

情報連携用語彙データベース に係るプロジェクトについて

経済産業省が実施した調査事業

2012年度

1. 「オープンデータに関する調査研究（2012年度）」（2012年3月29日）
 - オープンデータのニーズや経済効果やルール整備の調査
 - 経済産業省及び関係機関が保有するデータを対象とした二次利用可能化のための技術や法的課題の検討
2. 「公共データワーキンググループ報告書（2012年度）」（2013年6月）
 - 経済産業省の保有する公的データ（公共データ）を洗い出し、その一部をオープンデータとして整備。
 - データ公開に係る技術的課題、法的課題を検討
3. 空間位置情報に関連する公共データの活用実証事業（2012年度）
 - 自治体等で保有する公共データと空間位置情報との関係性の整理と有効的な組み合わせについて調査し、サービスの促進に必要な課題などを整理。
4. 「公共情報交換標準スキームの整備に関する調査研究（2012年度）」（2013年3月）
 - 行政の情報システムにおける情報連携の効率拡大のため、データの意味レベルでの相互運用性拡大のための手法につき、海外動向（NIEM(米国)、JOIONUP(欧州)）を中心に調査。
 - 情報交換用語彙セットのプロトタイプを整備
 - 情報交換スキームの有るべき姿として「IMI」を提案

⇒ 本「語彙基盤プロジェクトは」上記調査研究で提案されたスキーム「IMI」をベースに、情報共有の基盤の実現を目指すものである。

2013年度（実施中）

1. 「オープンデータの利用に係る民間、地方公共団体の連携基盤の開発・活用実証事業」（2013年度）
 - 地理情報を対象とした情報連携のためのツール（基盤地図情報等のWeb標準（JIS X 7197、JIS X 4197 等）準拠に変換するツール等）を整備し、実証実験を実施。
 - ツールは11月ごろプロトタイプを公開、3月末に正式版をオープンソースにより公開予定。

語彙基盤とは

情報の電子的な公開や交換の際、伝えたい内容が正しく伝わるようにするために、

1. 情報の構造に一貫性を持たせ、
2. そこに用いられる用語の意味を明確にし、
3. 用語の統一化または用語間の関連性の明確化を図る

事を目的とし、用語の意味定義や構想、データ作成時のルール等の基本情報を共有し、活用するための基盤。

構造、意味、表記の共通化

語彙基盤の効果

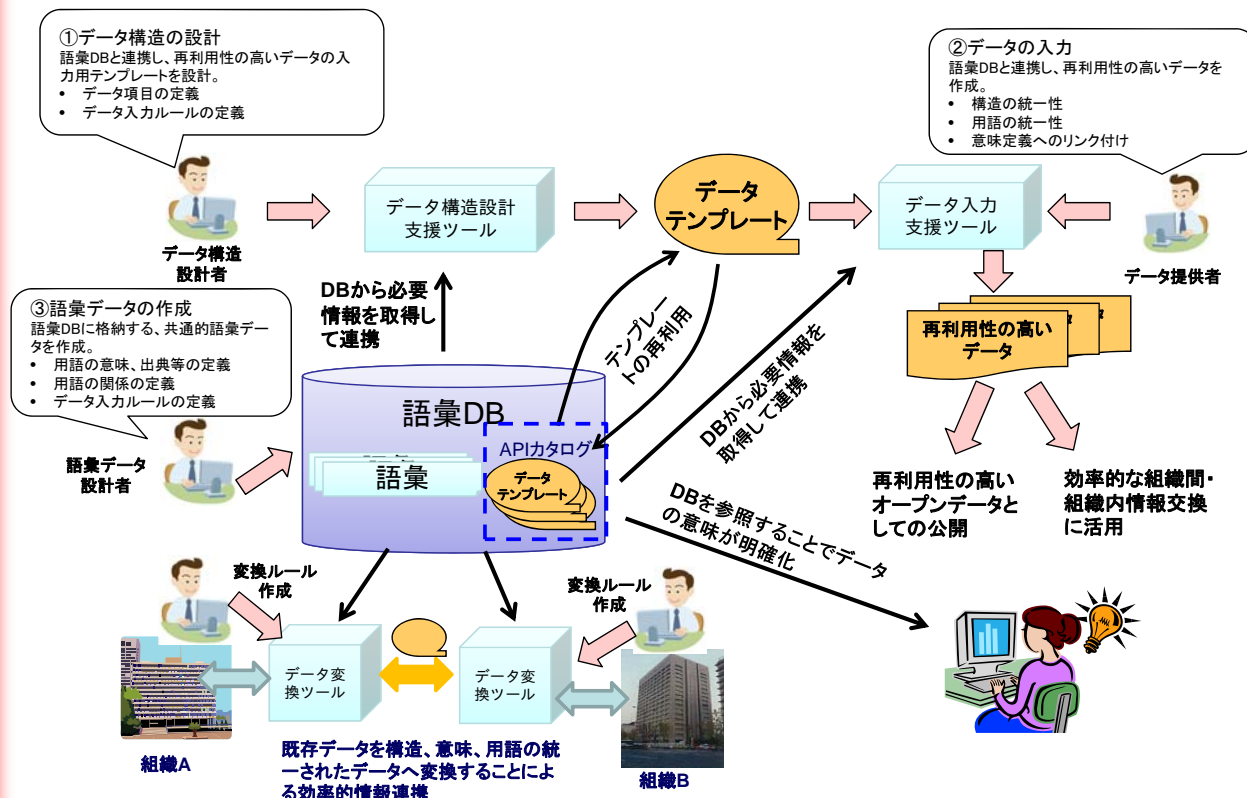
1 次的効果：

1. 行政機関等が、意味が明確で再利用性の高い情報を効率的に作成できる。
2. 行政機関などから公開された情報の意味の解釈が機械的に可能となる。

(再利用性の高い情報が普及することによる) 2 次的効果：

1. 情報交換効率の向上による行政の業務効率とサービス向上
2. オープンデータとして公開された再利用性の高い情報を活用し、価値を引き出す各種産業の発展

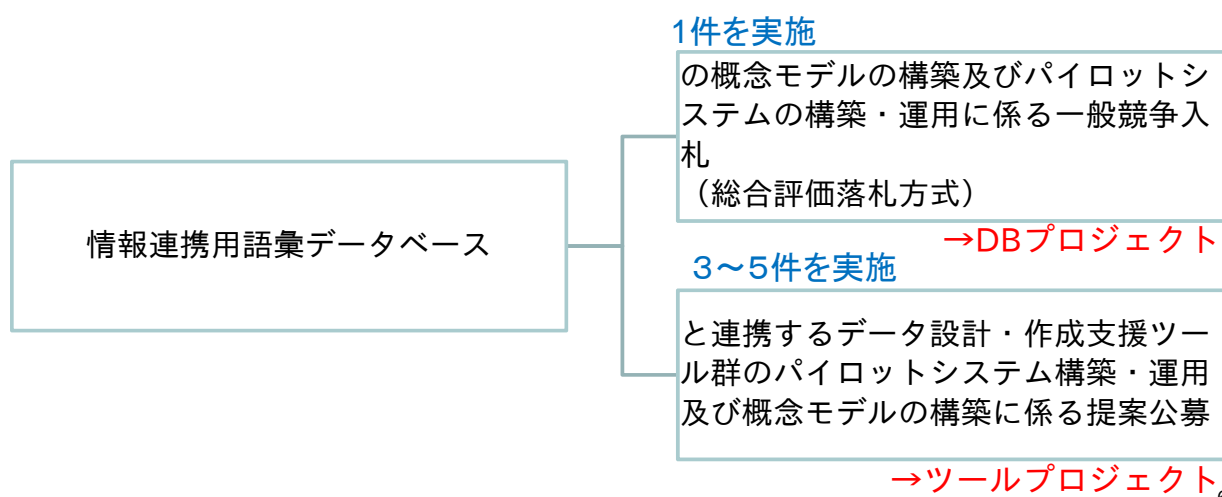
- 「用語」は、ある特定の意味あるいは概念を示すものであり、文字を組み合わせた表記を伴う。また、用語は、その意味あるいは概念の一部として、他の用語の関係（構造）を持つ場合がある。たとえば、「時刻」という用語を「時」、「分」、「秒」という下位の用語から構成すると定義することができる。また、「融資」と「貸し付け」のように、同一の意味を持つ用語が、異なる表記で示される場合がある。
- 用語の集合が「語彙」である。
- 用語の構造や意味、表記の制限、用語と用語との関係等、用語に係る様々な情報を総称してここでは「語彙データ」と呼ぶこととする。なお、本事業で対象とする用語は、原則、行政の現場で業務上必要とされるものとする。



5

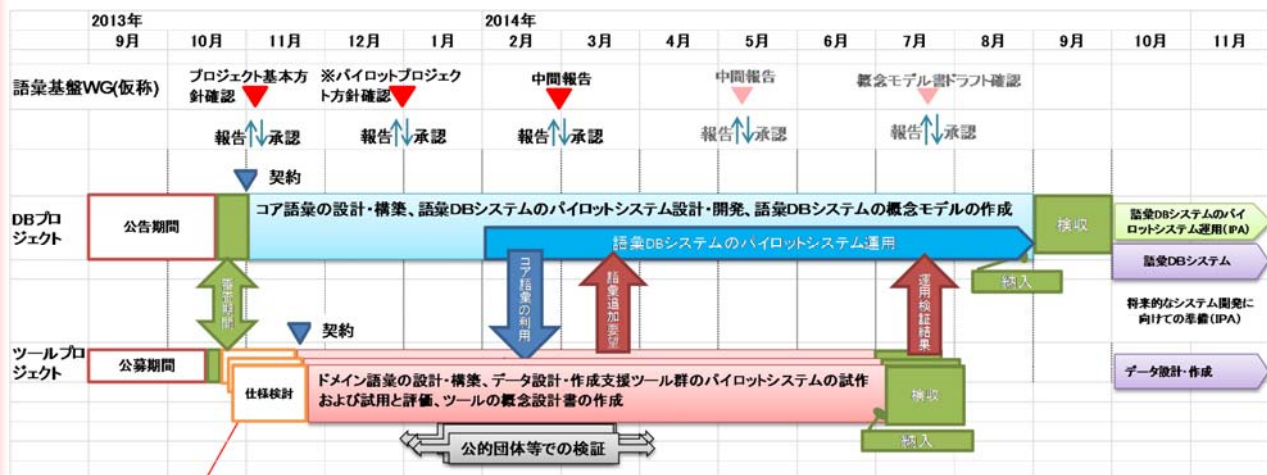
今回のプロジェクト構成

- 語彙基盤を支えるデータベースとツール群について、その要求要件を洗い出した「概念モデルの構築」を行う
- 「概念モデルの構築」にあたっては、パイロットシステムを運用し、実際のデータ作成を試行的に実施し、そこで得られた知見を盛り込む
- プロジェクトは以下のように「DBプロジェクト」と「ツールプロジェクト」の2種類から構成される。



6

「概念モデルの構築」の全体スケジュール(2013年10月頃～2014年9月頃)



DBパイロットシステムに用意されるデータ構造、APIおよびコア語彙の仕様がDBプロジェクト担当者から、ツールプロジェクト担当者に渡される。

公共情報交換標準スキームの整備に関する調査研究(2012年度) 参考資料1に収められた語彙とする

検証用の語彙データは極めて基本的な「コア語彙」以外については、ツールプロジェクト側が用意し、原則DBパイロットシステムへ登録し、運用。

評価はツールプロジェクト実施者が独自に行い、概念モデル中にまとめるとともに、それらをDBプロジェクト実施者がDB運用の視点から取りまとめる。

「概念モデルの構築」を含む、計画全体の概観



DBプロジェクト

情報連携用語彙データベースの概念モデルの構築及びパイロット
システムの構築・運用
に係る一般競争入札（総合評価落札方式）

9

DBの目的

- 情報の「構造、意味、表記の共通化」に必要なとなる基本的データ（語彙データ）を格納し、供給する。
 - － 語彙データの、想定される供給先：
 - データ構造設計支援のためのツール
 - データ入力支援のためのツール
 - データの意味確認のためのツール
 - 情報連携の際等にデータを変換するためのツール（「ツール」には一般的Webブラウザも含む）
- 語彙データのメンテナンスを支援する機能を提供する。
- 利用者が設計した「データテンプレート」の共有を支援・促進する。
- その他、データの「構造、意味、用語の共通化に資する資源(ツール類を含む)」を格納し、供給する。

10

1. DBの要求要件整理

- データ構造などに係る外部専門家を交え、客観的、かつ、国際的動向との整合性を配慮した検討を行う（10名程度による検討会を5回程度実施）
- 以下のドメインを設定し、各ドメイン及び共通ドメインにおけるニーズ評価及び利用シーンの洗い出しを行い、各シーンに基づいた要求要件を洗い出す

(ドメイン)

- 「電子行政オープンデータ推進のためのロードマップ」に掲げられた「オープンデータの重点分野（白書、防災・減災情報、地理空間情報、人の移動に関する情報、予算・決算・調達情報）を中心に対象とする分野を検討する。

(シーン例)

- 語彙データの配信に係る利用シーンとそれに対応した要求要件の洗い出し
 - 語彙データの作成に係る利用シーンとそれに対応した要求要件の洗い出し
 - データテンプレートの共有促進に係る利用シーンとそれに対応した要求要件の洗い出し
 - 省庁等が所有する既存データを、標準的データへ変換することによる情報連携
 - SNS的機能を活用した、語彙やテンプレートへの評価収集、意見交換機能の検討
- など

11

- データ構造の検討

- 「公共情報交換標準スキームの整備に関する調査研究(2012年度)」報告書第3章に示された「公共情報交換標準スキーム(IMI)」で示されたデータ構造をベースとし、さらに、
 - 「用語」の意味の記述法
 - 「用語」の構造の記述法
 - 「用語」の関係の記述法について、汎用性の高い構造を検討する。
- 既に定義された「用語」を使ってさらに別の新たな「用語」を定義できるといった、柔軟性、汎用性のある情報構造を検討する
- 国際標準（デファクトを含む）との整合性の検討（XML、RDF等）
- 諸外国先進事例（NIEM, JOINUP等）との整合性
- 既存データ（UN/CEFACT, Dublin Core等）との相互運用性
- 国等が統計等で用いている各種コードとの連携
- 観光・防災ドメイン等で重要となる地理情報の記述方法及び、地域コード等との連携
- データ入力ルール及びその記述法の検討
- 基本データ型の検討（文字コード／文字セットを含む）

12

- － APIの検討
 - ・ 語彙DBを活用するためのツール群と語彙DBが連携するためのAPIを検討。
- － その他
 - ・ 実運用時におけるセキュリティ要件の洗い出し
 - ・ コアとドメインへの分割など、将来の効率的、持続的運用を考慮した論理構成、物理構成及び管理機能等の要件の洗い出し
- － DB運用体制の検討

将来実現する語彙DBについて、語彙のメンテナンスを含めた運用全体の有り方に係る検討を行う。

 - ・ 語彙データの整備とメンテナンス、その他DBの運用全般を検討
 - ・ 語彙データの作成に係るルールの検討
 - ・ NIEMでとられているような、コアとドメインに分けるといった、検討構造につき、持続性の期待できる方法を検討。

13

2. ツール群に係る要件の洗い出し

- － 語彙データを活用し、情報連携、情報公開の効率化に資するツールの要件を洗い出す。
 - ・ データ構造設計支援のためのツール
 - ・ データ入力支援のためのツール
 - ・ データの意味確認のためのツール
 - ・ 情報連携の際等にデータを変換するためのツール

3. 語彙データの整備

- － 将来の中央省庁及び地方自治体等の情報公開、情報連携の基本となる用語を（3つ以上のドメイン及びそれらに共通するコア語彙につき、構造を持つ用語を合計100用語以上）リストアップ。
- － 以下の対象を調査あるいは参照し、対象データを設定する
 - ・ 公共情報交換標準スキームの整備に関する調査研究報告書(2012年度)
 - ・ NIEM(米国)、JOINUP(欧州)、UN/CEFACT Core Component Library (国連)、地域情報プラットフォーム(APPLIC)、中間標準レイアウト仕様（総務省平成23年度調査）
 - ・ 優先度の高い用語を、根拠を明確にして選択
 - ・ 対象分野関係者を交えた十分な検討を実施（10名程度による検討会を5回程度実施）
- － 検討したデータ構造に従い、上記でリストアップした用語を語彙データとして整備

14

4. DBパイロットシステムの構築と運用

パイロットプロジェクトで構築・運用するDBは、本プロジェクトで最終的に構築を目指す機能のうち、「ツールプロジェクト」を試行するために必要最小の機能を実現するものであり、最終的に実現を目指すDBとは必ずしも同一仕様である必要はない

- 語彙基盤の概念モデルの構築に必要な知見を収集するため、IPA共通基盤システム上にデータベースのパイロットシステムを構築し、運用する。
- 「公共情報交換標準スキームの整備に関する調査研究(2012年度)」報告書第3章に示された「情報交換基盤案」に基づくデータ構造に基づき、同報告書参考資料1に収められた「IMI コアボキャブラリ」の情報を格納し、WebAPIを通じて提供する。
- パイロットシステムで用いるAPI及び語彙データの詳細仕様を、契約後1か月以内に公表する（ツールプロジェクトが参照するため）。
- 「ツールプロジェクト」参加者が希望する語彙データのホスティングと配信を行う。ただし、データ構造は上記IMIコアボキャブラリと同等とする。
- その他、実施者が提案するDBテスト項目（**提案書へ提案**）について運用確認を行う
- 限定した利用者に対するアクセスを許可する
- （利用者限定の条件で）IPAが定めるセキュリティテストに合格すること

15

5. 概念モデルの取りまとめ

- データ構造等重要事項については、少なくとも米国NIEM、欧州ISAのJOINUP、英国data.gov.uk関係者に案を示し、意見聴取を行たうえで概念モデルをとりまとめる。
- とりまとめた概念モデルに基づく語彙基盤の実現により期待される効果につき、定量的、定性的面から根拠を明確にして取りまとめる。

16

ツールプロジェクト

情報連携用語彙データベースと連携するデータ設計・作成支援
ツール群の試作及び試用並びに概念モデルの構築
に係る提案公募事業

17

ツール群の目的

- 語彙データベースと連携し、共通化された構造を持ち、意味が明確化され、標記に統一化の図られたデータの作成を支援する。
 - － データ構造の設計を支援するツール
 - － 設計されたデータ構造へのデータ入力を支援するツール
 - － 既存のデータを、語彙基盤が規定する標準的構造、用語に変換することにより、情報連携の効率化を図るためのツール
- 等がある。

18

- データ入力のテンプレートとなるデータ（データテンプレート）を作成するためのツール。
 - － データ構造を定義する
 - 語彙DBに格納されている「語彙データ」、「データテンプレート」を参照し、適切なものを選択して構造定義に活用する。
 - 語彙DBに既存でない語彙データについては、その場で作成することを支援する。
 - － データ入力ルールを定義する
 - データの型、値の制限などを定義する。

19

- データテンプレートに基づき、指定された場所に指定された形式のデータを入力するようガイドするツール。併せて以下の機能を検討する。
 - － データの意味を明確化するためのリンク等作成支援
 - 語彙データベースが提供する語彙データへのリンクの他、地理座標情報の付与、コード情報、法令情報等の公的データベース等へのリンク作成による意味明確化等も含めた検討を行う。
 - － 標記の統一を図るための用語選択支援

20

- 省庁、自治体などが既に有しているデータを、標準的構造、用語に変換することにより、情報連携の効率化を図るツール。
 - － データ変換ルールの設定
 - 既存データの構造やそこで用いられる用語について、語彙データベースで提供される標準的な構造、意味等に変換するための変換規則を設定するためのツール
 - － データ変換ツール
 - 上記で設定されたルールに基づき、既存データと、標準形式のデータとの間でのデータ変換を行うためのツール

21

1. 対象者

- － 本パイロットプロジェクトの実施主体は、地方自治体等、公的な業務を行う団体（以後公的団体と呼ぶ）のシステム構築に関わる企業等とする。
- － パイロットシステムの構築、運用に際しては当該公的団体の職員等が構築への助言、および運用結果への評価に関与するものとする。

2. 対象業務（以下のいずれかまたは両方とする）

- － 再利用性の高いオープンデータの作成と公開
- － 業務連携のための再利用性の高いデータの作成とその交換

22

3. 業務内容

① ツールの要件整理

- 利用シーンの設定と、そのシーンに基づいたツールの要件の洗い出し
(例)
 - データ構造設計支援
 - データ入力ルール設計支援
 - データ意味定義とのリンク作成支援

② 語彙データの構築

- パイロットプロジェクトの実施に必要な語彙データを構築する。
- 語彙データ構築に当たっては「DBプロジェクト」が策定した要件に沿うものとする。
- 語彙データは、「DBプロジェクト」が運用するデータベースパイロットシステムに格納し、そこから参照することを原則とする。

③ 運用実験

- 上記①で要件整理したツールを実際に作成して運用し、データテンプレートの作成、データの入力に活用する。
- 地理データに係るツールについては、「平成25年度電子経済産業省構築事業（オープンデータの利用に係る民間、地方公共団体の連携基盤の開発・活用実証事業）」で実施されるプロジェクトで作成されたものが公開される予定なので、その活用、データとしては、総務省統計局の国勢調査データ（IDとポリゴン）、国土交通省位置参照コード（IDと緯度経度座標）等の活用の検討が推奨される。
- 実験で作成したデータは納品対象としない。データの作成は実験に必要最小限の範囲とし、それを超えたものは実施者の負担によるものとする。
- 作成した「語彙データ」は、納品の対象とする。ただし、実施者がもともと所有していたデータあるいは外部参照したデータについては、その限りでない。その場合データは納品対象とはならないが、当該データを本プロジェクトで活用するために変換するツールなどを本プロジェクトで作成した場合は、そのツールは納品対象とし、IPAからオープンなライセンスで公開する。
- ツールの運用を通し、ツールおよび、DB、語彙データ等の有効性、問題点を明らかにする。

④ 概念モデルの取りまとめ

- ツールを運用し、実際にデータテンプレート及びデータを作成し、作成効率、作成されたデータの質などに付き、評価を行う。
- ツールの運用実験結果を受け、ツール群及びそれが必要とするデータに係る概念モデルを作成する。

1	世界最先端IT国家創造宣言	http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20130614/siryou1.pdf
2	世界最先端 IT 国家創造宣言 工程表	http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20130614/siryou4.pdf
3	電子行政オープンデータ戦略	http://www.kantei.go.jp/jp/singi/it2/pdf/120704_siryou2.pdf
4	電子行政オープンデータ推進のためのロードマップ	http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20130614/siryou3.pdf
5	オープンデータに関する調査研究	http://datameti.go.jp/data/dataset/report-001-2012
6	公共情報交換標準スキームの整備に関する調査研究	http://datameti.go.jp/data/dataset/report-002-2012
7	IT融合フォーラム公共データワーキンググループ報告書	http://datameti.go.jp/data/dataset/report-003-2012
8	空間位置情報に関連する公共データの活用実証事業	http://datameti.go.jp/data/dataset/report-004-2012
9	平成22年度版「地域情報プラットフォーム標準仕様」	http://www.applc.or.jp/2012/tech/APPLIC-0002-2011/index.html
10	自治体業務アプリケーションユニット標準仕様V2.4	http://www.applc.or.jp/2012/tech/gyomu.html
11	総務省統計局の国勢調査データ	http://e-stat.go.jp/SG2/eStatGIS/page/download.html
12	国土交通省位置参照コード	http://nlftp.mlit.go.jp/isj/data.html
13	総務省 中間標準レイアウト仕様	http://www.soumu.go.jp/main_sosiki/jichi_gyousei/c-gyousei/lg-cloud/02kiban07_03000024.html
14	IPA 文字情報基盤整備事業	http://mojikiban.ipa.go.jp/
15	米国オープンデータの動向調査報告書	http://www.ipa.go.jp/about/research/20130830.html
16	NIEM	https://www.niem.gov/Pages/default.aspx
17	ISA Joinup	http://joinup.ec.europa.eu/
18	Schema Subset Generation Tool(SSGT)	http://tools.niem.gov/niemtools/ssgt/index.iepd
19	Data Model Browser	http://tools.niem.gov/niemtools/viewer/DataModelViewer.iepd
20	Work With IEPDs	http://tools.niem.gov/niemtools/iepd/index.iepd
21	Migration Assistance	http://tools.niem.gov/niemtools/migration/index.iepd
22	Conformance Testing Assistant(ConTesA)	http://tools.niem.gov/contesaNIEM/
23	Code List Generator	http://tools.niem.gov/niemtools/codelist/index.iepd
24	IEPD Clearinghouse	https://it.ojp.gov/framesets/iepd-clearinghouse-noClose.htm
25	Data.gov.uk	http://www.data.gov.uk/