

久野G研究紹介

東京大学大学院 情報理工学系研究科
数理・情報教育研究センター / 数理情報学専攻(兼担)
2024年



創発的研究支援事業
Fusion Oriented REsearch for disruptive Science and Technology



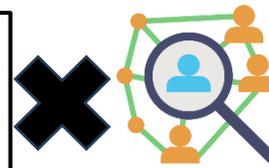
UTokyo
FSI
The University of Tokyo
Future Society Initiative



全体



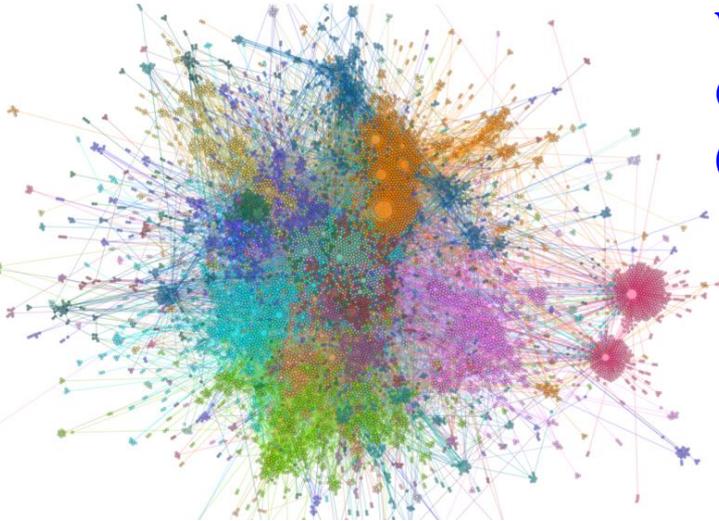
- ネットワーク
- セマンティックウェブ



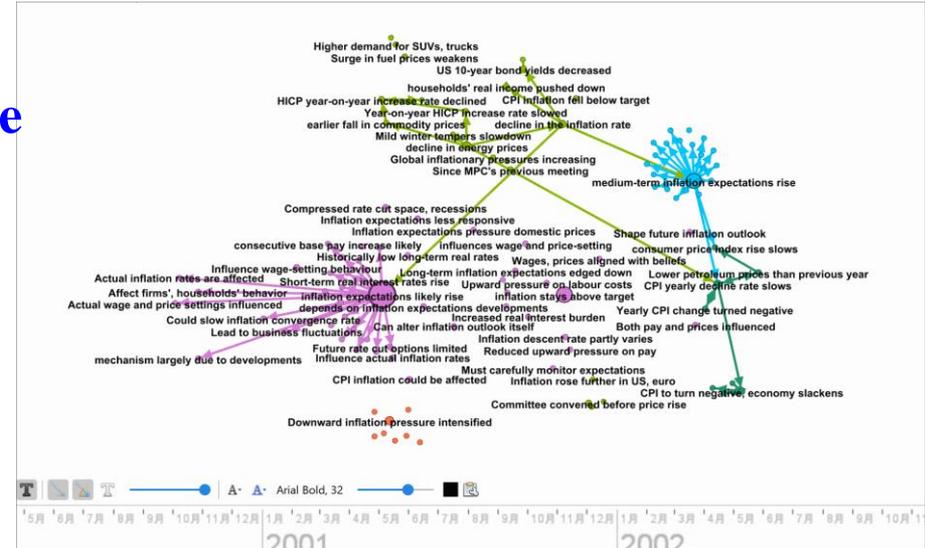
- 金融
- 法律

銀行送金ネットワーク

ナラティブの数量化



With
CJ Tessone
(UZH)



法とデータサイエンス

With
T. Altwicker(UZH)

経済ナレッジグラフ

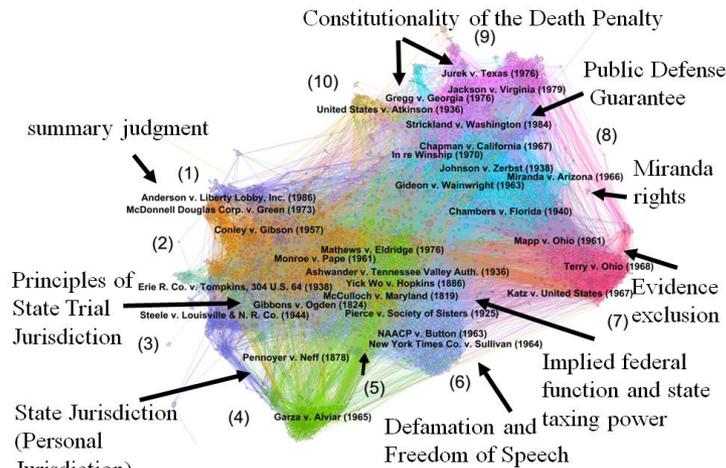


図 5.3 米国判例引用ネットワーク

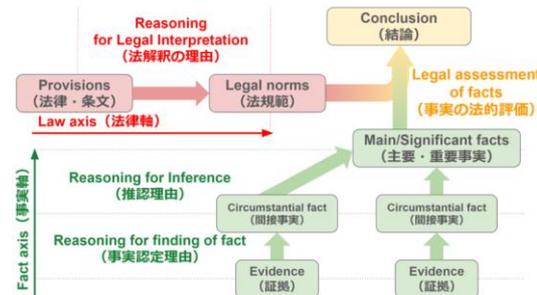
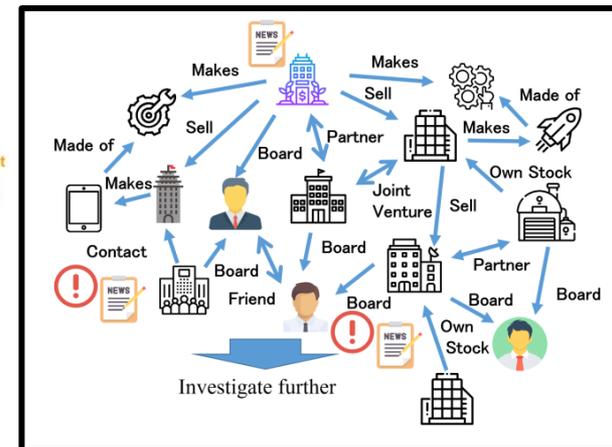


Fig. 1 Logical Structure in Judicial Decisions



オルタナティブデータとナウキャストイング

- これまで:「日本市場でそんな先端的なことをやっても…」
- 今:日銀総裁の交代、物価・賃金の関係、インフレによって 世界的に日本市場に関心が集まっている

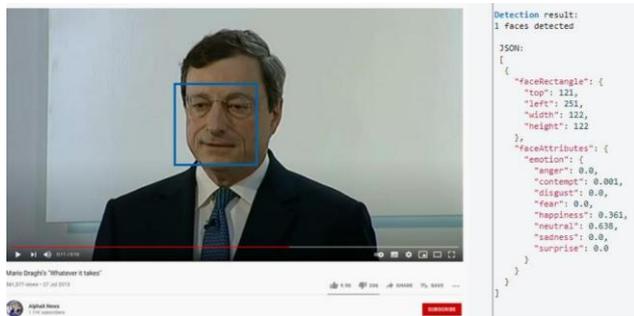
Activity Log

(Search, Electronic Payment)



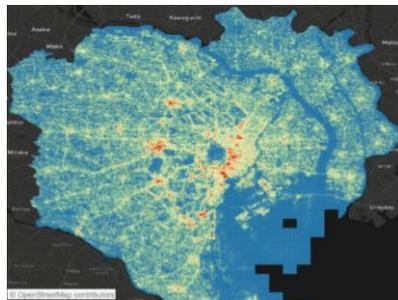
Non-verbal

Communication [Curti2023]



Mobile GPS

[Furukawa2022]



Surveillance

Footage [Nikkei]



Satellite

Image [Stanford]



Job postings [Hrog]

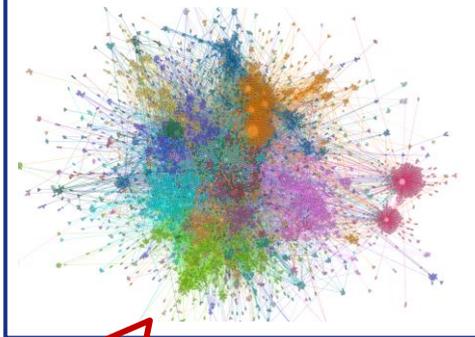


銀行送金ネットワークと指標開発

目的：銀行送金ネットワークデータを市場予測に活かす

送金データを市場予測に活かすイメージ

銀行送金ネットワーク



指標

\hat{y}_t

市場予測



口座を頂点、送金を
辺とするネットワーク
(動的・重み付き・有向)

いくつかのネットワーク統計量が経済の状況と
ある程度マッチしていることを確認済み
→マクロ指標・産業指標の導出,
経済指標のナウキャストイング

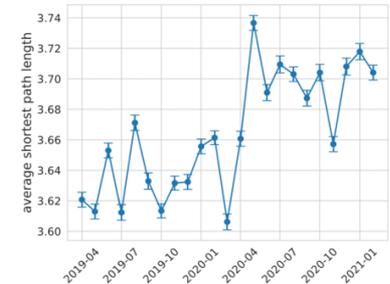
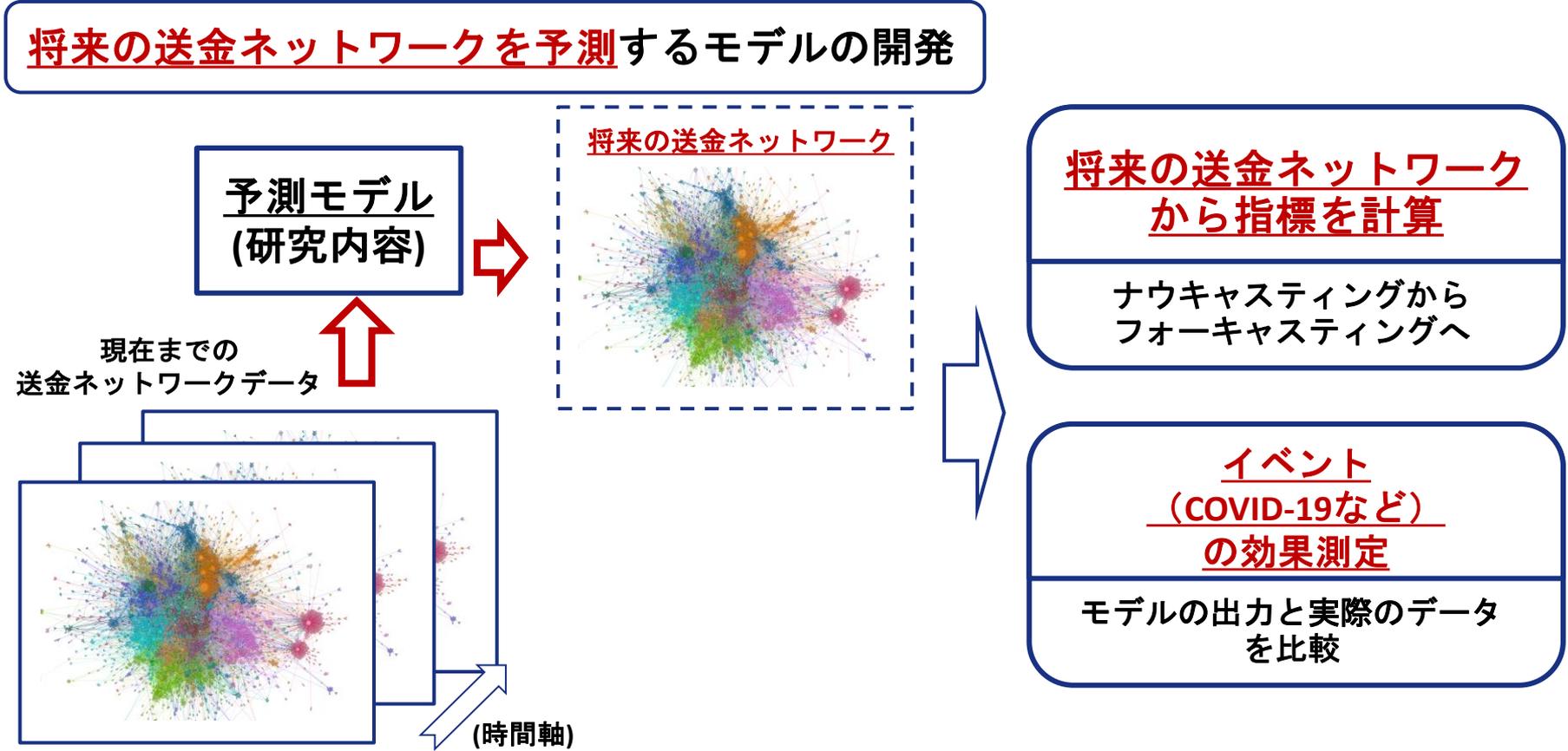


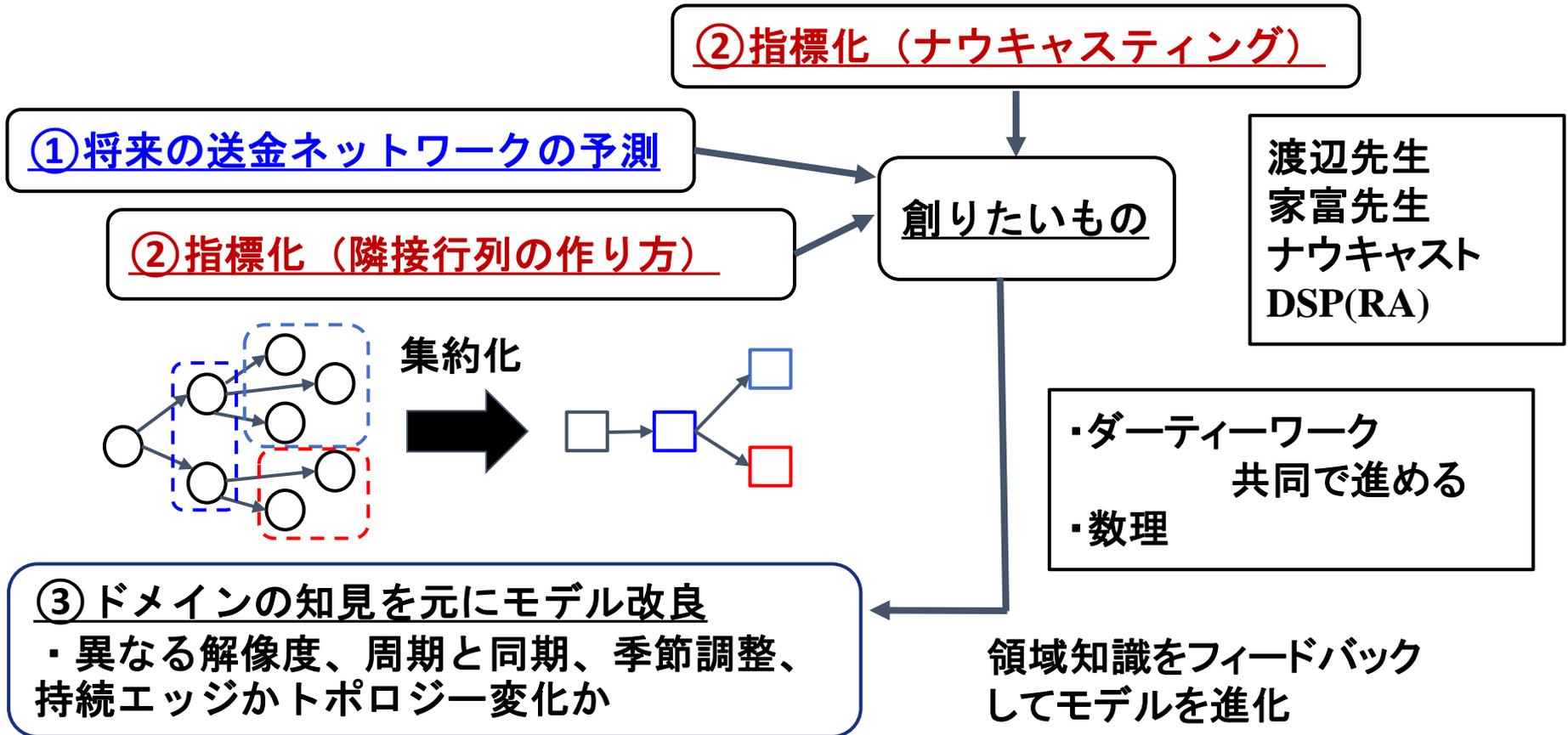
図1：口座間の平均最短距離
COVID-19流行(2020年2~3月)
に合わせて距離が増加→経済停滞を示唆

将来の送金ネットワークの予測



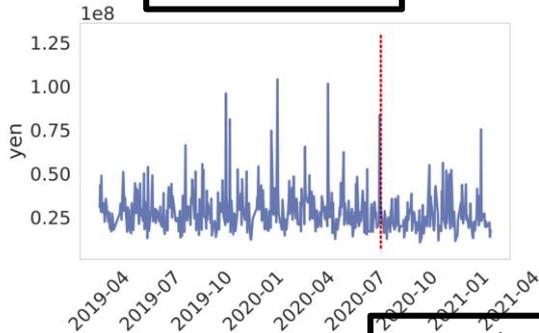
2024年度

□ ②は順調、①も頑張りたい

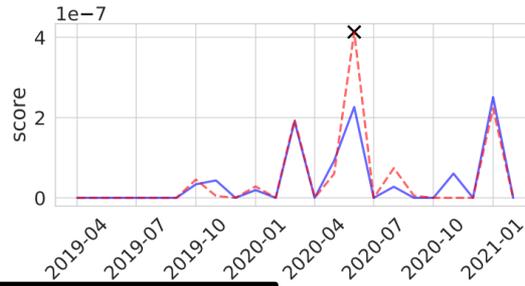


銀行送金データ

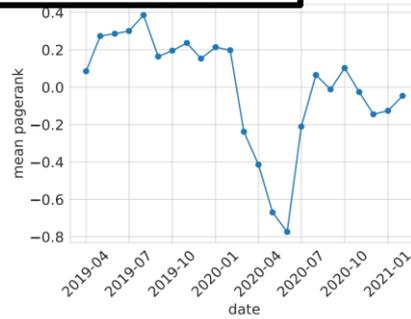
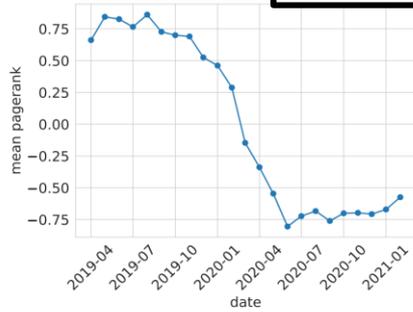
平均送金額



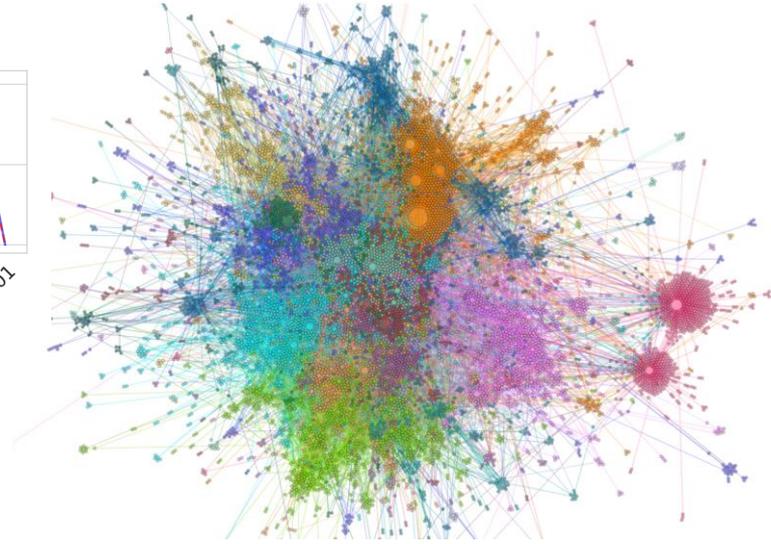
異常検知



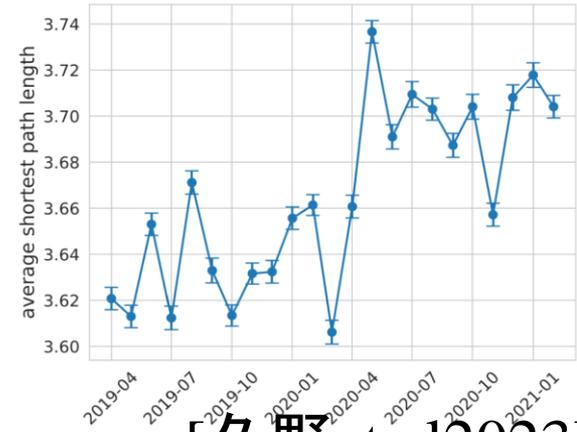
ページランクの時間変化



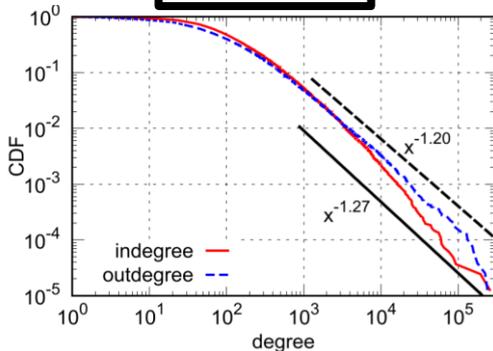
企業ネットワーク



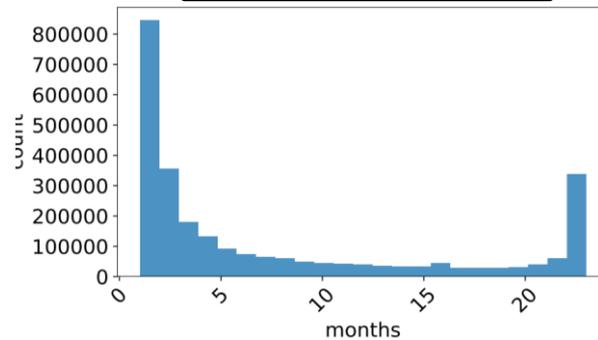
平均最短経路長



次数分布



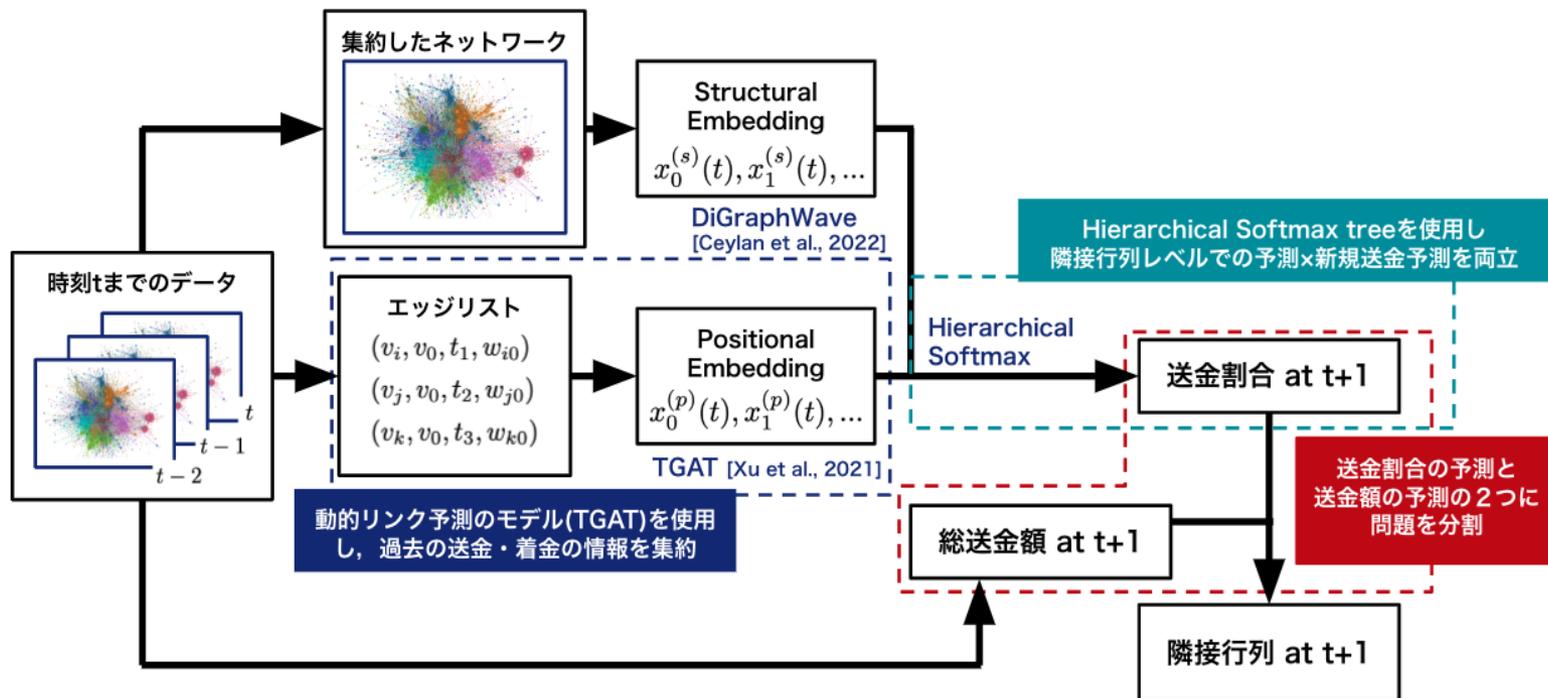
エッジの出現月数



[久野et al2023]

動的リンクフロー予測

- 動的リンク予測のモデルを拡張し重みを予測できるように改良
 - ベンチマークモデルが欲しい
 - 深層学習でどこまで行けるか



- 目覚ましい精度を出すためには？

法とデータサイエンス

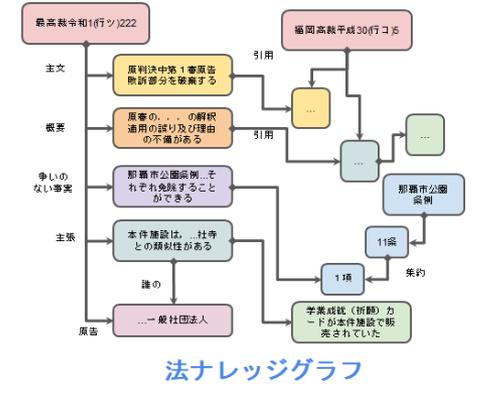
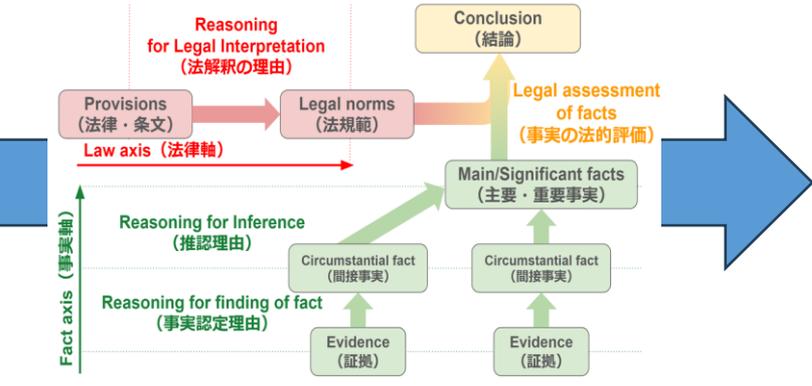
金融契約文書: Trust

検索エンジン・LLMの限界打破: 弁護士、法科生、(株)TKC 地裁判決の構造化 セマンティック・ウェブ

判決文

利息制限法1条1項所定の制限を超える利息を受領した貸金業者が、その預金口座への払込みを受けた際に貸金業の規制等に関する法律18条1項に規定する書面を債務者に交付していなかったために同法48条1項の適用を受けられない場合において、

...
上記認識を有するに至ったことについてやむを得ないといえる特段の事情があるとはいえ、上記貸金業者は民法704条の「悪意の受益者」であるとする推定を覆すことはできない。



数量比較法学/司法の法創造 機能: 法学者

- 日独瑞比較
- 制定法・判例法の収束
- Prof Dr T. Altwickerと共同

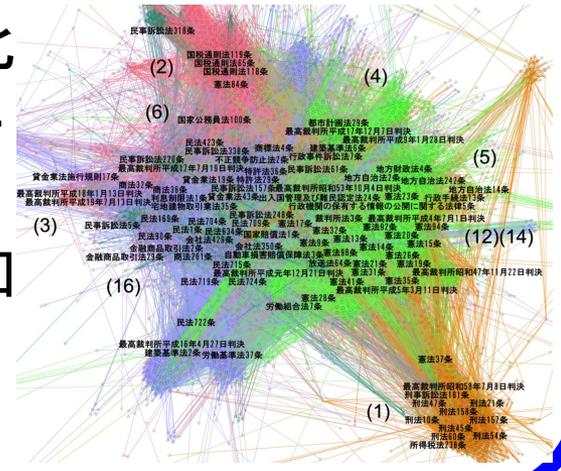


University of Zurich^{UZH}

CLDS
Center for
Legal Data Science

ネットワーク科学

- 規範進化
- ハイパーグラフ変化点検知



法分野におけるAIの技術的課題

with 情理・
近藤先生

- LLMは専門的な領域だと不十分な点が多い
 - 見落としが多い、幻覚
- 統計的パターンの学習

①知識が未整理

前提としてほしい法律・判決フレーム



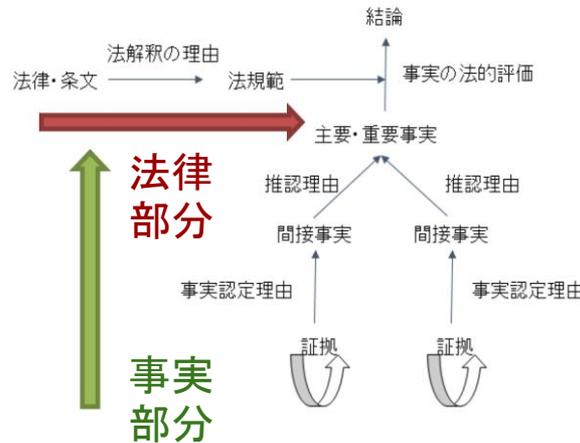
架空の判決

架空の法律

根拠不明な回答

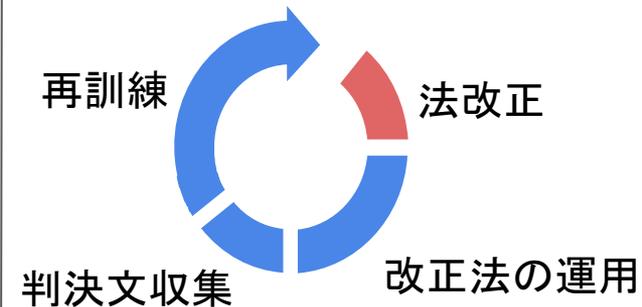
前提が間違った回答

②構造が複雑



通時的・共時的に
構造はますます複雑化

③学習が非効率



都度、有効な法規範で
再訓練は非現実的

規範進化

3. 規範進化の過程

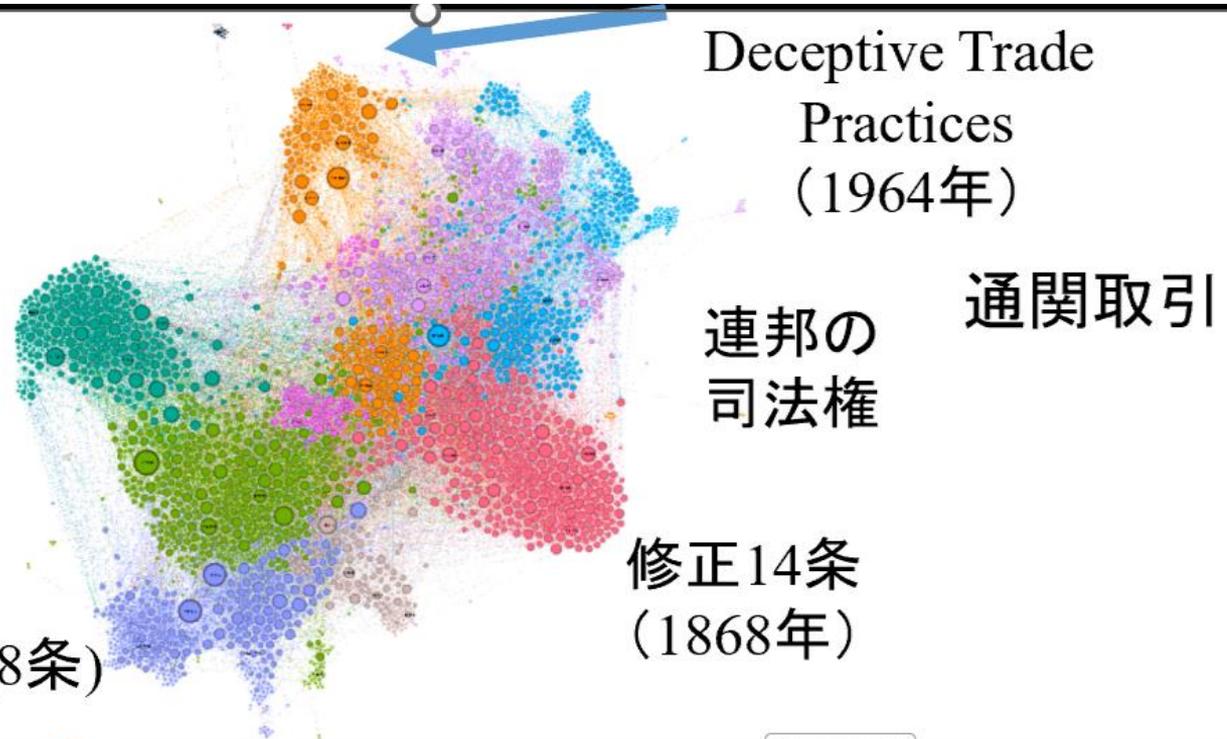
□ 規範進化

□ 判例法 → 引用は判決が多い したため18-19世紀に判例が少ない

米国判例の引用関係[Freelaw]

* 次数500以上に限定 = 3K, 43K (元はノード, エッジ数 = 3M, 30M)

→ 描画の都合で3Mノードから3Kノードに減ら



00:26.49

00:26.49

日本の法・判例構造

- 1998年～2018年までの判決文
- 左：刑事判決の法・判例の共起
- 右：民事・刑事の全体

ハイパーグラフの
クリーク展開

判決文

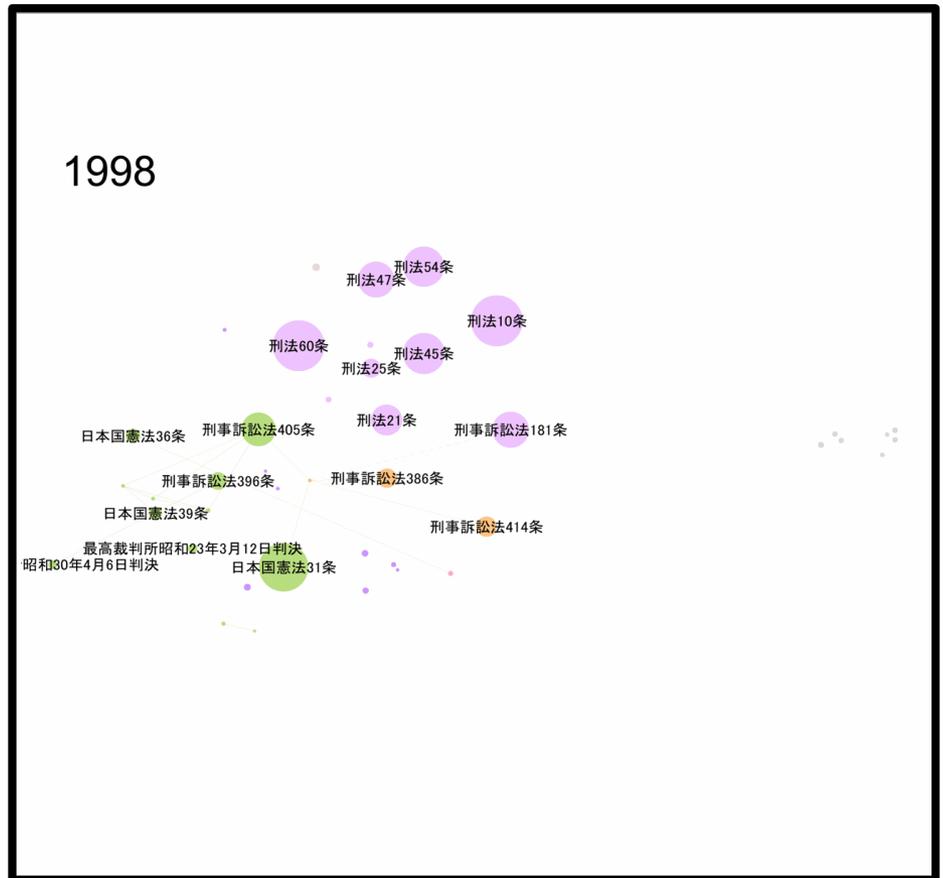
抽出



商標法

不正競争防止法

最高裁判所昭和
38年12月5日



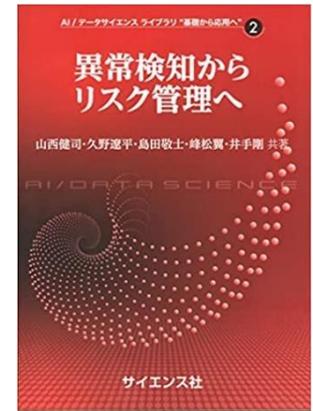
Overruling (判例変更)

□ 動的ネットワーク構造から変化を抽出できるか？

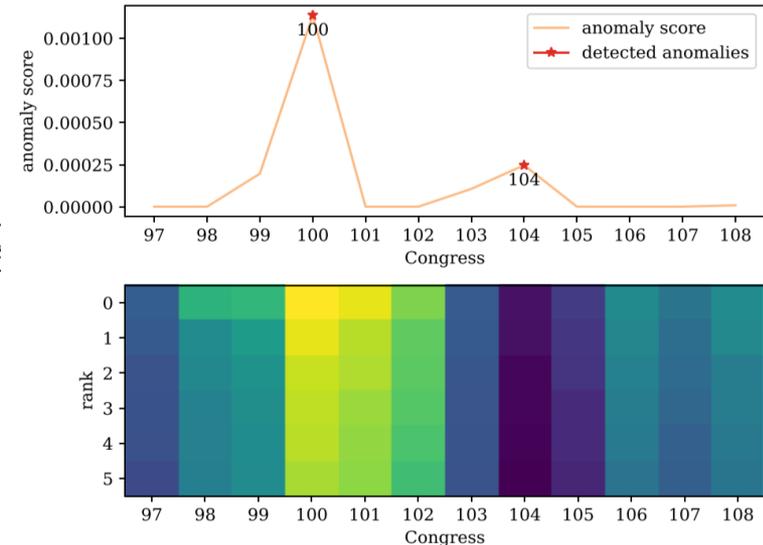
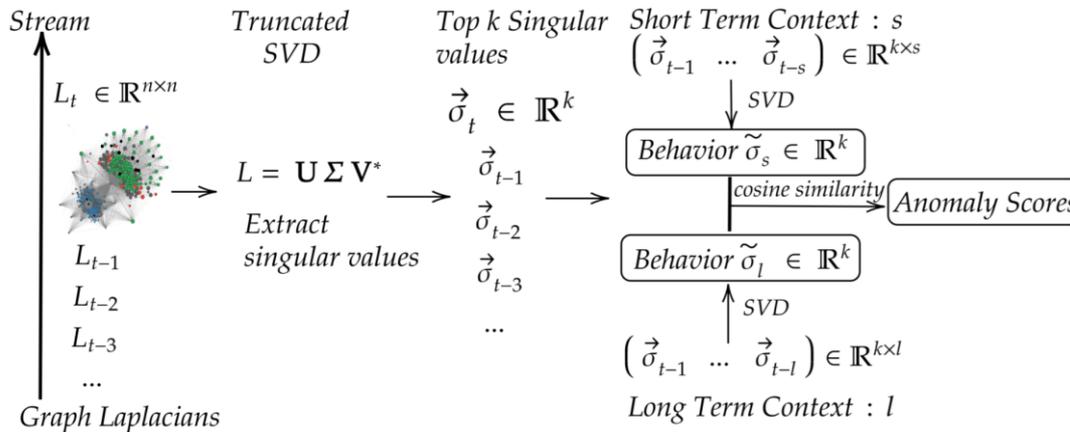
- Schenck v. United States (1919) “Clear and Present Danger” → Brandenburg v. Ohio (1969) “Imminent lawless action;“
- Goldberg vs Kelly (1970) : ”interest-benefit ratio” → Mathews vs Eldridge (1976) : ”cost-benefit analysis”
- Regina vs Hicklin (1868) : ” Subjective determination of obscene material” → Roth vs US (1957) / Miller vs California (1973) : ”average man”
- Betts vs Brady (1942) : ” No criminal defense attorney is not a violation of the 14th Amendment” → Johnson vs Zerbst (1938) / Gideon vs Wainwright (1963) : ” The right to have an attorney”
- Conley vs Gibson (1957) : ” Plain and concise statement” → Ashcroft vs Iqbal (2009) : ”plausibility”
- Pennoyer vs Neff (1878) : ” theory of territorial sovereignty” → International Shoe Co. vs Washington (1945) ” Theory of Fairness”

ネットワーク変化点検知

- ネットワークの場合: signature vectorを抽出
- グラフラプラシアンの特クトルを用い異常スコアを計算



- コンテキスト行列: $C = \begin{pmatrix} \vec{\sigma}_{t-l-1} & \vec{\sigma}_{t-l-2} & \dots & \vec{\sigma}_{t-1} \end{pmatrix} \in \mathbb{R}^{n \times l}$
- Zスコア: $Z = 1 - \sigma_t^\top \bar{\sigma}_t$



- ハイパーグラフの場合は?

ハイパーグラフ変化点検知

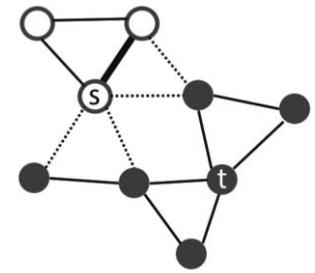
- [Bolla1993]のハイパーグラフラプラシアン

$$L_b := D_v - HD_e^{-1}H^\top$$

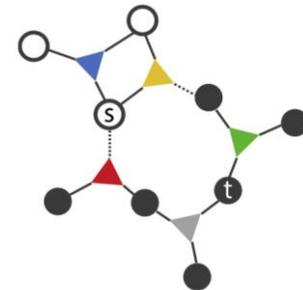
- [Zhou et al.2005]のハイパーグラフラプラシアン

$$L_z := I - D_v^{-1/2}HWD_e^{-1}H^\top D_v^{-1/2}$$

- クリーク展開やスター展開とほぼ同じ[Agrwal 2007]



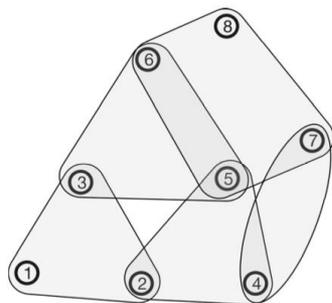
(b) Clique expansion



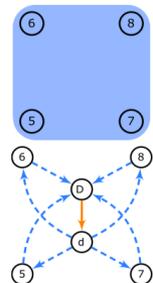
(c) Star expansion

Algorithm	Graph	Matrix
Bolla	Clique	Combinatorial Laplacian
Rodríguez	Clique	Combinatorial Laplacian
Zhou	Star	Normalized Laplacian
Gibson	Clique	Adjacency
Li	Star	Adjacency

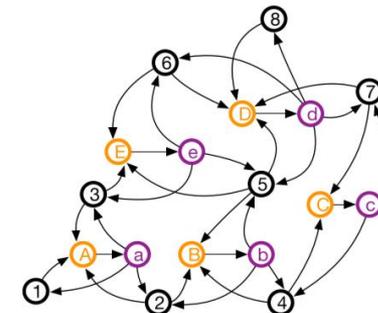
- ハイパーグラフで変化点検知を精度よく実行するには？



(a) original hypergraph



(b) single hyperedge reduction gadget

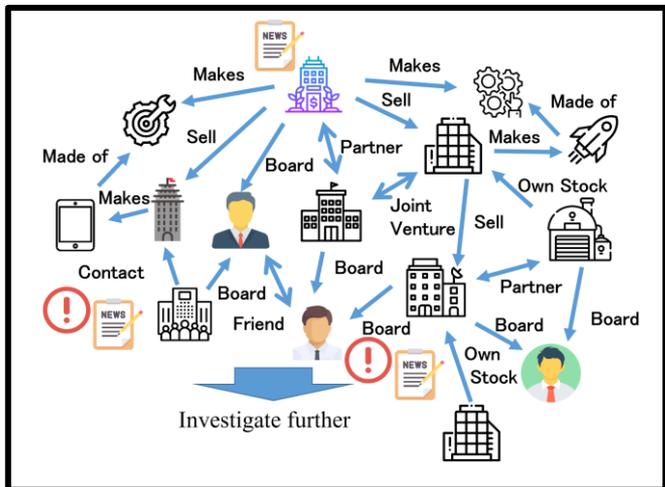


(c) expanded graph

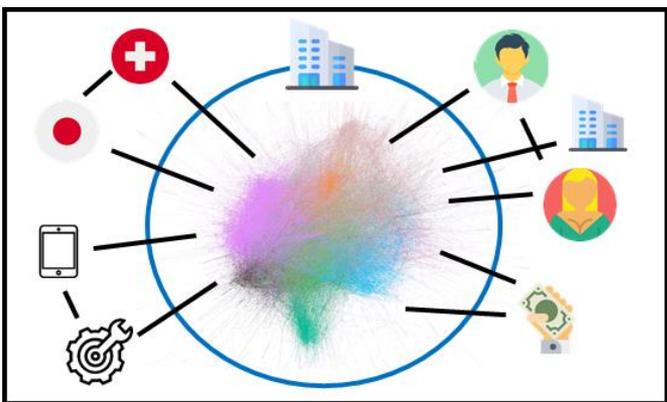
経済ナレッジグラフ

関係に注目した世界中の企業・人・製品のデータベースの構築

- 金融データをベースにナレッジグラフ作成 → **不正検知**[Hisano++2020]
- リスク管理、情報の信頼性、データ駆動型政策支援



● ノード: **5千万**、エッジ: **4億**

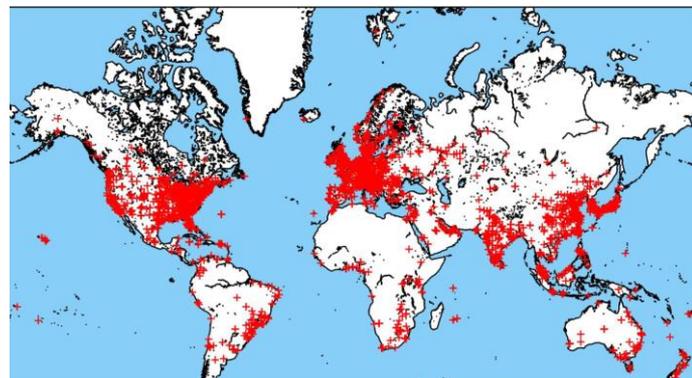


● 金融データ7種と公開情報2種でデータベース作成

Source	Date of Acquisition	Node types	Relation types	Num Nodes	Num Edges
Dow Jones Adverse Media Entity	Dec 2016	Firm	Location, Homepage	132,127	390,320
Dow Jones State Owned Companies	Dec 2016	State Owned Firms	VIP, Employee, Owner	280,995	702,172
Dow Jones Watchlist	Dec 2016	VIPs, specially interested person	social relations	1,826,273	8,322,560
Capital IQ Company Screening Report	Dec 2016	Firms	Buyer-Seller, Borrower etc	505,789	2,916,956
FactSet	Dec 2015	Firm, Goods, Industry	Parent-child firm, Issue Stock	613,422	8,213,225
FactShip	Jan 2017	Firm, Goods, Invoice etc	Overseas trade etc	16,137,550	36,345,381
Reuters Ownership	Dec 2016	Owners, Stocks	Issue, Own	1,560,544	121,769,151
Panama papers	Jan 2017	Entities, Officers	shareholder of, director of	888,630	1,371,984
DBpedia	Apr 2016	Various	Various	35,006,127	249,429,771

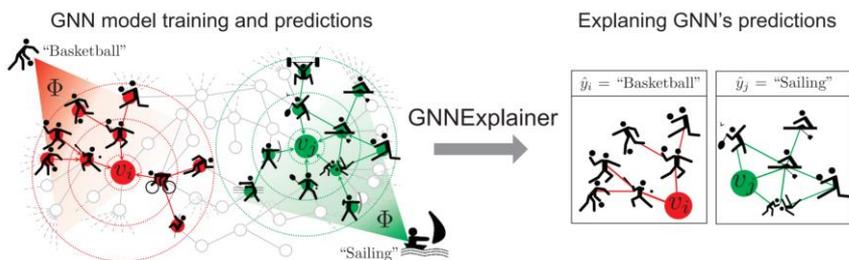
● 関係型(頻出上位25) ● 予想対象(約3.5万)は世界中に存在

Rank	Relation	Number
1	located_in	2,723,162
2	customer	717,019
3	supplier	713,434
4	own_stock	493,316
5	belongs_to_industry	359,425
6	strategic_alliance	348,352
7	creditor	339,184
8	recieve_goods	330,311
9	send_goods	319,292
10	issue_stock	187,498
11	make_products	181,574
12	competitor	174,487
13	part_of_industry	172,621
14	borrower	153,203
15	domain	131,153
16	distributor	116,262
17	subsidiary	107,119
18	parent-company	107,117
19	associated-person	100,699
20	international_shipping	95,050
21	associate	72,685
22	landlord	62,904
23	http://dbpedia.org/ontology/party	55,653
24	employer	47,901
25	employee	47,184

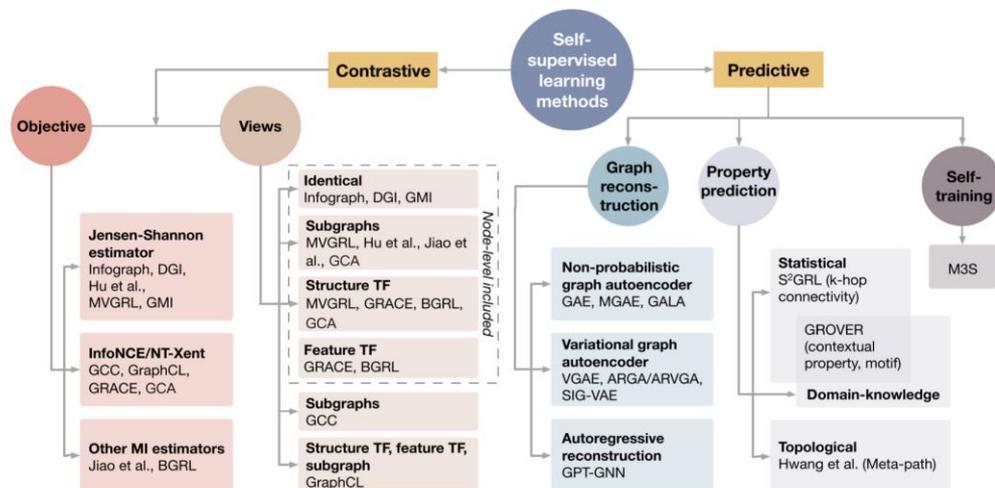


xAIとSelf supervised learning

- 旧データ → データ更新(2023)
- 経済応用を考えると解釈性は重要
 - GraphExplainer[Ying++2019]
- Self supervised learning[Xie++2021]
 - あるネットワークの学習結果を元のデータは共有せずに他で使える → 経済価値が大きい



[Ying++2019]



[Xie++2021]

ナラティブの数量化

ネットワークの集団・動的変化

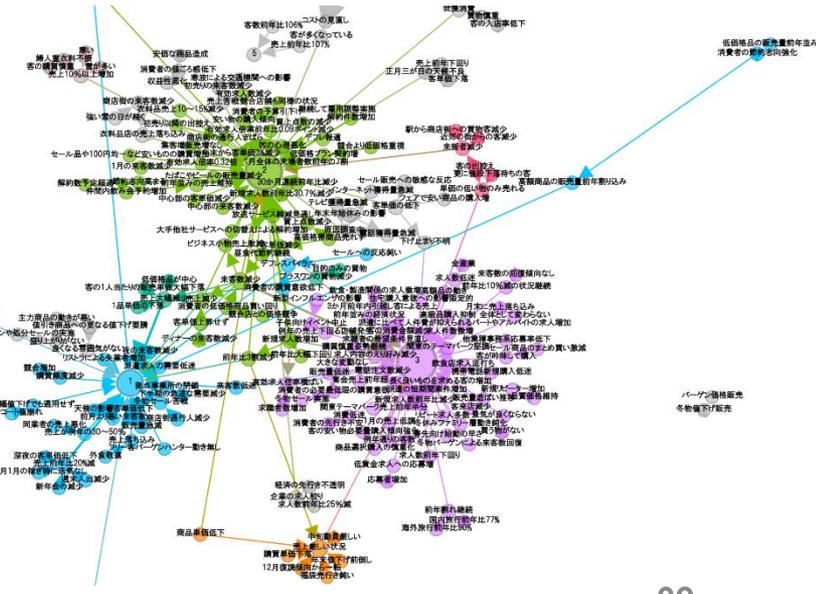
意見(生声・パブリックコメント)の構造化 → 行政利用

北海道;2.0;・客の出控えもあり、商店街を訪れる来街者は3か月前と比べて減少している。时期的な要因もあるかもしれないが、特に駅から商店街へと流れてくる買物客の減少が著しく、近郊の街から訪れる客が少なくなっていることがうかがえる。

{'ノードリスト': ['客の出控え', '来街者減少', '时期的な要因', '駅から商店街への買物客減少', '近郊の街からの客減少', '2.0'], 'エッジリスト': ['客の出控え -> 来街者減少', '时期的な要因 -* 来街者減少', '駅から商店街への買物客減少 -> 来街者減少', '近郊の街からの客減少 -> 来街者減少', '来街者減少 -> 2.0']}

「AI は本格的に規制し、もし AI を投稿したものに対しては容赦なく罰金 1 億、懲役 100 年か死刑にすべきだ。何故なら AI に絵を食われて、あたかも自分の商品だと偽って商売したり、公式絵を食ってあたかも自分が描いたかのように見せつけたりと悪用が多発している。実際私の相互さんの相互さんが被害にあってしまい、絶筆してしまった。AI のせいでもう素晴らしい作品を拝見できないとなると非常に腹立たしい許せぬ行為だ。このような惨状を防ぐのにはイラストサイトに投稿しないなど規制すべきである。実際海外では AI 規制を取り掛かっている中日本だけやっていないのはどうかと思う。いい加減クリエイターのことを考えろ。」

(A, 罰金や懲役や死刑で規制すべきだと思っている, AI),(AI, 悪用の原因である, 絵),(AIによる悪用, 非常に腹立たしい許せぬ行為であるとされている, 素晴らしい作品),(規制, 防ぐべきだとされている, AIによる悪用),(海外, AI規制を取り掛かっている, AI),(日本, AI規制をしていない, AI),(A, 考えるべきだと思っている, クリエイター)



まとめ

□ ネットワーク × 金融・法律

□ 実データをベースにモデル開発したい

- 今日あげたものに関してはデータ整備の作業はほぼ全て終わっているののですぐモデルの話に取り掛かれる

久野遼平、大西立顕、渡辺努

「ネットワーク学習から経済と法分析へ」(サイエンス社)
が近日発売