

国内の匿名加工情報を一覧する利活用ポータルサイト「匿名加工情報目録」の開発

梶間 大地 †

明治大学総合数理学部 先端メディアサイエンス学科 菊池研究室 †

1 はじめに

自社の有する顧客情報などの個人情報を匿名加工してビッグデータビジネスに活用する動きが盛んである [1]。その一方、取扱事業者が匿名加工情報を作成する際には届け出は不要で、提供する情報に含まれる個人に関する情報の項目の公表が義務付けられているに過ぎないため、利活用者が望む匿名加工情報を探し求めることが困難であった。そこで藤田らは、匿名加工情報のカタログを開発している [2]。金子らは、匿名加工情報を公表している事業者を収集するクローラの開発を行った [3]。しかし、公表形式がまちまちであり、業種ごとの取得項目や提供方法のトレンドを知ることができなかった。

そこで、本研究では、ウェブで公表されている匿名加工情報のページを収集してデータベースに格納して、検索などのサービスを実現したポータルサイトを開発をする。前述の問題点に対して、SQL のあいまい検索などを工夫した点に新規性がある。本稿では、開発サイトの設計と実現した機能について報告する。

2 ポータルサイトの開発

2.1 概要

図 1 にシステム構成図、表 1 にポータルサイトに実装されている機能を示す。

ポータルサイトの開発には PHP ver 5.3.3 と MySQL ver 5.1.73 を用いた。ポータルサイトは情報一覧ページ、情報を追加するページ、情報を集計した統計情報ページの 3 つで構成されている。

2.2 匿名加工情報一覧ページ

図 2 に開発したポータルサイト*の情報一覧ページを示す。ここでは、匿名加工情報公表ページへのリンク、事業者の業種と詳細ページへのリンクを表示している。

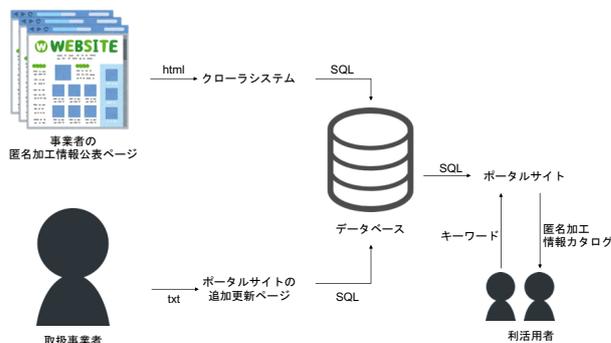


図 1 システム構成図

表 1 ポータルサイトの機能

機能	概要
情報一覧	登録されている企業名、業種、公表されている URL を示す。 詳細から提供項目、提供方法を確認する。
追加	新たに情報を追加する。
検索	企業名、業種、提供項目、提供方法を選択して登録されている情報に検索をかける。
編集	既に登録されている情報を編集する

詳細ページから匿名加工情報の項目と提供方法を確認することができる。

2.3 情報追加更新ページ

情報を追加するページでは、誰でも匿名加工情報を提供している企業の情報を追加することができる。ポータルサイトの利用者の登録は不要とし、登録する際には登録データに対してパスワードを設定している。また、エージェントによる「荒らし」等への対策するため、Google 社の提供する reCAPTCHA[4] を導入している。

2.4 ユースケース「統計情報の集計」

統計情報ページでは、任意のキーワードを与えてデータベースに格納されている業種ごとの事業者一覧と件数を提供する。検索機能を用いてポータルサイト内に格納

†Department of Frontier Media Science, School of Interdisciplinary Mathematical Science, Meiji University, Kikuchi Laboratory.

*<https://windy.mind.meiji.ac.jp/~kajima/2020>

されている取得項目と提供方法を集計した結果を表2に示す。業種分類は、日本標準産業分類を基にした独自分類[3]を採用した。登録されている情報は[3]で収集されたものである。

2.5 考察

表2より、業種ごとに取得項目や提供方法の偏りがあることから、匿名加工情報公表には業種ごとに共通のテンプレートがあると考えられる。

表2 業種別の全てのデータ数と特定のキーワードを含む数

業種	数†	取得項目		提供方法				
		氏名	健診	メール	サーバ	外部	セキュ 媒体	セキュ リティ
医療業(病院)	89(6)	1	0	0	64	62	3	
小売業	11(5)	1	0	2	3	2	1	
小売業(薬局)	56	25	0	2	4	1	47	
複合サービス業	50(43)	1	0	1	1	0	0	
情報サービス業	31(8)	2	3	5	6	3	5	
健康保険組合	19(1)	0	12	0	0	1	3	
保険業	15(9)	0	2	2	0	4	1	
金融業	15(5)	0	0	0	2	0	3	
情報通信業	9(6)	0	0	1	1	1	0	
サービス業	9(4)	0	1	1	1	1	1	
医療業(製薬)	7(1)	0	1	2	4	1	2	
製造業	6(2)	0	1	0	2	1	1	
運輸業	5(4)	0	0	0	1	1	0	
年金保険センター	3	0	0	0	1	1	0	
電気・ガス・ 熱供給・水道業	3(3)	0	0	0	0	0	0	
不動産業	2	0	0	0	1	1	0	
健康保険協会	2	0	1	0	0	0	0	
教育, 学習支援業	2	0	0	0	2	1	1	
映像・音声・ 文字情報制作業	2	0	0	2	1	0	0	
化学工業	2	0	0	0	0	0	0	
社会保険, 社会 福祉, 介護事業	5(1)	0	0	0	2	1	0	
一般社団法人	1	0	0	0	0	0	1	
未分類	18(8)	0	2	1	2	2	4	
合計	378	28	23	19	98	84	73	

†内は取得項目, 提供方法のどちらも欠損しているデータ数

3 おわりに

本研究では、国内の匿名加工情報取扱事業者の公表情報を収集して、検索などの機能を加えたポータルサイトを開発した。データの追加, 更新等を適時行うことを今後の課題とする。



図2 匿名加工情報一覧ページ

参考文献

- [1] 国内ビッグデータ活用事例 - 総務省 (<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc254330.html>, 2020年12月参照)
- [2] 藤田真浩, 飯田泰興, 服部充洋, 山中忠和, 松田規, 伊藤聡志, 菊池浩明, “匿名加工情報取扱事業者による公表情報を利用した匿名加工カタログの提案と実装”, コンピュータセキュリティシンポジウム, pp.1214-1221, 2020.
- [3] 金子侑紀, 小野敦樹, 伊藤聡志, 菊池浩明, 服部充洋, 飯田泰興, 藤田真浩, 山中忠和, “匿名加工情報取扱事業者を調査するクローラーシステムの開発”, 情報処理学会第82回全国大会, pp.3.447-3.448, 2019.
- [4] reCAPTCHA(<https://www.google.com/recaptcha/about/>, 2020年12月参照).
- [5] 匿名加工情報 - 個人情報保護委員会 (<https://www.ppc.go.jp/personalinfo/tokumeikakouInfo/>, 2020年12月参照).