の通りである。

勉強会を希望する事業は、25事業中17事業と多いが、アドバイスについては、当該アンケート実施時点で希望する事業は、6事業であり勉強会受講後に判断するとしている事業が10事業である。

①勉強会の希望有無



②アドバイスの希望有無

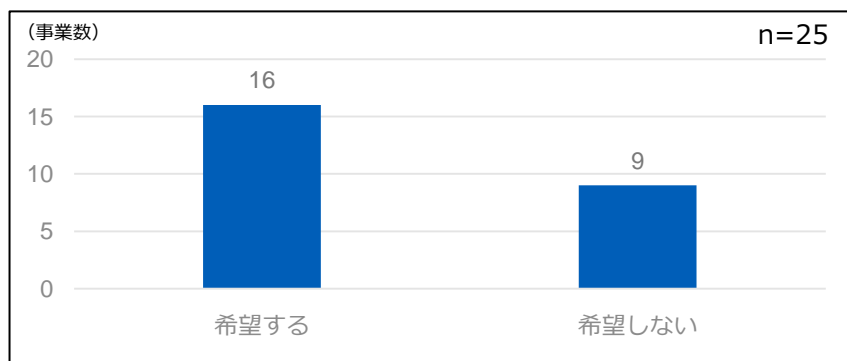


(2) 共通語彙策定

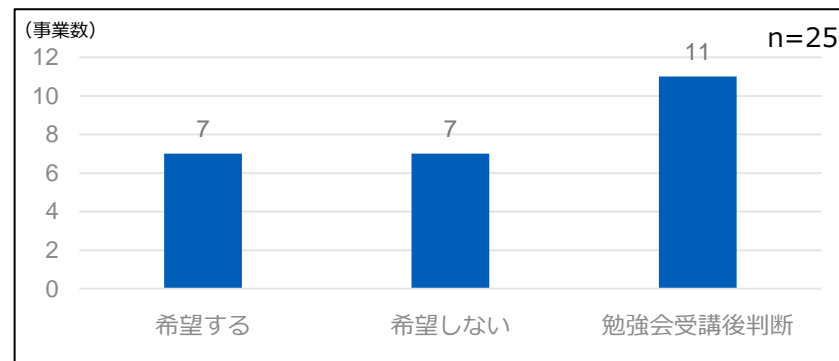
共通語彙に係る勉強会希望有無のアンケート結果は、以下の通りである。

勉強会を希望する事業は、25事業中16事業と多い。アドバイスについては、当該アンケート実施時点で希望する事業は、7事業、勉強会受講後に判断するとしている事業が11事業と、データカタログと同様の傾向となっている。

①勉強会の希望有無



②アドバイスの希望有無



3. 事業者支援教材整備

3. 事業者支援教材整備

3.1 データ標準化に関する教材の開発

(1) 開発の目的

産業データの共有の促進に向けては、ユーザーのデータ検索や比較が容易となるように、データカタログ基準を用いて、データカタログを運用する各所でデータカタログ形式を共通化することや、組織間、分野間でデータを共有しやすいように、IMI 共通語彙基盤（以下、IMI）等のIT規格を用いて、データの意味理解可能な環境を目指すことが望まれる。

本教材では、上記を踏まえ、産業データの共有促進に向けて、データカタログDTA基準やIMIを題材としたデータ標準化に関わる教材の開発を実施した。また、本教材は、初回の勉強会として、データ標準化の必要性の理解に重点を置き、産業データ共有の課題やIT規格適用のメリットなどのコンテンツに力を置いて構成している。

(2) 対象者

経済産業省の「平成29年度補正予算 産業データ共有促進事業費補助金」に採択された25の補助事業者を対象とする。

(3) 開発方法

IT規格として対象とするデータカタログDTA基準の開発状況や、独立行政法人情報処理推進機構（以下、IPA）のIMIの取組を踏まえつつ、データ標準化に関するIT規格の概要や標準化の利点等を含めた教材として開発した。

(4) 教材の内容

本教材は、事業者との初回の説明会の資料として構成したため、上記のデータ標準化の必要性に関わるコンテンツに加え、今回の事業における「IT標準化支援の位置づけ」や、「IT支援プログラムの概要」を含めた以下の構成とした。なお、教材自身は、【別冊1】に示す。

<教材の構成>

1. 事業における「IT標準化支援」の位置づけ
2. IT標準化支援プログラムの概要
3. 産業データ共有の効果とそのためIT基準
 - 3.1 産業データ共有の効果
 - 3.2 産業データ共有の課題
 - 3.3 産業データ共有の課題への対策 [データカタログ]
 - 3.4 産業データ共有の課題への対策 [共通語彙]
 - 3.5 産業データ共有の実現（データカタログ、共通語彙の整備）

3. 事業者支援教材整備

3.2 データカタログを適用する教材の開発

(1) 開発の目的

産業データの共有促進に向けては、データ提供者とデータ利用者間で、齟齬なく、データのやり取りができるようなデータカタログが運用されることが望まれる。また、データカタログは、データ利用者からの検索が容易となるように、データカタログ基準を用いて、データカタログを運用する各所で、共通のデータカタログ形式を用いることが望ましい。

本教材では、上記考えを踏まえ、データカタログやデータカタログに関わるIT規格（データカタログDTA基準）を理解、適用しやすいように、データカタログの必要性、IT規格（データカタログDTA基準）の概要、データカタログの作成手順を含めた教材を開発した。

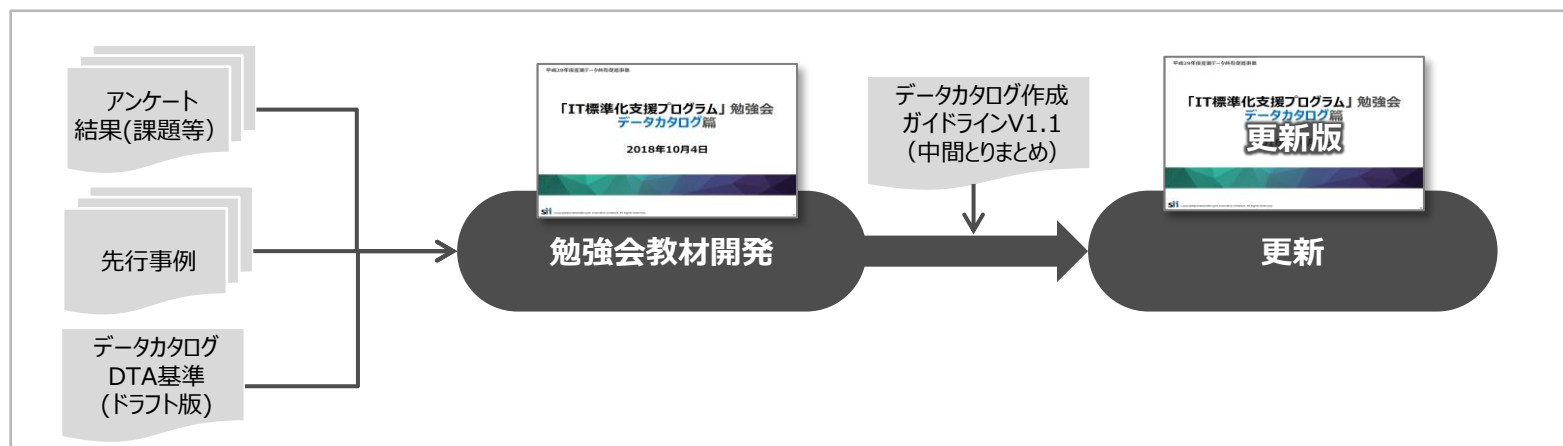
(2) 対象者

経済産業省の「平成29年度補正予算 産業データ共有促進事業費補助金」に採択された25の補助事業者を対象者として開発した。なお、補助事業者のIT知識が事業者によって異なることや、補助事業者以外にデータカタログを検討する事業者でも活用できるように、「データカタログとは」といった基礎的な内容を含めて、初心者でも理解しやすい教材を心掛けた。

(3) 開発方法

2018年10月の勉強会実施に向け、DTAで作成した「データカタログDTA基準（ドラフト版）」や、先行事例（国内外）及び、本事業で実施した補助事業者へのアンケート結果を踏まえ、作成した。本教材はSIIより、Web公開され、広く一般の事業者が参照可能な形となっている。

2019年2月には、DTAより、「データカタログDTA基準（ドラフト版）」の後継として、「データカタログ作成ガイドラインV1.1（中間とりまとめ）」が公開されたが、当該内容により、教材の更新を行い、更新版としてWeb公開を行っている。



3. 事業者支援教材整備

3.2 データカタログを適用する教材の開発

(4) 教材の内容

データカタログを適用する教材を、【別冊 2】に示す。教材の構成を以下の通りである。

<教材の構成>

1. データカタログとは何か？
2. データカタログとDTA基準の位置づけと意義
3. データカタログDTA基準の解説
4. 分野ごとのデータカタログ作成手順

3. 事業者支援教材整備

3.3 ドメイン語彙を策定する教材の開発

(1) 開発の目的

産業データの共有促進に向けては、データの共有を行う際に、組織間、分野間で異なる用語の意味理解が促進されるように、IMI 共通語彙基盤（以下、IMI）等のIT規格を用いて、人やコンピュータでもデータの意味理解可能な環境を目指すことが重要となる。

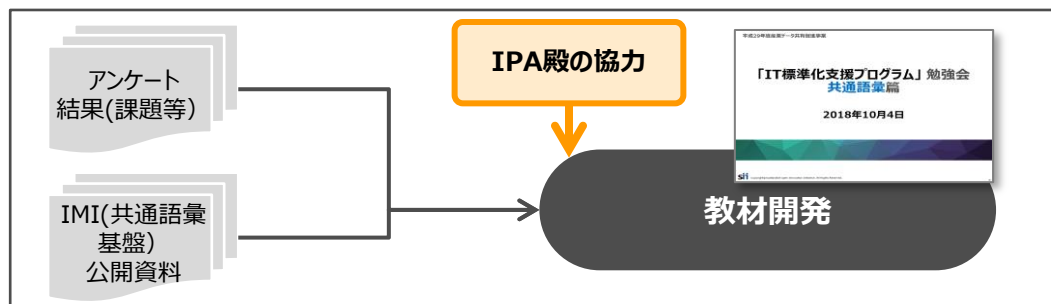
本教材では、上記考えを踏まえ、語彙に対する理解促進や、IMI適用を含めたドメイン語彙策定を行いやすいように、語彙を揃える意義やドメイン語彙の策定手順を含めた教材を開発した。

(2) 対象者

経済産業省の「平成29年度補正予算 産業データ共有促進事業費補助金」に採択された25の補助事業者を対象者として開発した。なお、補助事業者のIT知識が事業者によって異なることや、補助事業者以外にドメイン語彙を検討する事業者でも活用できるように、「そもそも語彙とはなにか」といった基礎的な内容を含めて、初心者でも理解しやすい教材を心掛けた。

(3) 開発方法

本事業で実施した補助事業者へのアンケート結果をもとに、IMIに関する公開資料や、IMIを運用しているIPA殿の協力のもと、教材を作成した。



(4) 教材の内容

ドメイン語彙を策定する教材を、【別冊3】に示す。教材の構成は以下の通りである。

<教材の構成>

1. そもそも「語彙」とは何か？
2. 「語彙」を揃える意義
3. IMI（共通語彙基盤）の概要
4. ドメイン語彙の策定（作業概要）
5. ドメイン語彙の策定 作業ステップ

4. 事業者支援

4. 事業者支援

4.1 データ標準化の教育実施

4.1.1 教育実施概要

(1) 説明会の実施

IT標準化支援プログラムによる補助事業者に対し、IT標準化に関する説明会を実施した。

No.	区分	実施日	対象者	実施内容
1	説明会	2018.6.5	平成29年度補正予算 産業データ共有促進事業費補助金 一次公募採択事業者（13事業）	作成した『「IT標準化支援プログラム」説明会資料』を用いて、以下を説明 1. 事業における「IT標準化支援」の位置づけ 2. IT標準化支援プログラムの概要 3. 産業データ共有の効果とそのためIT基準
2		2018.7.24	平成29年度補正予算 産業データ共有促進事業費補助金 二次公募採択事業者（12事業）	

説明会実施風景を以下に示す。



「IT標準化支援」の位置づけ、
IT標準化支援プログラムの概要説明



「産業データ共有の効果とそのため
IT基準」説明

4. 事業者支援

(2) 勉強会の実施

勉強会の参加を希望する補助事業者に対し、データカタログ・語彙策定に関する勉強会を実施した。データカタログ、共通語彙の両方を希望する補助事業者が多かったため、同一日に実施した。

No.	区分	実施日	対象者	参加者数		実施内容
				参加事業数	参加者数	
1	勉強会 データカタログ篇	201810.4 13:00 ～15:00	平成29年度補正予算 産業データ共有促進事業費補助金 採択事業者のうち、勉強会を希望する事業者	16事業	31人	・作成した『「IT標準化支援プログラム」勉強会資料-データカタログ篇-』を用いて以下を実施 1. データカタログ策定手順について 2. データカタログ利用イメージ 3. 今後の予定 4. アドバイス希望有無アンケート
2	勉強会 共通語彙篇	201810.4 15:30 ～17:30	平成29年度補正予算 産業データ共有促進事業費補助金 採択事業者のうち、勉強会を希望する事業者	16事業	29人	・作成した『「IT標準化支援プログラム」勉強会資料-共通語彙篇-』を用いて以下を実施 1. 共通語彙策定手順について 2. 今後の予定 3. アドバイス希望有無アンケート

説明会実施風景を以下に示す。



<データカタログ策定手順説明>



<データカタログ利用イメージ紹介>



<共通語彙策定手順説明>

なお、勉強会実施後、アドバイス希望有無のアンケートを実施した。次ページにアンケート結果を示す。

4. 事業者支援

4.1 データ標準化の教育実施

4.1.2 勉強会実施後のアンケート結果

勉強会終了時に、データカタログ策定、ドメイン語彙策定に関する理解度及び今後アドバイスを希望するかどうかのアンケートを行った。

(1) アンケート概要

理解度及び今後アドバイスを希望有無のそれぞれの概要を以下に示す。

アンケートの種類	目的	対象	回答状況
理解度アンケート	勉強会に参加した個人の「データカタログ策定」「ドメイン語彙策定」に関する理解度の把握	勉強会に参加した個人 ・「データカタログ篇」：28人 ・「共通語彙篇」：27人	・「データカタログ篇」：27人/28人 ・「共通語彙篇」：25人/27人
アドバイス希望有無アンケート	本事業内でのアドバイス実施希望補助事業者の把握。 アドバイス実施計画に利用する。	勉強会に参加した補助事業者：15事業	15/15補助事業者

4. 事業者支援

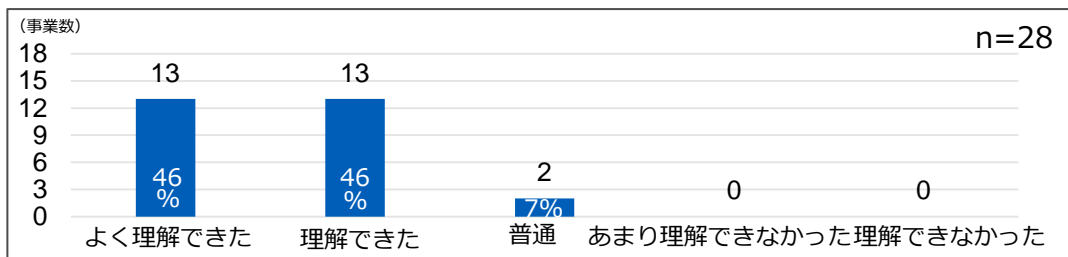
(2) 理解度アンケート結果

(i) データカタログ篇

勉強会の各章について理解度を5択（よく理解できた/理解できた/普通/あまり理解できなかった/理解できなかった）の選択方式と「疑問点、質問等」を調査した。調査結果を以下に示す。

いずれのコンテンツについても、「普通」以上が90%以上であり、高い評価を得ることができたと考える。

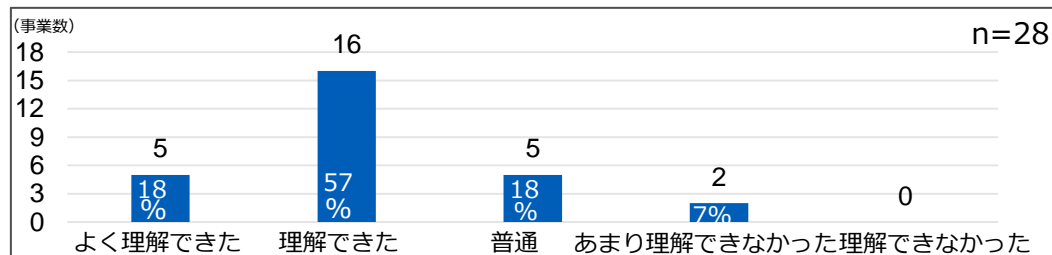
① データカタログとは何かについて



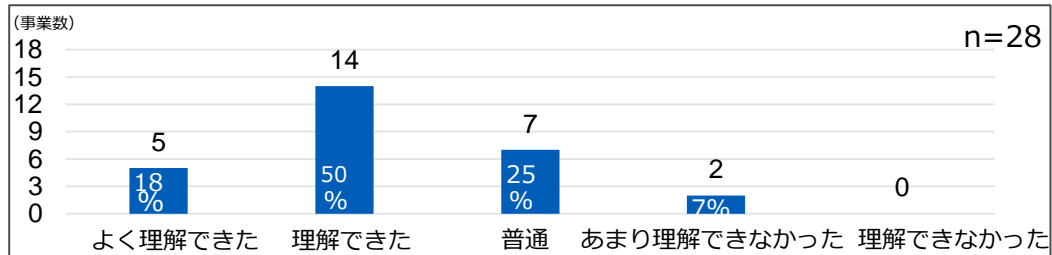
④ 疑問点、質問等（自由記入）

番号	コメント
1	概念は良く理解できた。あとはやはり実務にのせて見てと感じている。
2	今日の時点では質問事項はありません。まずは自分のテーマに関してデータカタログ（データプロファイル）を作成し、わからない点について後日質問させて下さい。
3	分野ごとのデータカタログ策定手順についての説明が、非常に実践的で良かったです。
4	人間がデータをやりとりするためのカタログとして、基準を出して頂けるのは有用であると感じた。「項目をチョイスする」というのが将来的に未知な共有や、データのやりとりに問題が起きないか、というのが少し気になる。
5	データカタログの各データ項目に何を書くかは、かなり自由度が高そうなので、データカタログの相互運用面で課題になりそうに感じた。
6	複数のデータを重ねてデータ化している。別々にカタログ化するか、まとめてカタログ化するのか検討したい
7	データカタログがないことよっての不都合が現状ではあまり想定出来ていないが、どの程度普及しているものなのか
8	後半のEverySenseの説明資料を電子データで配布してほしい。
9	具体的な例によりカタログ作成手順を示していただけるとイメージがしやすいのですが、抽象的なご説明あったのでとても難しく感じました。
10	カタログの構造化に意味はあると思うが、設定する値の標準化についてどう考えればよいか知りたかった。

② データカタログDTA基準について



③ 分野ごとのデータカタログ基準策定手順について



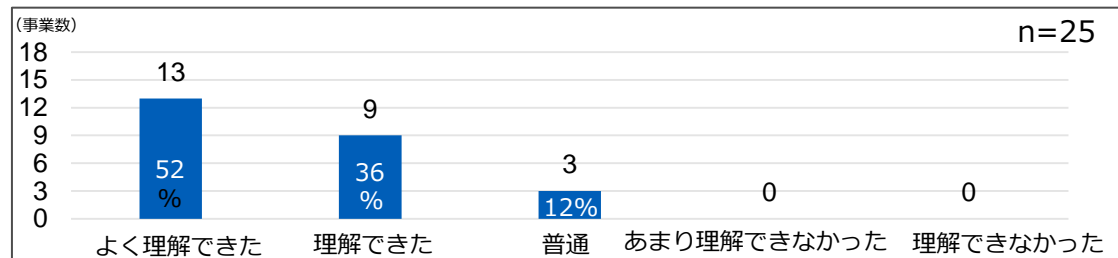
4. 事業者支援

(ii) 共通語彙篇

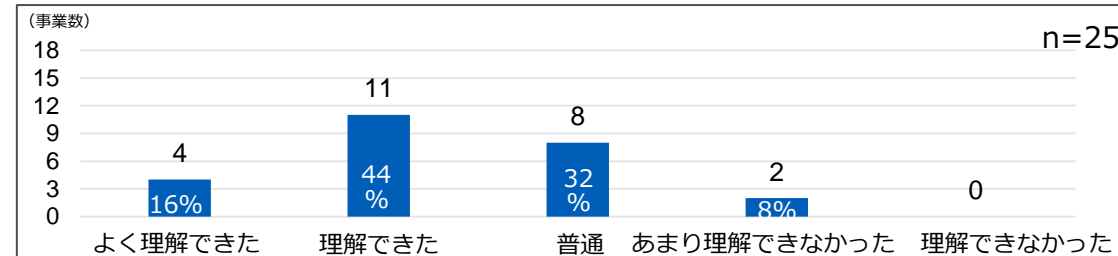
勉強会の各章について理解度を5択（よく理解できた/理解できた/普通/あまり理解できなかった/理解できなかった）の選択方式と「疑問点、質問等」を調査した。調査結果を以下に示す。

いずれのコンテンツについても、「普通」以上がほぼ90%以上であり、高い評価を得ることができたと考えられる。ただし、データカタログ篇と比べると「よく理解できた」、「理解できた」という回答の割合が減っており、データカタログより難しかったことが伺える。

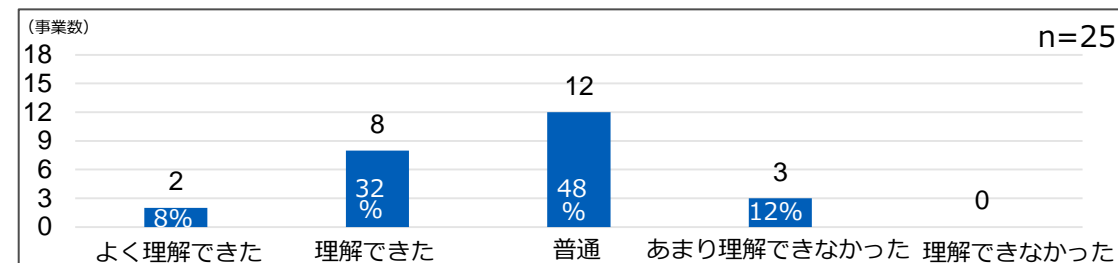
① 語彙とは何かについて



② IMI（共通語彙基盤）について



③ ドメイン語彙策定の策定ステップについて



④ 疑問点、質問等（自由記入）

コメント	
1	データプロファイルを作っているので、共通語彙に則って作成できればと考えています。
2	特にありません。 現在私達はSTEP2の段階ですが、STEP3に入るにあたりお力添えをいただければ、と思います。
3	スコープの定義、語彙の網羅性のチェック方法、別の辞書のTransfer等どう考えたら良いか知りたかった。
4	数値データの「単位」の階層がすでに策定されているのか？
5	住所など項目として1つになっている場合、複数項目に別れている場合があるが、プロパティ用語としてはどのように設置すべきか？

4. 事業者支援

(3) アドバイス希望有無アンケート結果

アドバイスを希望するとした補助事業と希望するアドバイスの種類を以下に示す。勉強会に参加した15補助事業のうち、アドバイスを希望する補助事業は、8事業となっている。

No.	補助事業名	幹事社名（共同申請者及びコンソーシアム事業者）	データカタログ	共通語彙
1	PHR及びNon-PHR領域における統合データ収集・活用を実現する基盤構築	国立研究開発法人 国立がん研究センター (株式会社ファインデックス、 デジタルデータソリューション株式会社)	○	○
2	海外衛星データ連携調査および国内地上空間データ統合API等環境整備事業	さくらインターネット株式会社	○	-
3	中小サービス事業者への高度なデータ利活用推進プラットフォーム構築運営事業	株式会社 J T B (エブリセンスジャパン株式会社)	○	○
4	集中型プラットフォームの社会実装に向けた検討事業	一般財団法人 石油エネルギー技術センター	-	○
5	映像価値データ流通促進事業	日本電気株式会社 (中部電力株式会社、関西電力株式会社、 アイホン株式会社)	○	○
6	音声データを基にした製造業パイプラインのつまり予知・予兆診断事業	横河ソリューションサービス株式会社 (Hmcomm株式会社、日本ゼオン株式会社)	○	○
7	マスターセンター事業	株式会社インテージテクノスフィア (株式会社ジャパン・インフォレックス、ジーエフ ケー マーケティングサービスジャパン株式会社、 株式会社Payke)	○	○
8	ブロックチェーン活用によるヘルスケアデータ共有モデル構築事業	阪急阪神ホールディングス株式会社 (国立研究開発法人 理化学研究所、 株式会社OKEIOS)	○	○

4. 事業者支援

4.2 アドバイス実施

4.2.1 アドバイス実施概要

勉強会終了後にアドバイスを希望する事業に対し、アドバイス会を実施した。実施事業と実施日は以下の通り。データカタログと共通語彙の両方を希望する補助事業者に対しては、同一日に実施した。

No.	補助事業名	幹事社名（共同申請者及びコンソーシアム事業者）	データカタログ	共通語彙	実施日
1	PHR及びNon-PHR領域における統合データ収集・活用を実現する基盤構築	国立研究開発法人 国立がん研究センター (株式会社ファインデックス、 デジタルデータソリューション株式会社)	○	○	11/12
2	海外衛星データ連携調査および国内地上空間データ統合API等環境整備事業	さくらインターネット株式会社	○	—	1/9
3	中小サービス事業者への高度なデータ利活用推進プラットフォーム構築運営事業	株式会社 J T B (エブリセンスジャパン株式会社)	○	○	12/14
4	集中型プラットフォームの社会実装に向けた検討事業	一般財団法人 石油エネルギー技術センター	—	○	12/4
5	映像価値データ流通促進事業	日本電気株式会社 (中部電力株式会社、関西電力株式会社、 アイホン株式会社)	○	○	・12/19 ・1/9 ・1/18
6	音声データを基にした製造業パイプラインのつまり予知・予兆診断事業	横河ソリューションサービス株式会社 (Hmcomm株式会社、日本ゼオン株式会社)	○	○	11/12
7	マスターセンター事業	株式会社インテージテクノスフィア (株式会社ジャパン・インフォレックス、ジーエフケー マーケティングサービスジャパン株式会社、株式会社 Payke)	○	○	11/20
8	ブロックチェーン活用によるヘルスケアデータ共有モデル構築事業	阪急阪神ホールディングス株式会社 (国立研究開発法人 理化学研究所、 株式会社OKEIOS)	○	○	11/12

4. 事業者支援

4.2 アドバイス実施

4.2.2 アドバイス会概要

アドバイス会では、データカタログやドメイン語彙に対するアドバイスを中心としつつも、その他のデータ共有に関わる課題についてもアドバイスを実施した。以下にアドバイスを希望した各補助事業のアドバイス会の概要を示す。

(1) PHR及びNon-PHR領域における統合データ収集・活用を実現する基盤構築

- ✓ 各所から取得するデータが異なる標準等を採用しているため、それに対応したシステム構成についてアドバイスを行うとともに、データ抽出にあたっての単位などの項目（メタデータ）付与の必要性や、利用シーンの整理に向けたデータカタログDTA基準のデータジャケット部の活用についてアドバイスを実施。

(2) 海外衛星データ連携調査および国内地上空間データ統合API等環境整備事業

- ✓ データカタログの適用箇所や、データカタログの構成方法についてアドバイスを実施。

(3) 中小サービス事業者への高度なデータ利活用推進プラットフォーム構築運営事業

- ✓ 実装するデータカタログとのギャップ分析を実施するとともに、ドメイン語彙の作成にあたって参考になる語彙についてアドバイスを実施。

(4) 集中型プラットフォームの社会実装に向けた検討事業

- ✓ IMIやData Model Description（以下、DMD）の概念の理解促進や、ドメイン語彙の策定方法についてアドバイスを実施。

(5) 映像価値データ流通促進事業

- ✓ 本事業の検討会メンバーに対し、データカタログ並びにドメイン語彙の理解促進に向けた勉強会を実施するとともに、勉強会の前にデータカタログ記入の事前トライアルを実施。当該活動を通じて、データカタログの適用箇所の整理、データカタログの構成方法、記入方法等についてアドバイスを実施。

(6) 音声データを基にした製造業パイプラインのつまり予知・予兆診断事業

- ✓ データカタログに掲載する項目や、段階的なデータ共有の可能性についてアドバイスを実施。

4. 事業者支援

4.2 アドバイス実施

4.2.2 アドバイス会概要

(7) マスターデータセンター事業

- ✓ 商品マスターを統合したカタログに加え、各マスターの出自などを管理した内部系のカタログの必要性や、データカタログの記入方法（データ詳細部の項目整理等）についてアドバイスを実施するとともに、コア語彙を踏まえた語彙の整理方法についてアドバイスを実施。

(8) ブロックチェーン活用によるヘルスケアデータ共有モデル構築事業

- ✓ データの共有の難しいヘルスケアデータについて段階的なデータ共有の可能性についてアドバイスを行うとともに、ID/認証に関わる課題を共有。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

本節では、アドバイス会における主なアドバイスを下記の観点から示す。

- (1) 業界における課題、 (2) IT基準自体に関わる課題、 (3) IT基準の活用に向けた課題、 (4) その他課題等

(1) 業界における課題

(i) 各所での異なる標準等の採用

各所からデータを収集する際に、各所で異なる標準等を採用しているため、収集したデータの活用が困難な状況になっている業界もあり、当該事業に対しては、データの差異を吸収する辞書の整備や、データ分析がしやすいデータマート、データレイクなどのシステム構成についてアドバイスを実施した。

- 診療データは、全国からデータを集めようとする病院ごとにローカルコードがある。
- 領域に特化しており、各学会が疾病ごとにデータを定めており、学会ごとにも差がある。また、国際標準、国内標準も差異があり、疾病に応じて採用する標準が異なっている。言葉の表記（赤血球、RBC）、などもそろわない。
- 標準化する際に、どこを目指していけばよいか非常に難しい。すべてのコードの統一化も難しい。

- ✓ 各所から収集するデータの差異を吸収する辞書を整備することが考えられる。
- ✓ 具体的には、下記の整備が考えられる。
 - ✓ 辞書
 - ✓ データをためるデータレイク
 - ✓ 外部提供に向けたデータレイクのデータカタログ
 - ✓ 用途に応じてデータレイクのデータを加工するデータマート
 - ✓ 外部のデータ提供先のニーズをまとめた辞書（どの疾病のどのレベルのデータが必要で、どのような支援が必要かを示したもの）
- ✓ 辞書やデータマートの項目は、IMIを活用すると、他業界との互換性の高い語彙になる。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(1) 業界における課題

(ii) データ共有への抵抗感

一部の補助事業では、匿名化を行い、機微情報を削除するなどの対応も視野に入れているが、医療データ等の個人の機微データを扱う事業や、保安データ等の企業の機微データを扱う事業では、データ共有への抵抗感が課題となっている。NDAの締結前後等のタイミングでデータを見せる範囲が異なることや、外部に見せる範囲と内部管理するデータの範囲が異なることから、データカタログにより、レベルを分けた管理を行いつつ、段階的にデータ共有を行う可能性について、アドバイスを行った。

- 他社に保安系のデータを共有することに抵抗がある
- 各社ともに保安データを外に出すことに抵抗感がある。仮に出すのであれば、どの施設のどの場所どのパーツかがわからないようにしてほしいといわれている。
- 医療データのような機微なデータは個人が納得した上でないとデータを共有したがる

- ✓ サマリーなど、開示できる範囲で見せる範囲と、NDA締結後に見せる範囲、そして、内部で保有しておく範囲と3層程度になる可能性があり、レベルを分けたカタログ管理もありうる。
- ✓ データジャケット部を開示し、次に、NDAを結んでデータ詳細部といった見せ方も想定できる。
- ✓ 開示できるデータが少ないだけに、後から細かいデータがほしいという話になることは容易に類推される。そのため、ここまですら出せる、交渉次第といったものを事前に提供者サイドと整理しておくことが考えられる。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(1) 業界における課題

(iii) データのマッシュアップ

① データの粒度の統一

データをマッシュアップしていくためには、時間やエリアなどのデータの粒度をそろえることが望ましいが、どう揃えていくかが課題として挙げられている。当該課題に対しては、なるべくスタンダードを使うことで、無理なく共通化していくことを推奨している。

- 各所からデータを収集する場合に、データを統合分析しやすいように、時間やエリアの単位などの粒度をそろえていく必要があるが、どう揃えていくかはまだ見えていない。

✓ エリアを特定するスタンダードがあるなら、それを使うのがよいだろう。

② ライセンスの統一

データをマッシュアップしていくためには、データが使いやすいように、データのライセンスが統一されていることが望ましいが、データ提供者毎にライセンスにばらつきがあり、ばらつきの抑制に苦慮しているとの意見が挙げられている。データカタログを作成するという本来の用途とは異なるが、データカタログDTA基準のデータ利用条件部は、データの契約等にあたっての検討事項が記載された項目群と捉えることもでき、当該項目群の活用の可能性についてアドバイスを実施した。

- データ提供者ごとにばらつきのあるデータライセンスだと、利用者が不安がって使い勝手が悪いため、事業者ごとに異なるライセンス条件の調整をしているが、ばらつきの抑制に苦慮している。

✓ データ提供者と購入契約を結ぶ際のやりとりや、データ提供者の社内での検討等のたたきとして用いることもできるだろう。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(2) IT基準自体に関わる課題

(i) データカタログの拡張

補助事業でデータカタログDTA基準の試用をする中で、各種データの詳細についての表現を、データ詳細部で表現するように、アドバイスを実施した。ただし、現行のデータカタログDTA基準では、データの詳細が表現しにくい部分もあり、今後、データ詳細部の拡張が必要となる。

また、補助事業の中には、解析モデルの販売や、ツールの販売を視野に入れる事業もあり、サービスの表現に対するニーズもあると想定され、今後、データカタログDTA基準をサービスに対応していくことも考えられる。

さらに、ヘルスケア関連のデータは、データ提供者からデータを公開しにくいことから、データ利用者からのオファーにより、データを提供するリクエスト型のモデルを実装している補助事業もあり、データカタログDTA基準をリクエスト型のデータカタログに対応させる必要もあると考える。

- データ詳細部はセンサー情報に関する項目であったため、現状、統合したマスタに関する代表的な項目をそのまま取り込んでいるが、どう表現すればよいか。

- ✓ マスターをそのまま項目として取り入れると、カタログには実際の値を入力することになるため、統合マスターのスキーマの解説をデータ詳細部を書くことが考えられる。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(3) IT基準の活用に向けた課題

(i) データカタログ

① データカタログの構成方法

データカタログDTA基準は、データカタログ本体の中に、データカタログと、データセットという構造を有している。データカタログは、当該構造をベースに、構成していくこととなるが、データカタログやデータセットの単位や、組み方に対する質問が散見され、データカタログの構成方法について説明を実施し

#	課題	課題概要	アドバイス
1	データカタログの構成方法	・データセットはどのように組むのか。	・データセットの商品設計をすることが重要である。その上でユーザーにとって使い勝手のよい単位となるようにカタログと、データセットの階層を仕分けする。
2	データカタログの構成方法	・データを販売する単位を想定した際に、データカタログの単位で設計すればいいのか、それとも、データセットの単位で設計すればいいのか。	・ベースはデータセットでの売り買いだが、グロスで売る場合には、データカタログでの売り買いも想定される
3	データカタログの構造の選択	・データ流通事業者が複数存在する際に、どこを共通化した方がよいか。	・データカタログ本体部をデータ詳細部や、データ利用条件部で補足する形をとっている。データカタログ本体部を考えた上で、データ詳細部をどこまで書くかを考える必要がある。データ詳細部を書くことで利用者がデータにたどり着きやすくなる。全体的に俯瞰したいのであれば、データジャケット部を使うことも考えられる。

② データカタログ項目

既に実装しているデータカタログの項目に関して、問題の有無を確認してほしいという依頼があり、ギャップ分析を行い、差分を共有した。既存のデータカタログが実装され始める中、次期バージョン以降のデータカタログ基準では、既存ビジネスを阻害することのないように、既存のデータカタログを考慮した項目検討が必要となる。

#	課題	課題概要	アドバイス
1	既存のカタログ項目との差異	・24項目のデータカタログ項目を定めているが、問題はないかを確認してほしい。	・DTAのデータカタログ作成ガイドラインとギャップ分析を行い、差分については、その内容を共有する。市場で実装しているものなので、市場の形成を阻害しないという観点で検討し、次期バージョン以降で考慮していく。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(3) IT基準の活用に向けた課題

(i) データカタログ

③ 記入方法

データカタログ基準の項目の理解や、項目の記入方法についての疑問が挙げられており、それらの疑問に対して詳細な説明を実施した。挙げられた疑問の中には、現行のデータカタログDTA基準の項目追加や説明の見直しなどにつながる事項もあり、基準改定の際に考慮していく必要がある。

#	課題	課題概要	アドバイス
1	カタログレコードの用途の理解	・カタログレコードはどのように使うのか	・カタログの更新履歴を記載する
2	カタログIDの記入方法	・カタログIDは何を記載すればよいのか	・将来、事業者間でデータカタログを交換することになった際には、グローバルでユニークな識別子を入れる必要があるが、社内管理用とするのであれば、任意の番号を採番するなどすればよい。
3	データ提供者の記入方法	・データ提供者については複数事業者をすべて記入するのか。	・データ提供者は今後増加が想定されることや、全てを記載すると、長大になることから、別建てで管理することが考えられる。
4	データの分析・加工方法の記入方法	・利用者に提供するものは、映像データではなく、映像分析ツールから出力されたデータを提供する場合、どう書けばいいのか。	・データ詳細部を、カメラと解析エンジンのセットの位置づけとして、データ詳細部書き入れるということが考えられる。 ・現在の項目定義書のデータ詳細部は、「観測活動」に関わる項目が主だが、DTAとして、分析・加工に関わる項目を設けることも考えられる。
5	サンプルデータの記入方法	・データのやり取りを行う際には、サンプルデータを見ないと分からないこともある。そうした情報は、どこかに書けるのだろうか。	・サンプルを書く部分については、課題としては認識しているが、例えば、サンプルデータセットを作って、データカタログ本体部の「配信」にある「ダウンロードURL」という項目から飛ばすということが考えられる。
6	カタログの分類テーマの記入方法	・「カタログの分類テーマ」が分かりづらかった。KOS自体は分かるが、どのキーワードを持ってくればよいか分からなかった。	・今後、統一的な分類の設定などが必要であると認識している。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(3) IT基準の活用に向けた課題

(ii) ドメイン語彙

① 策定方法

IMIやDMD等の概念の理解や、語彙の調整、オーソライズの仕組みなどに対する疑問が出ており、それらの疑問に対して詳細な説明を実施した。類似した分野間の語彙の調整については、まだスキームが不明確であるため、今後検討が必要となる。

また、アドバイス会の中で、デモを行いつつ、ドメイン語彙の策定（コア語彙マッピング等）を行ったが、“自分でドメイン語彙を作ろうとするとかなりハードルが高いと感じた”という意見も挙がっており、手厚い支援が必要となると考えられる。

#	課題	課題概要	アドバイス
1	IMIの概念の理解	IMIを使うとはどういうことなのか知りたい	<ul style="list-style-type: none">・企業や業界が違うが、同様の意味を表しているものに対してコア語彙をマッピングすることで、紐づけが可能になる。コア語彙で関連付けは行うが、ラベルは業界のラベルを使うことも想定される。・コア語彙マッピング、DMD作成、JSONなどの物理形式での出力、ライブラリ（準備中）の活用という段階があり、コア語彙マッピングまで行えばパブリックドラフトとして公開も可能となる。
2	DMDの概念の理解	DMDはどういうものか知りたい	<ul style="list-style-type: none">・DMDはデータモデルをIMIで決められた記法であり、汎用的なデータモデル定義である。DMDを作成し、IMIツールに入力することで、RDFやJSON、XMLなどを出力できるところがメリットとなる。
3	コードリストのルール理解	コードリストの作成に関して守るべきルールはあるか。	<ul style="list-style-type: none">・コードリスト型のプロパティは、名称、発行者、バージョンの3つで識別し、バージョンを管理する必要がある。
4	類似した分野間の語彙の調整	類似した分野間の語彙はどう合わせていくのか。	<ul style="list-style-type: none">・IPAのIMIにおいては制度がまだ固まっていないと聞いている。IPAにパブリックドラフトを出せば、その分野がドメイン語彙を策定していることはわかるだろう。
5	語彙のオーソライズ	作成した語彙はどのようにオーソライズするのか。	<ul style="list-style-type: none">・IMIにはパブリックドラフトの登録制度があるので、IMIパートナーになって登録することで、オーソライズすることとなるだろう。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(3) IT基準の活用に向けた課題

(ii) ドメイン語彙

②参考となる語彙

語彙の策定にあたり、参考となる語彙があれば知りたいといった要望が挙げられており、可能な範囲で、アドバイスを実施した。

#	課題	課題概要	アドバイス
1	参考になる語彙	・他の産業で参考になるような語彙はないか。	・EDI業界は様々な標準化が行われ、観光EDIなども行われているので、その中に活用できるものがあるかもしれない。その他に汎用的な語彙として、IPAのコア語彙があり、コア語彙をドメイン語彙に用いて汎化していくという形もある。
2	多言語展開に向けて参考になる語彙	・多言語展開していくときに参考になるものはあるのか。	・日本語の表現が、その国の表現に意味や値的に一致しているかは、文化や商習慣の違いが存在するといわれおり、翻訳は、大きな検討テーマである。辞書を介してお互いの言葉を紐づけあう方が汎用的な部分もあるが、お互いの言葉を紐づける概念については、統一的なものが必要だろう。

4. 事業者支援

4.2 アドバイス実施

4.2.3 アドバイスの状況

(4) その他課題等

その他に、アドバイス会を行う中で、データ共有にあたってのいくつかの意見が散見されている。以下に、当該意見を示す。

(i) 分野横断でのID/認証の仕組み

安心してデータ共有を行うために、データ提供者とデータ利用者でお互いを特定するためのID/認証の仕組みの整備について意見が挙げられている。

- 今の仕組みは、データ提供者や個人を認証する仕組みがなく、なりすましや改ざんを防止する仕組みが将来的な課題である。どこまで厳密性を求めるかがポイントである。
- 認証等はどこかでディスカッションされているか。認証の仕組み等があると公的なインフラとして使えるようになると思う。

(ii) データ収集（データの品質向上等）に対する国の支援

データ学習等をするには、収集するデータの品質向上が重要であり、データ収集部分に対する国の支援がほしいという意見も挙げられている。

- 機械学習に注目が集まっているが、収集するデータの品質の向上が必要であり、データの収集部分のデータの整備に対して国が支援をしてほしい。

4. 事業者支援

4.2 アドバイス実施

4.2.4 まとめ

アドバイスを希望する8事業に対して、アドバイス会を開催した。アドバイスは、データカタログやドメイン語彙を中心としつつも、その他のデータ共有に関わる課題についても、アドバイスを実施した。以下に、アドバイス会の結果のまとめを示す。

#	区分	課題	アドバイス会結果
1	業界における課題	各所での異なる基準等の採用	・データの差異を吸収する辞書の整備についてアドバイスを実施した。
2		データ共有への抵抗感	・データカタログによりレベル（NDAの前後等）を分けた管理を行いつつ、段階的なデータ提供することで、緩やかにデータ共有をする可能性がある旨をアドバイスを実施した。
3		データのマッシュアップに向けた粒度・ライセンスの統一	・単位等の粒度の統一に向けたスタンダードの採用や、ライセンスの統一に向けたデータ利用条件部の活用についてアドバイスを実施した。
4	IT基準自体に対する課題	データカタログの拡張	・各種データの詳細をデータ詳細部へ表現する旨をアドバイスを実施した。 ・様々なデータが表現できるように、データ詳細部の拡張の必要性に対する示唆を得た。 ・データカタログのサービスへの対応や、リクエスト型データカタログへの対応に対する示唆を得た。
5	IT基準の活用に向けた課題	データカタログの構成方法	・データカタログやデータセットの単位や、データカタログ基準の構造の考え方等についてアドバイスを実施した。
6		データカタログ項目	・既存のデータカタログとのギャップ分析を実施し、差分を提示した。 ・既存のデータカタログとの差分につき、データカタログDTA基準へ考慮が必要である旨、示唆を得た。
7		データカタログの記入方法	・データカタログ基準の項目並びに、記入方法の理解促進に向けての説明を実施した。 ・カタログの分類テーマに対して、統一的な分類設定など検討が必要である旨、示唆を得た。 ・今後の基準改定の際に項目追加や説明の見直しなど、考慮が必要である旨、示唆を得た。
8		ドメイン語彙の策定方法	・IMI等の概念や、策定方法の理解促進に向けて説明を実施した。 ・分野間の語彙の調整については、今後検討が必要である旨、示唆を得た。 ・ドメイン語彙の策定（コア語彙マッピング等）に対する手厚い支援が必要である旨、示唆を得た。
9		参考となる語彙	・参考となる語彙の手がかり等についてアドバイスを実施した。

5. まとめ

5. まとめ

5.1 本事業の成果

本事業における成果を以下に示す。

(1) 補助事業者へのアンケートの実施

事業者支援の実施に向け、補助事業者に対してアンケートを実施し、補助事業者の検討状況や、補助事業者が抱える「データカタログ策定」「語彙」等に関する課題の把握を実施した。本アンケート結果は、勉強会の教材や、アドバイスの実施に役立っている。

(2) 事業者支援教材整備

「データ標準化教材」、「データカタログ策定に関する教材」、「共通語彙に係る教材」の3種類の教材を開発した。

本教材は、補助事業者の支援教材とすることにポイントを置きつつも、初心者でも理解しやすい形で作成しており、「データカタログ策定に関する教材」「共通語彙に係る教材」を用いた勉強会のアンケートでも、一定レベルの評価が得られていることから、データカタログや共通語彙の普及に向けた導入教材として、広く活用することが可能と考えている。

(3) 事業者支援

全補助事業者に対する「データ標準化」説明会、希望する補助事業者への「データカタログ策定」「共通語彙」に関する勉強会、及びアドバイス会を実施した。アドバイス会については、希望する全補助事業者に対して実施した。

これらの事業者支援を通して、データ標準化の重要性や、データカタログ、共通語彙に対する理解を促すとともに、データカタログ、ドメイン語彙の策定を中心としたデータ共有に関わる補助事業者の課題把握、アドバイスを実施することで、課題解消の一助として貢献できたものとする。

また、これらの事業者支援を通じ、現行のIT規格にフィードバックが必要な事項を整理した。当該内容は、「5.2 今後の取組」に示す。当該フィードバック事項をIT規格ブラッシュアップの材料として検討することで、IT規格の更なる利便性・実効性の向上に貢献できるものとする。

