

LODのリンクに起因する Covert Channelの解析

木下研究室 横山鳳我

研究背景

近年では、米国や英国をはじめとした世界各国で標準化されたデータ形式であるLinked Open Data(LOD)に基づいたデータ公開が行われている。LODはデータのオープン化、分野内でのデータ共有、そして分野を横断したデータの共有を促進するという特徴を持っている。



そこでLODのリンクに起因するCovert Channelにより情報漏洩が起きるのではないかと考えた。

問題点

コミュニティでLODを扱う時にコミュニティ内の機密情報がCovert Channelにより漏れてしまう可能性がある。コミュニティでのLODと公共のLODのリンクによってどのような情報漏洩が起きるのかを考慮する必要がある。

目的とオリジナリティー

目的

コミュニティでのLODと公共でのLODのリンクにより Covert Channelがどの様に起きるのかを分析する。

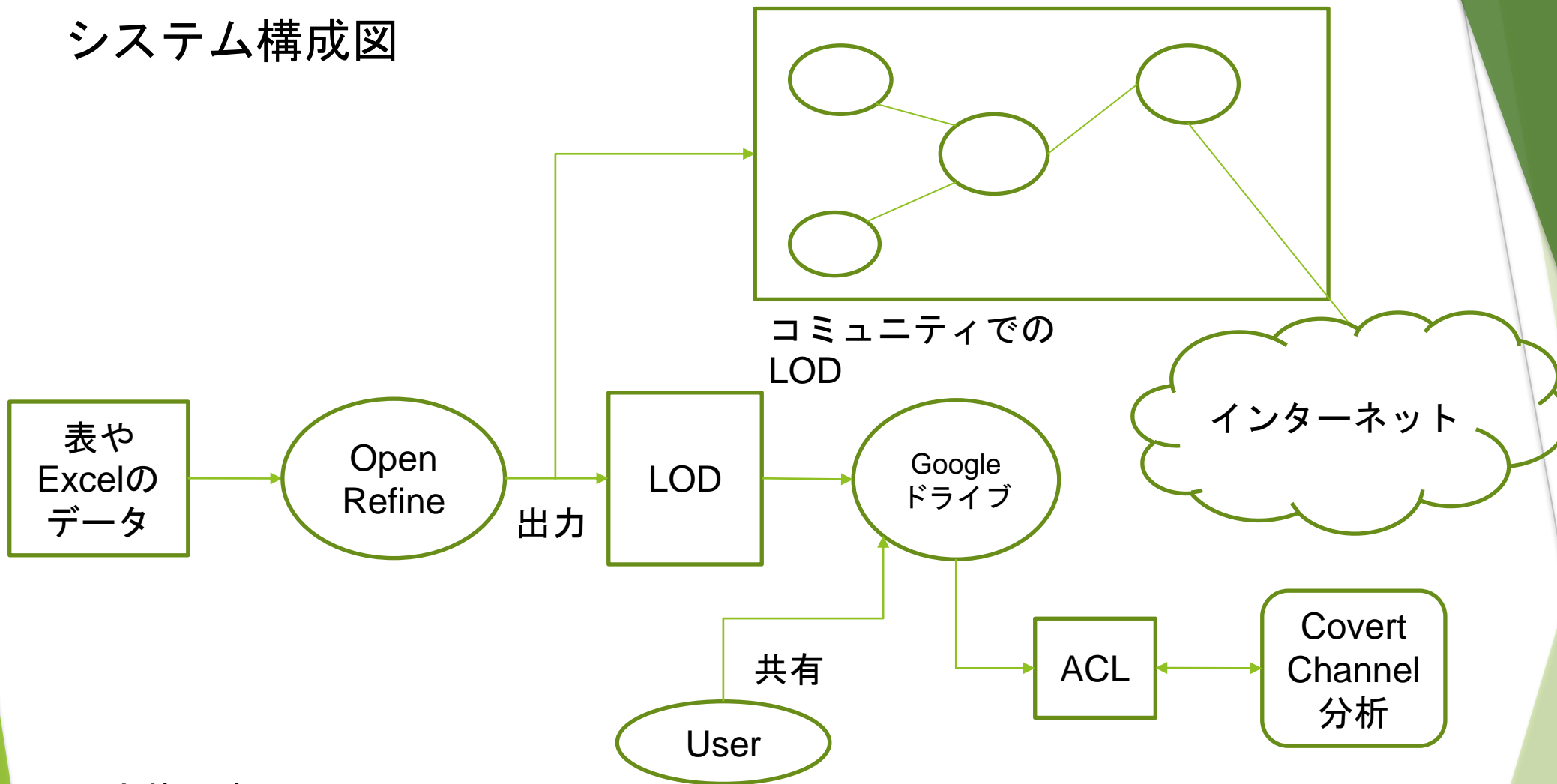
オリジナリティー

OpenRefine により LOD を実装し Google ドライブ 環境 下で LOD を利用するコミュニティでの Covert Channel 分析を提案する。

提案手法

コミュニティでのLODと公共のLODを実装するためにOpenReneの機能を使って、実際にのLODを実装し、実装する過程でCovert Channelの分析をGoogleドライブの環境下で行う。

システム構成図



実装環境

Ubuntu16.04

OpenRefine 2.8

Googleドライブ

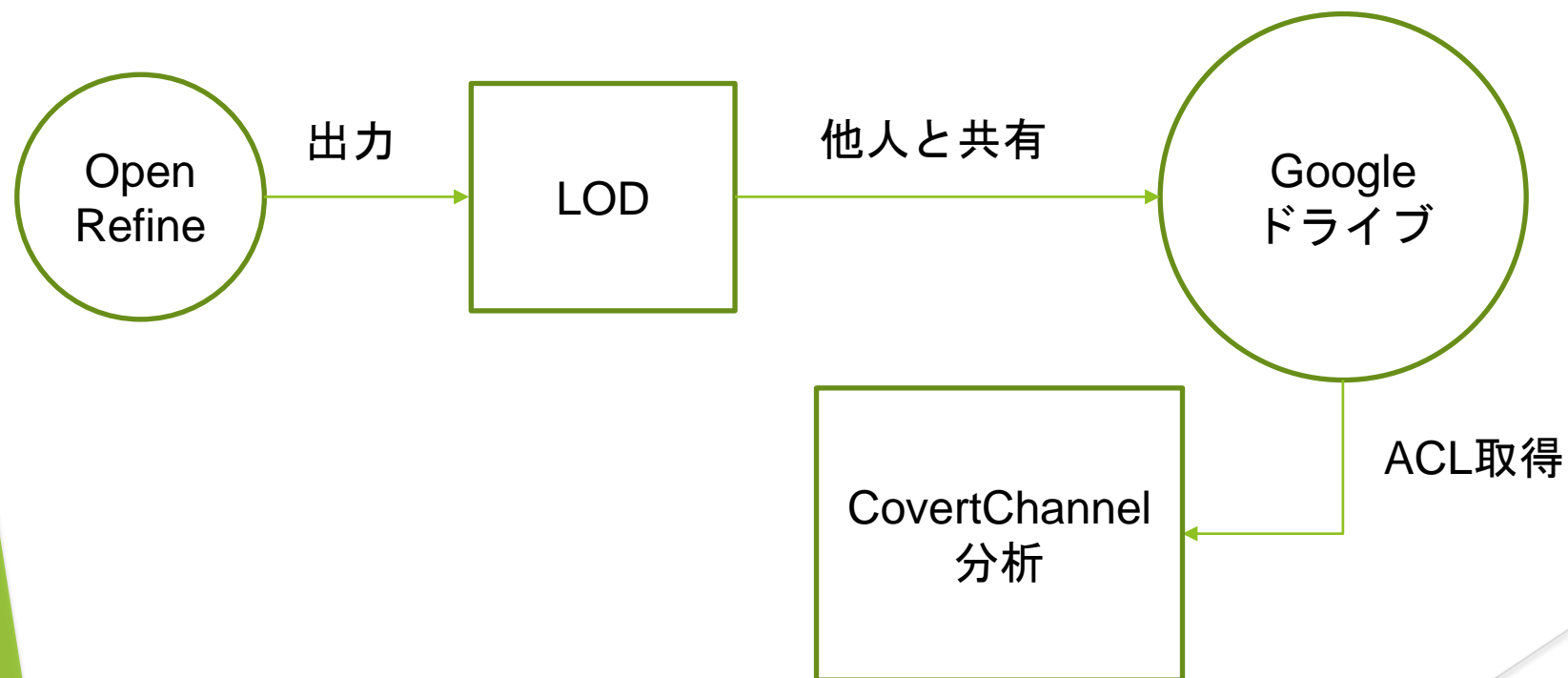
LODとは？

Linked Open Data(リンクト・オープン・データ)の略で、LOD(エルオーディー)と読み、ウェブ上でデータを公開して共有するための方法。
主語(Subject),述語(Predicate),目的語(Object)の組み合わせでリンクしている。



Google ドライブ

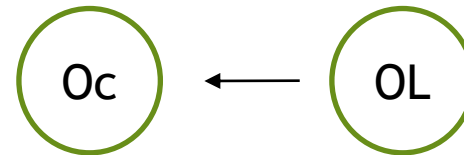
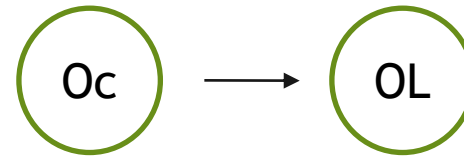
Googleが提供する、オンラインストレージサービス。様々なファイル作成、アップロード、編集、それらの共有、共同編集ができる。Googleドライブを用いてACLを取得し、Covert Channel分析を行う。



Covert Channel

Oc(コミュニティ Object) Sc(コミュニティ Subject)
OL(公共のLOD Object) SL(公共のLOD Subject)

	Sc	SL
Oc	R ↓	Φ
OL	W →	R



アクセス行列からScがOcをreadし,その情報をOLにwriteしてしまふと,そのOLをSLがreadすることによってOcに記述されていた情報がSLに漏えいしてしまふ

OpenRefineと名寄せについて

OpenRefineはデータ整形ツールで様々な機能がある。その中の機能に名寄せというものがある。

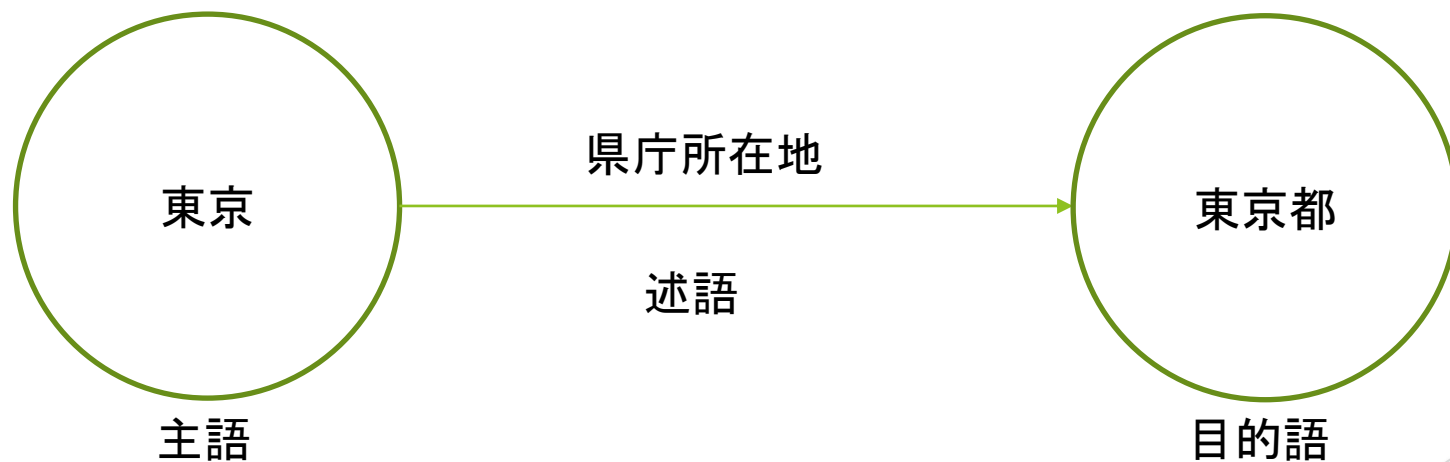
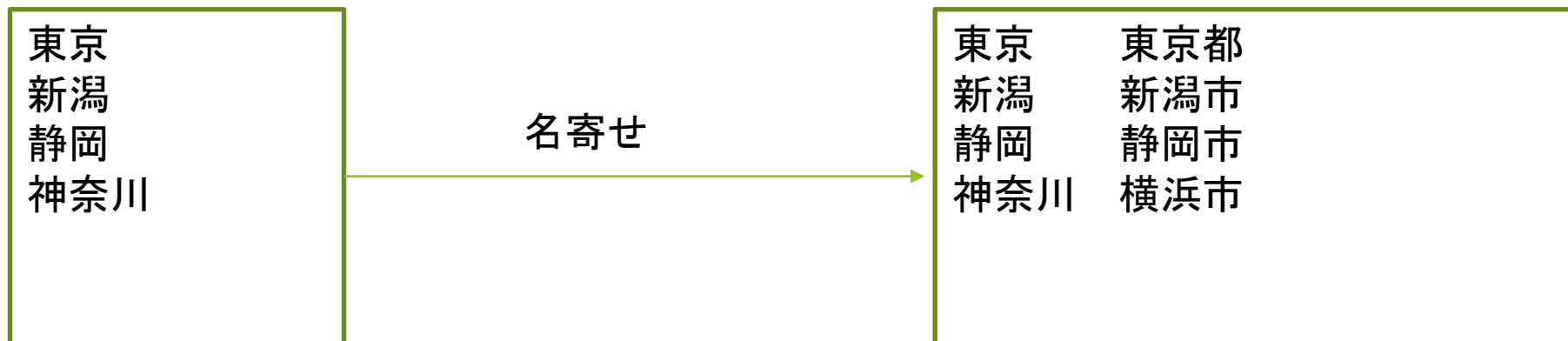
名寄せ(Reconciliation Service API)の仕様に従うウェブサービスを提供しているデータベースに対して照合できる。

名寄せは、何かの名前またはラベルであるテキスト、及びオプションで追加の詳細情報が与えられた場合に、基準に一致する潜在的なエンティティのランク付けされたリストを返すWebサービス。

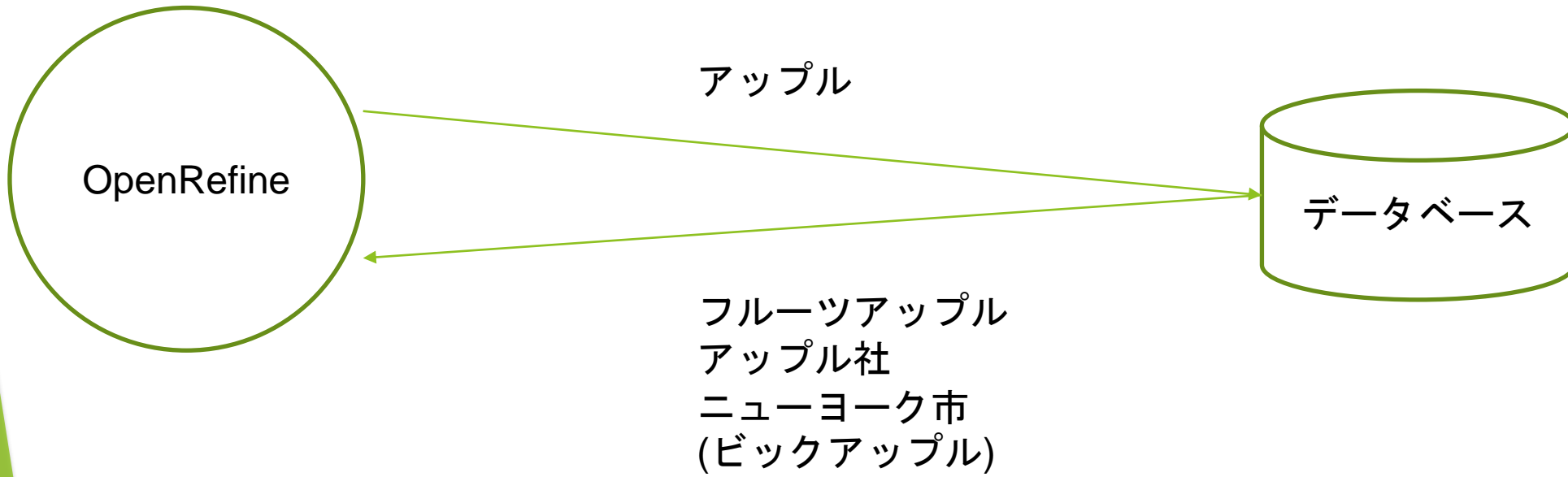
名寄せについて

表や
エクセルデータ

URL追加できる

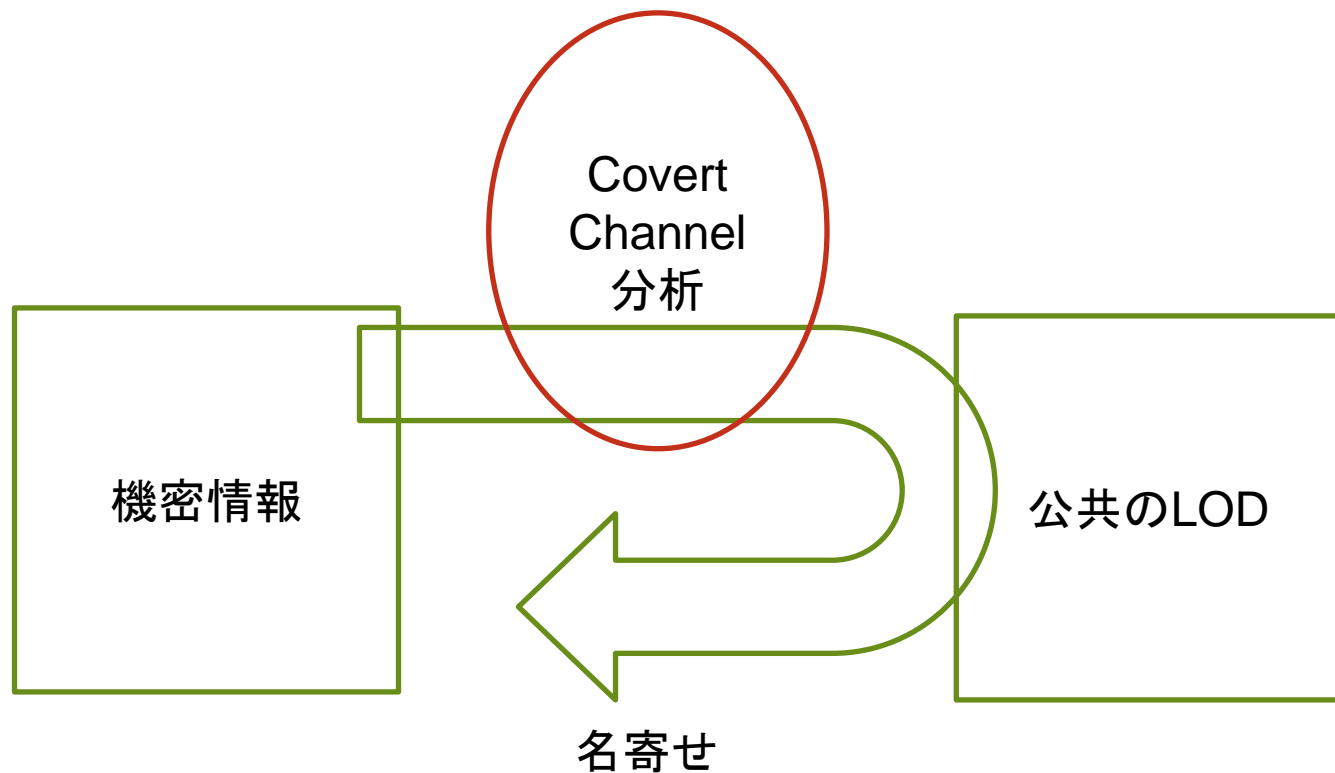


名寄せの例



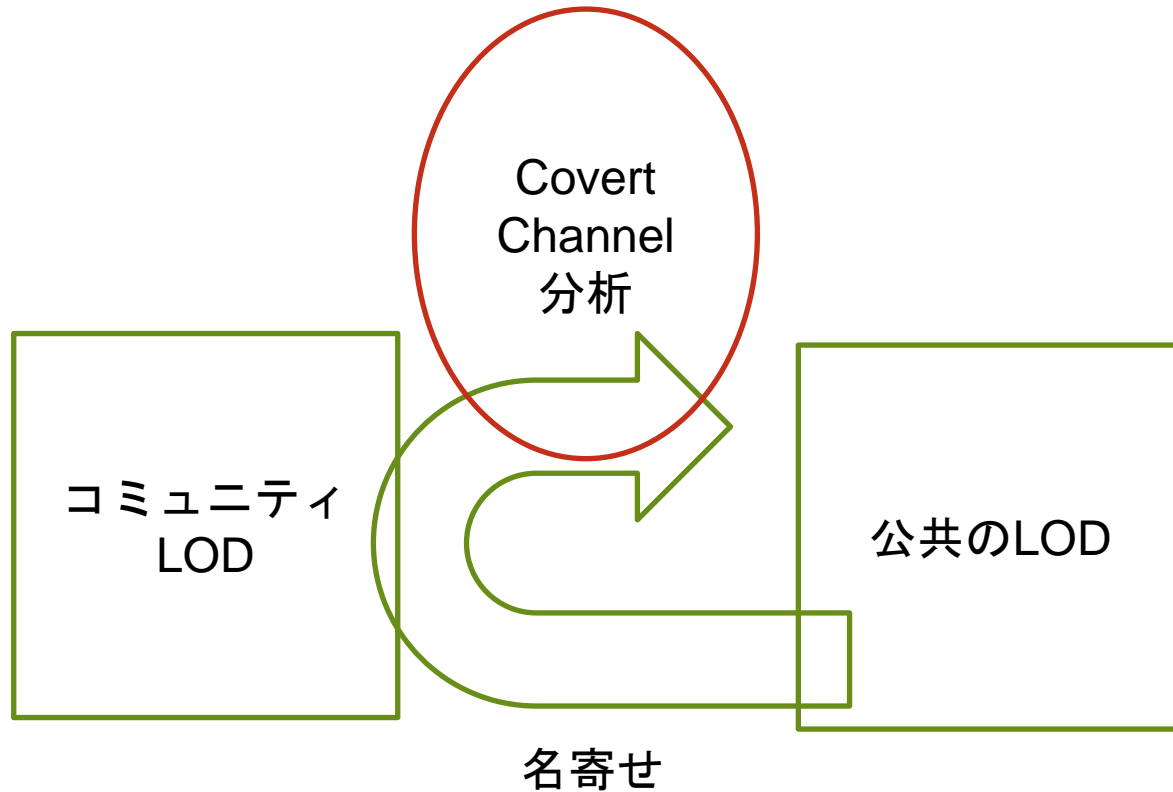
Covert Channel分析

コミュニティでLODを利用したい場合
機密情報に対して名寄せを行う場合にCovert Channelが起きる
可能性がある。



Covert Channel分析

コミュニティでのLODをOpen(おおよけにする)場合
コミュニティでのLODが公共のLODに対して名寄せが行われる場
合にCovert Channelが起きる可能性がある。



名寄せを行う工程

1. OpenRefineの言語を日本語にセットしている。
2. 名寄せを行う。
3. OpenRefineをそのまま使うとマッチが不十分(リンク先が多くなること)があり、人手を必要とする。
4. OpenRefineで照合を行うには、列名の左側にあるアイコンをクリックして表示されるドロップダウンメニューから、「名寄せ」→「名寄せ開始...」の順に選択する。
5. 対象になる列の一部のセルに対してだけ照合させたい場合は予めテキストフィルターやファセットを利用して絞っておく。
6. 続いて「名寄せ」ダイアログボックスの「サービス」エリアに「Wikidata Reconciliation for OpenRefine (ja)」が表示されていればそれを選択する。(より信頼性のある照合結果が得られるように、照合処理をするためのオプションが幾つか用されている)
7. 名寄せを開始する。
8. **Covert Channel分析**
9. 問題なくマッチした場合は、紺色のリンクが一つだけ表示される。この場合は処理結果の正確性が高く、人手で改めて確認作業をする必要はない。

OpenRefineで実装したLOD

19304 レコード

表示: 行 レコード 表示: 5 10 25 50 レコード << 最初 < 前へ 1 - 50 次へ > 最後 >>

	All	NISTEP機関ID	レコード番号	日本語正式名	公式ウェブサイト	位置する行政区画	Facebook Place	面積	学生数	設立	創設者	ロゴ画像	ISNI	英語正式名	セクター番号	セクター分類
☆	🔗	1.	NID201200199399167	01-0001	お茶の水女子大学 新しいマッチを選択	http://www.ocha.ac.jp	東京都 新しいマッチを選択	2187.66	2075	1949-01-01T09:00:00Z		Ochanomizu University (south gate).jpg	000000012192178X	Ochanomizu University	1	国立大学
☆	🔗	2.	NID201200434617623	01-0001-001	お茶の水女子大学生活環境研究センター new 新しいマッチを選択									Institute of Environmental Science for Human Life, Ochanomizu University	1	国立大学
☆	🔗	3.	NID201200072764185	01-0002	愛知教育大学 新しいマッチを選択	http://www.aichi-edu.ac.jp/	刈谷市 新しいマッチを選択	50390000	3844	1943-01-01T09:00:00Z		Aichi University of Education main gate ac.JPG	0000000121114080	Aichi University of Education	1	国立大学
☆	🔗	4.	NID201200808829421	01-0003	愛媛大学 新しいマッチを選択	http://www.ehime-u.ac.jp/	松山市 新しいマッチを選択	103134126393509	428.86	1949-01-01T09:00:00Z			0000000110113808	Ehime University	1	国立大学
☆	🔗	5.	NID201200488738503	01-0003-001	愛媛大学医学部附属病院 new 新しいマッチを選択									Ehime University Hospital, Ehime University	1	国立大学
☆	🔗	6.	NID201200016556519	01-0003-002	愛媛大学沿岸環境科学研究センター new 新しいマッチを選択									Center for Marine Environmental Studies, Ehime University	1	国立大学
☆	🔗	7.	NID201200139365669	01-0003-003	愛媛大学地球深部ダイナミクス研究センター new 新しいマッチを選択									Geodynamics Research Center, Ehime University	1	国立大学
☆	🔗	8.	NID201200540848841	01-0004	旭川医科大学 新しいマッチを選択	http://www.asahikawa-med.ac.jp/	旭川市 新しいマッチを選択	747.6	978	1973-01-01T09:00:00Z			0000000086382724	Asahikawa Medical University	1	国立大学
☆	🔗	9.	NID201200296397387	01-0004-001	旭川医科大学 新しいマッチを選択	http://www.asahikawa-med.ac.jp/	旭川市 新しいマッチを選択	747.6	978	1973-01-01T09:00:00Z			0000000086382724	Asahikawa Medical University Hospital, Asahikawa Medical University	1	国立大学
☆	🔗	10.	NID201200248230458	01-0005	一橋大学 新しいマッチを選択	http://hit-u.ac.jp/	東京都 新しいマッチを選択	2187.66	6200	1875-01-01T09:00:00Z	森有礼 新しいマッチを選択		0000000123479884	Hitotsubashi University	1	国立大学
☆	🔗	11.	NID201200829790881	01-0005-001	一橋大学経済研究所 ☑️ 一橋大学経済研究所 (50) ☑️ 新しいトピックを作成 マッチを検索									Institute of Economic Research, Hitotsubashi University	1	国立大学
☆	🔗	12.	NID201200341506245	01-0006	茨城大学 新しいマッチを選択	http://www.ibaraki.ac.jp/	水戸市 新しいマッチを選択	217320000		1920-01-01T09:00:00Z		Ibaraki Univ-Ami-MainBldg01.jpg	0000000099490476	Ibaraki University	1	国立大学
☆	🔗	13.	NID201200894539974	01-0007	宇都宮大学 新しいマッチを選択	http://www.utsunomiya-u.ac.jp/	宇都宮市 新しいマッチを選択	416.84		1949-01-01T09:00:00Z		Udai-seikyuu.jpg	0000000107224435	Utsunomiya University	1	国立大学
☆	🔗	14.	NID201200942671133	01-0008	横浜国立大学 new	http://www.ynu.ac.jp/	横浜市	105589566142652	437380000	1949-01-01T09:00:00Z		YNU	0000	Yokohama	1	国立大学

facetとfilterを使う

データのサブセットを選択してファセットやフィルターを使えます。それぞれのデータカラムのメニューからファセットとフィルターメソッドを選択してください。

ご不明な場合は
取説動画を御覧ください

拡大図

表示: 5 10 25 50 行

レコード番号	日本語正式名	位置する行政区画	学生数	業種	公式ウェブサイト	国
01-0001	お茶の水女子大学 新しいマッチを選択	東京都 新しいマッチを選択	2075		http://www.ocha.ac.jp	日本 新しい マッチを 選択
01-0001-00 edit	お茶の水女子大学生生活環境研究センター <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> 新しいトピックを作成 マッチを検索					
01-0002	愛知教育大学 新しいマッチを選択	刈谷市 新しいマッチを選択	3844		http://www.aichi-edu.ac.jp/	日本 新しい マッチを 選択
01-0003	愛媛大学 新しいマッチを選択	松山市 新しいマッチを選択			http://www.ehime-u.ac.jp/	日本 新しい マッチを 選択
01-0003-001	愛媛大学医学部附属病院 <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> 新しいトピックを作成 マッチを検索					
01-0003-002	愛媛大学沿岸環境科学研究センター <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> 新しいトピックを作成 マッチを検索					
01-0003-003	愛媛大学地球深部ダイナミクス研究センター <input checked="" type="checkbox"/> <input checked="" type="checkbox"/> 新しいトピックを作成 マッチを検索					
01-0004	旭川医科大学 新しいマッチを選択	旭川市 新しいマッチを選択	978		http://www.asahikawa-med.ac.jp/	日本 新しい マッチを 選択
01-0004-001	旭川医科大学 新しいマッチを選択	旭川市 新しいマッチを選択	978		http://www.asahikawa-med.ac.jp/	日本 新しい マッチを 選択
01-0005	一橋大学 新しいマッチを選択	東京都 新しいマッチを選択	6200		http://hit-u.ac.jp/	日本 新しい マッチを 選択

結論

本研究では、OpenRefineの機能を用いてLODを実装し、Googleドライブによって他人と共有するところまで実装した。OpenRefineの名寄せ(Reconciliation Service API)アルゴリズムの中にコミュニティでのLODと公共のLODの間のCovert Channelを分析し、アクセスを制御するアルゴリズムを考察した。

今後の課題としては名寄せアルゴリズムの中にCovert Channel分析ができるようなプログラムを考える必要がある。