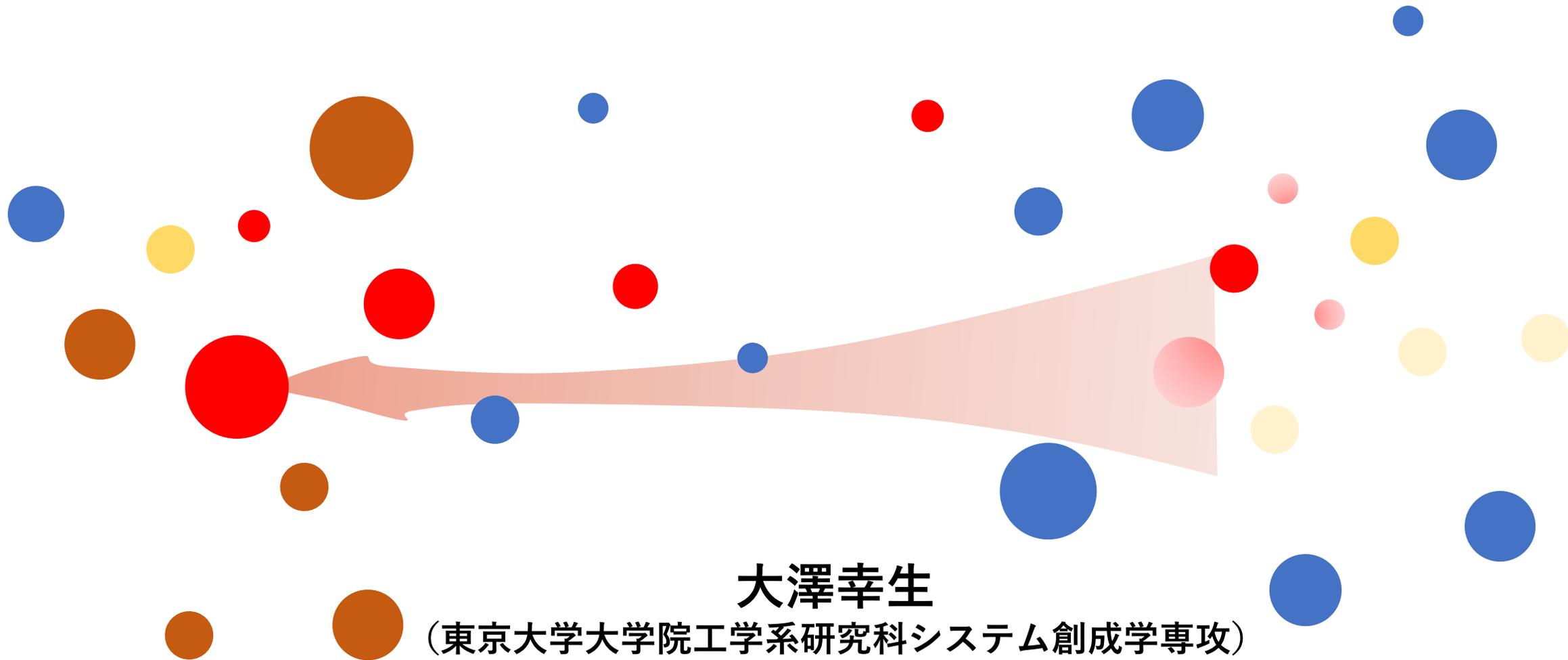


新・チャンス発見学

～変化の『説明』でチャンスを創る～



大澤幸生

(東京大学大学院工学系研究科システム創成学専攻)

東京大学大澤研究室

が繋ぐ多業種多国籍ビジネスデータネットワーク(一部のみ記載)

ニッセイ アセットマネジメント(株)
市場における異常検知

FRONTEO 行動情報科学研究
所
人工知能の基盤モデル・アドバイザー

共同印刷 (株)
異業種間のデータ流通・活用の実証実験

日本データ取引所と
広告・食品業
テレビ・WEB・SNS・店舗・食卓をまたいだコミュニケーション効果の研究

(株)構造計画研究所
KeyGraph商品化、IMDJ伝承等
多様な共同研究

Abeam Consulting トッパン・フォームズ(株)
TEEDAによるデータ構想 データ駆動ソリューション
teradata
データ活用手法を理解するワークショップ

SDGsオープンイノベーション ユナイテッド・スーパーマーケット・
プラットフォーム(多数社) ホールディングス(株)
マーケティングへのデータ利活用

エステートテクノロジーズ(株)
不動産情報提供・技術顧問

(株)ウフル
データ連携による
観光地の安全安心化

下田市:コミュニティシステム
(下田市、東京大学大澤研究室等)

民公

横浜データ共創ラボ
(横浜市、東京大学、富士通研)

内閣官房

“医療データ×天候データ×市場データ ×電力データ×人流データ×・・・”で新サービスの創出

UTokyo joins COVID-19 data
exchange

A new platform for data collaboration
with Dawex Systems

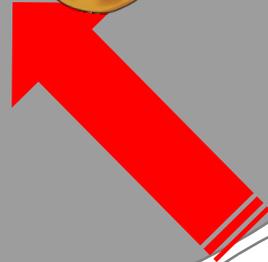
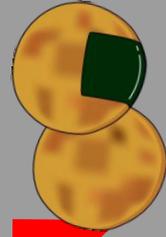
国際

実業界におけるデータジャケットの国際標準化
内閣府SIP2

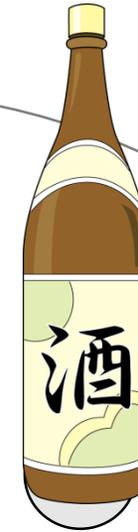
チャンス発見 (2000-)とは

あのスルメ、
そそるけど...高
いよなあ

スナックコーナー



¥1000



酒コーナー

Books on チャンス発見



素材: <http://www.civillink.net/fsozai/fish.html>, <http://illpop.com> から引用

KeyGraph® を適用し専門性の異なるマーケッターが会議

臼井・大澤:組織内合意形成のためのチャンス発見:生地メーカーにおける触視覚補完型KeyGraphの効果, ファジィ学会論文誌, Vol.15 No.3, pp.275-285

プロセスは「**データを作る**」ところから

サンプル発注票

貴社名: ○○○○

2001/10/04

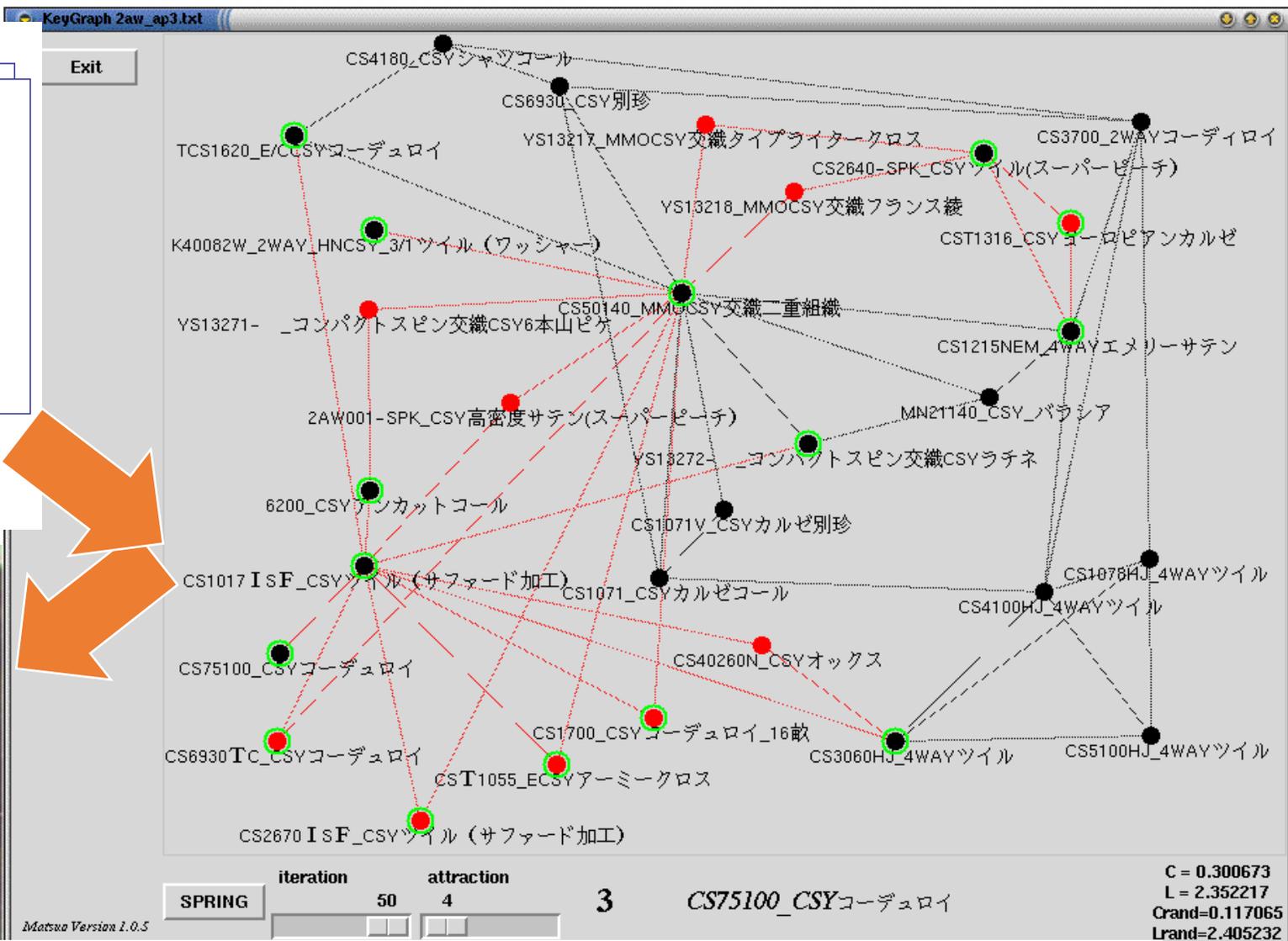
シリアル番号	商品番号	商品名
....	CS4550N	4WAY...
2AW116	...	MMO CSY
2AW371	CS2016N	...
...

A社: CS3700, CS4444, CS5100, ...

これがデータ

B社: CST1316..., CS3700, ...

C社: ..., CS3700, MN21140, ...



KeyGraph (Ohsawa, Benson, Yachida 1996)だが 本バージョンのKeyGraphはLinuxで、松尾豊 君が作製(当時D2でJST大澤研研究員, 現東京大学教授)

データは「機械の学習→『予測』『検出』」だけでなく 「人の覚醒→『創造』『行動』」に使う

売れ筋(1): ビジ
ネスウェア等に使う
「きれい目」生地



売れ筋(2): 若さを感じ
させる「着古し系」生地

13位 (/800商品)のヒット

「若く爽やかに変身したい」
「仕事帰りには着替えたい」
→コーデュロイの新ジャケット



人気のスーツ



17年前のあるメディアのAI観

「AIが示せるのはせいぜいヒント止まり、ということはビジネスマンとしてまた人間として、喜ぶべきことといえるだろう」

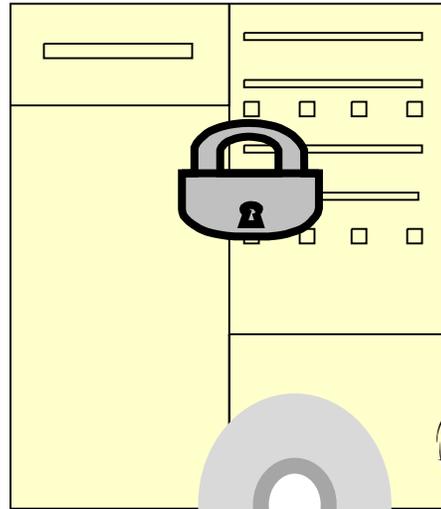
(2003年、AERA編集長 久保田裕氏「人工知能でチャンス発見」 AERA 2003.4.7, p.63-64 より)

本質は変わっていないのですが、
「期待」には羽があって足が無く、
どこまでも飛んでゆきます・・・



データジャケット： 素人でも「期待」から入るデータ市場 Innovators Marketplace on Data Jackets (IMDJ)

Ohsawa, Y., Liu, C., Suda, Y., and Kido, H., Innovators Marketplace on Data Jackets for Externalizing the Value of Data via Stakeholders' Requirement Communication AAAI 2014 Spr. Sym. on "Big data becomes personal: Knowledge into Meaning, Stanford (2014)



過去の気象
日々の気温や湿度等生活に関する情報
@場所
@日時
@気温
@湿度
...

DJ(a)

Step1) データジャケットをシェア：

記載・公開可能な範囲でデータ概要(変数集合、データ概要、期待するユースケース)を記載

DJの実例

<DJ1758:自転車利用可能道路>	
概要	ある時刻における、年齢、性別、車種などに合わせた適切な自転車ルートを表示するナビ。東京五輪開催時に混雑が予想される地域において、自転車用レーンが設置されている区画、もしくは車道・歩道が十分に広く、自転車が安全で快適に走行できることがわかった区画のマッピングデータ。年齢、車種など多様な条件で自転車で区内を走り回って収集した。国土交通省から自転車レーン設置区画のデータをもらう。自転車利用者からの情報提供。
変数	丁, 区, 安全な区画の連続性, 年齢, 性別, 時刻, 歩道の広さ, 自転車の種類, 車線形態, 車道の広さ, 運動経験
共有ポリシー	With anyone.
データタイプ	Text, Table
データフォーマット	PDF



ID-POS
個人の生活における購買内容
個人別購買)
@顧客ID番号
@年齢・性別
@購入品目
@日時

DJ(b)

データジャケット(DJ): メタデータ的一种だが、主観を含めてひとつのデータに対し複数のDJを与え、仮想的なマイニングストーリーを創造することによってデータの利用価値まで評価する点が特徴

(Ohsawa et al, Data Jackets for Synthesizing Values in the Market of Data, *Procedia Computer Science* 22, 709-716 (2013))

データ市場：意味レイヤーの階層構造

真の要求：生活に健康要因がある疑い

↑Why

社会的要求（「健康変化の要因を知りたい」）

IMDJ

利用者
→考案者：
要求と評価

実現手法（分析計画・実施条件・コストなど）の提示



利用（事業）者

考案者→利用者：

考案者(Data scientists etc)

データ結合・利用の方法を提案(例:提案1, 2)

Proposed plan: Combine data behind DJ(A), DJ(B), and DJ(C) to learn rules such as
 (1) $temperature(t) > 25c \rightarrow consumption_beer(t) > 1L$
 (2) $consumption_beer(t) > 0.7L \rightarrow fat(t) > \alpha \rightarrow \gamma-GTP \beta.$

DJ結合

プラットフォーム

Why

How

計画・報酬の提示による信頼の提供

データ提供者



Data(A)

過去の気象
@場所
@日時
@気温
@湿度
...
DJ(A)

POSと通信
@顧客ID番号
@年齢・性別
@住所
@購入品目
@日時
@メッセージ内容
DJ(B)

健康診断結果
@年月・日時
@CHE
@AST
@ γ -GTP
...
DJ(C)

データ提供・送信

Data(C)

機械学習による予測、計画
...
@...
@...
@...
@...
...
TJ(x)

関連ルール分析
@購入品目
@状況(日時、場所など)の基数
TJ(y)

Tool(x)

Tool(y)



データの真の権利者(controller)



データ収集

照明種別c	管理No	柱種c	適合ランプc	町名c	丁目	番地	号	緯度	経度	ラ
街路灯	1	ポール大型	CM220	湯島	1	13	2	354209902	1394603203	1
街路灯	2	ポール大型	CM220	湯島	1	12	5	354209792	1394601561	1
街路灯	3	ポール大型	CM190	湯島	1	12	3	354209675	1394559999	2
街路灯	4	ポール大型	CM220	湯島	1	11	13	354209923	1394559225	1
街路灯	5	NTT共架	LED33W	湯島	2	11	15	354215362	1394555127	2b39
街路灯	6	ポール大型	CM220	湯島	1	11	10	354210382	1394557204	19800
街路灯	7	ポール大型	CM220	湯島	1	7	12	354210922	1394554851	19800
街路灯	8	ポール大型	CM220	湯島	1	7	9	354211217	1394553007	19800
街路灯	9	ポール【139	HF125	湯島	1	7	9	354211036	1394552377	6000
街路灯	10	ポール大型	CM220	湯島	1	6	7	354211245	1394551131	19800
街路灯	11	東電共架	LED33W	湯島	2	4	14	354212575	1394552722	2635
街路灯	12	東電共架	LED33W	湯島	2	4	1	354213096	1394551953	2635
街路灯	13	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	14	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	15	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	16	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	17	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	18	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635
街路灯	19	東電共架	LED33W	湯島	2	4	10	354213457	1394551383	2635

IMDJ事例A: 街灯データとGoogle mapの結合(Data(a)+Data(b))

街灯設置データを地図にマッピングし、最明経路を表示(早矢仕・池上ら)



DJ(b)区役所等へIMDJ結果を示し入手した街灯設置データ(非公開データが提供された)



利用者の声

- 後楽園へ抜ける近道が常々暗いと思っていた
- マップを見て明るい道を探し新しい行き方を発見した
- 飲み会の帰りに安全なルートを確認できる
- 近道だと思って通った道が意外と暗くて怖かった

DJ(a) Google map

変化を検出？予測？それとも説明？

- **検出:** 時刻 t_c の前後で、学習された前提条件（パラメータ、構造など） f が大きく変化したかどうかで検出可能 $|\theta(t < t_c) - \theta(t > t_c)| > c$.

技術分野: センシング（これも広義）、統計、機械学習、...

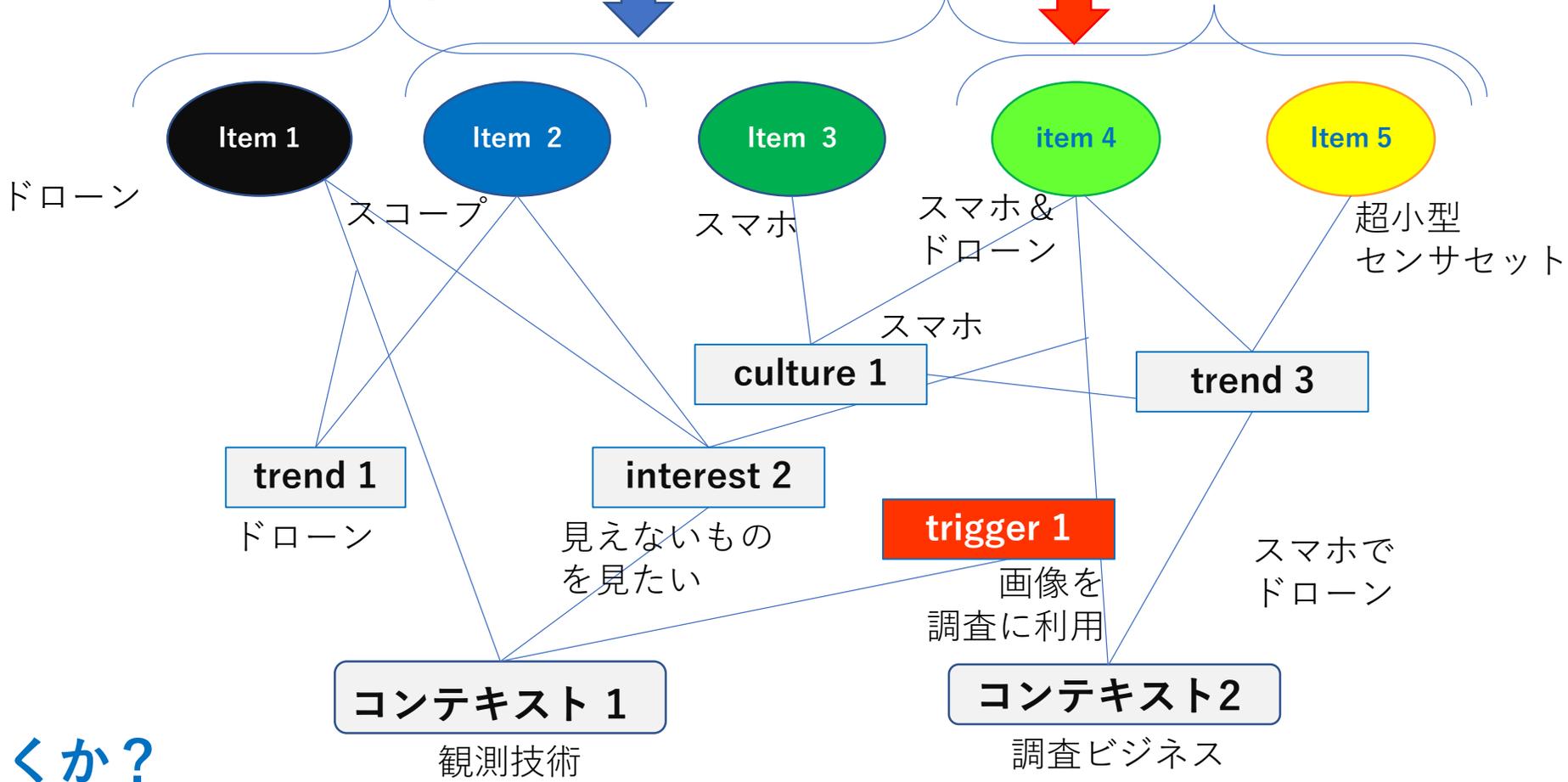
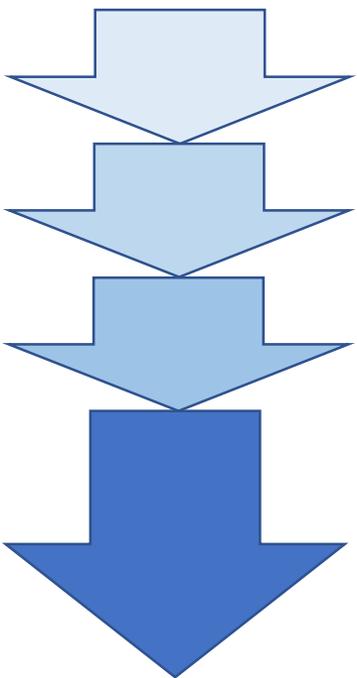
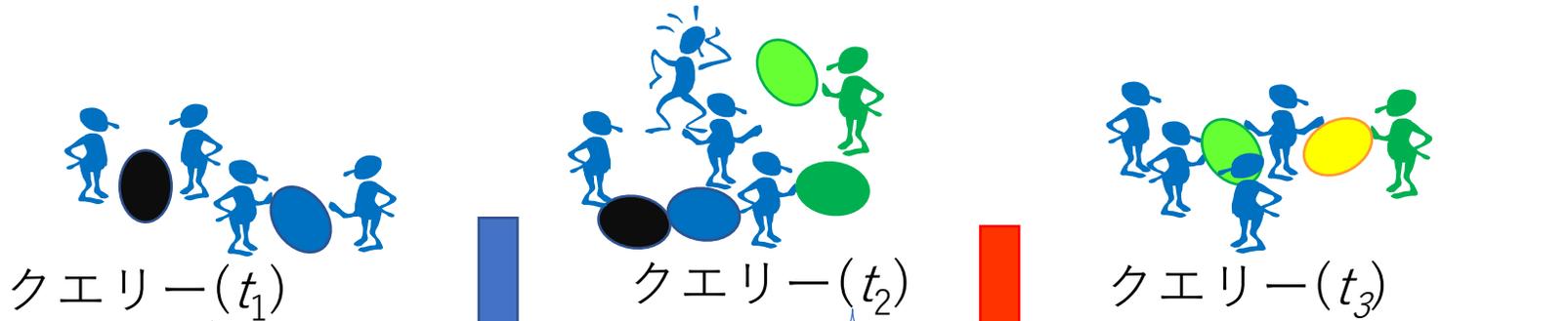
- **予測:** $f(t_p + \Delta t)$ を、 t_p 以前のデータから学習した $df(t)/dt$ から求めるなど（ d^2/dt^2 や離散的規則を用いる場合も含めた広い意味で）.

技術分野: センシング、統計、機械学習、...

- **説明:** $f(t_c)$ の変化点の候補を求めた上で $f(t_c + \Delta t) - f(t_c)$ の原因に接続。

技術分野: （上記に加え）データ可視化、モデル科学、**データ連携**

変化の検出 vs 説明



どう掘ってゆくか？

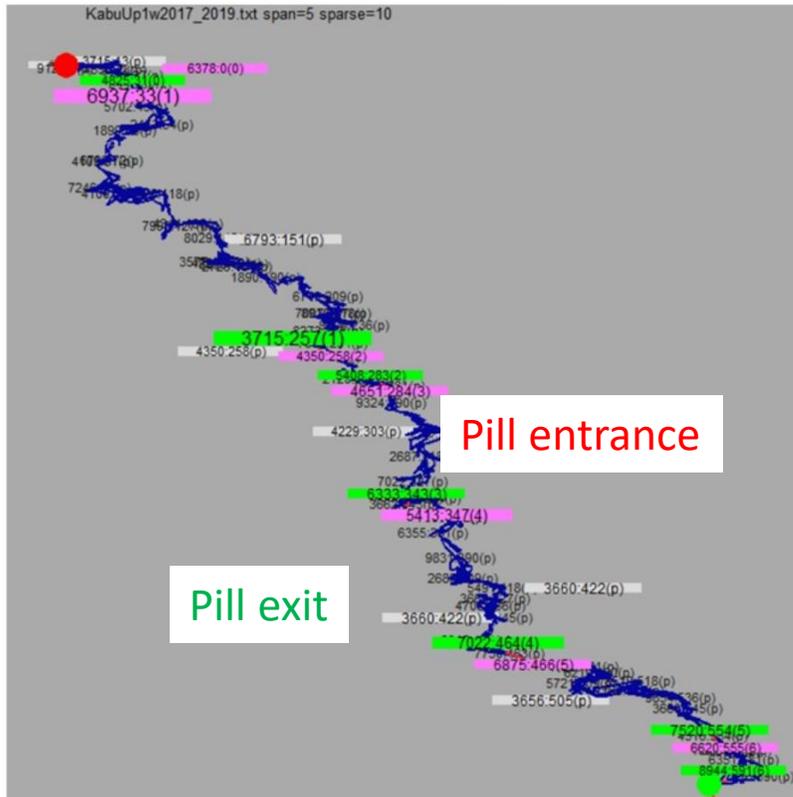
事例B マーケター、証券アナリストとのコラボ(変化を説明し、「売り時」をつかむ)

要求 R: 市場(全体と個別商品)変化の原因(商品等?)が知りたい
 ソリューション(回答案) S: 変化のタイミングを把握し、外部情報と合わせて検討

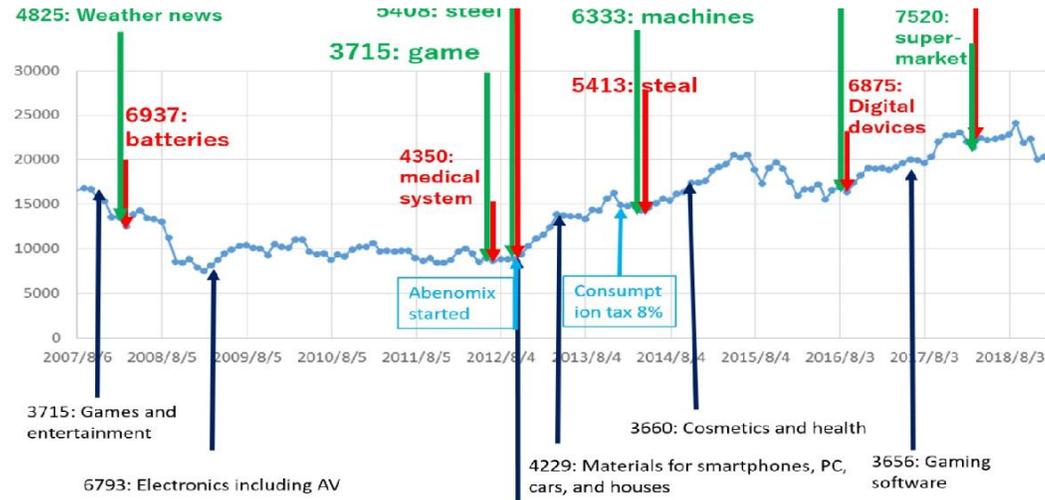
予兆・変化の説明

Data for realizing S: $\{DJ_{567}$: 購買履歴データ, $D(T)J_5$: Tangled String}

(a) Tangled String (TS)でタイムスケール(考慮する変化の遅さ)を5週間に設定した東証銘柄株のシーケンス



(b) TSでの赤・緑の時点に現れた銘柄を日経平均の変動に併せたもの



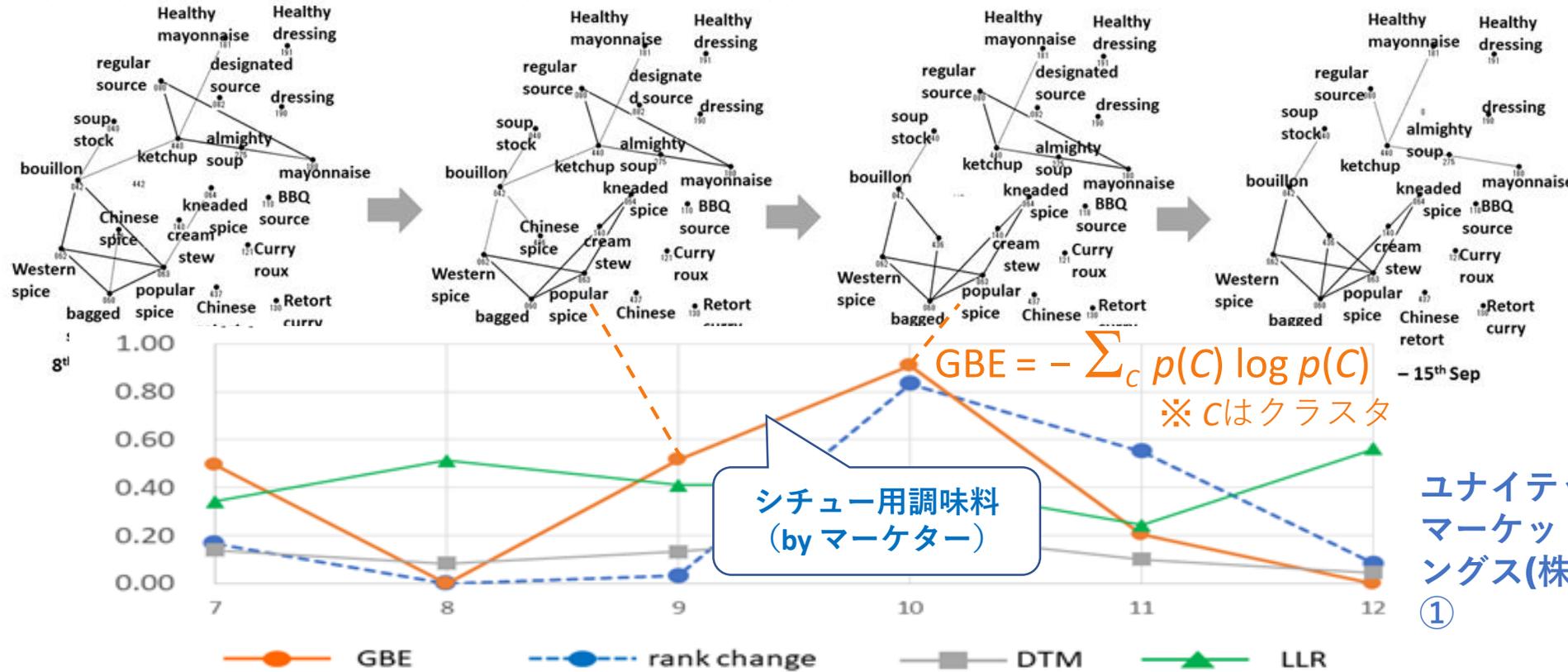
(c) TSにおける赤ノードが実際の株価上昇の直前を検知する効果



事例C 要求「お客様の行動が変化するのは何故？きっかけを知りたい」

新しいAI: データ連結による「変化の説明」

KeyGraphの構造とリンクしたGraph based Entropyの変化を捉える



ユニテッド・スーパーマーケット・ホールディングス(株)様との共同研究
 ①

Google Trend 「シチュー」の結果



事例D 「地震の予兆を発見し背景を説明したい」

34 DJs 7 Solutions 1 Requirements

ALL sharable condition specific purchase research undecided other

震源リスト

地震の震源となった位置とマグニチュードのデータを日別で取得可能。...

小売業のポイントカード会員購買履歴データ

ポイントカード会員の購買履歴データ...

生活者行動分析サービスDo-Cube

性別・職業・趣味属性を推定済みの24万人のAmebaブログユーザーDBを用い、キーワードに対する発言...

地震計設置箇所

県内の地震計設置箇所の位置情報です...

台湾地震資料(~20140514)

中央気象局提供之台湾1995年1月~2014年5月14日之地震資料 参考連結 : http://www....

Mobile Phone Activity Log

The usage of mobile phones has been popular in mos...

Numbeo - Crime

How serious you feel the level of crime is. Change...

日々の天気図

2002年からの全日の天気図が月別に取得可能。各日毎に気圧配置の説明、その日に起こった気象に關す...

Sol: 50 Sol: 46 Sol: 340 DJ: 311

小売業のポイントカード会員購買履歴データ

概要
ポイントカード会員の購買履歴データ商品ごとの売上推移、購買購入層の把握。マーケティング施策の検討材料。

変数ラベル

性別。※会員番号、買上店舗コードはマスキング済み

会員番号の生年月日。※会員番号、買上店舗コードはマスキング済み 会員番号に紐付いた買上商品(IAN)の買上日時

会員番号に紐付いた買上店舗コード 会員番号に紐付いた買上点数 会員番号に紐付いた買上金額

データの種類

NUMBER TIMESERIES

共有条件

条件付で共有可

HOW TO USE

早矢仕, 大澤, “Data Jacket Store: データ利活用知識構造化と検索システム, 人工知能, Vol.31, No.5, 2016

34 DJs 7 Solutions 1 Requirements

新しい食品開発に役立つ健康に良い美味しい水を発見し、輸送体系を確定できる。

ソーシャル分析(発見)、QAサイト、広告アクセス(アクセス数)、ID-POS(リピート率)

東京防災ブックの各国語版の作成

地震の揺れを体験できるシミュレーター装置

犯罪率と地震発生率と汚染が少ない都市を発見

発信された土地・時間ごとにテキスト(SNS)をクラスタリングしてみる(発見的に)

電子書籍を買うポテンシャルがあるのに買っていない人を発見

Sol: 50 Sol: 46 Sol: 340 DJ: 311

Solution 46

ソリューション概要

電子書籍を買うポテンシャルがあるのに買っていない人を発見

組み合わせたDJ

小売業のポイントカード会員購買履歴データ

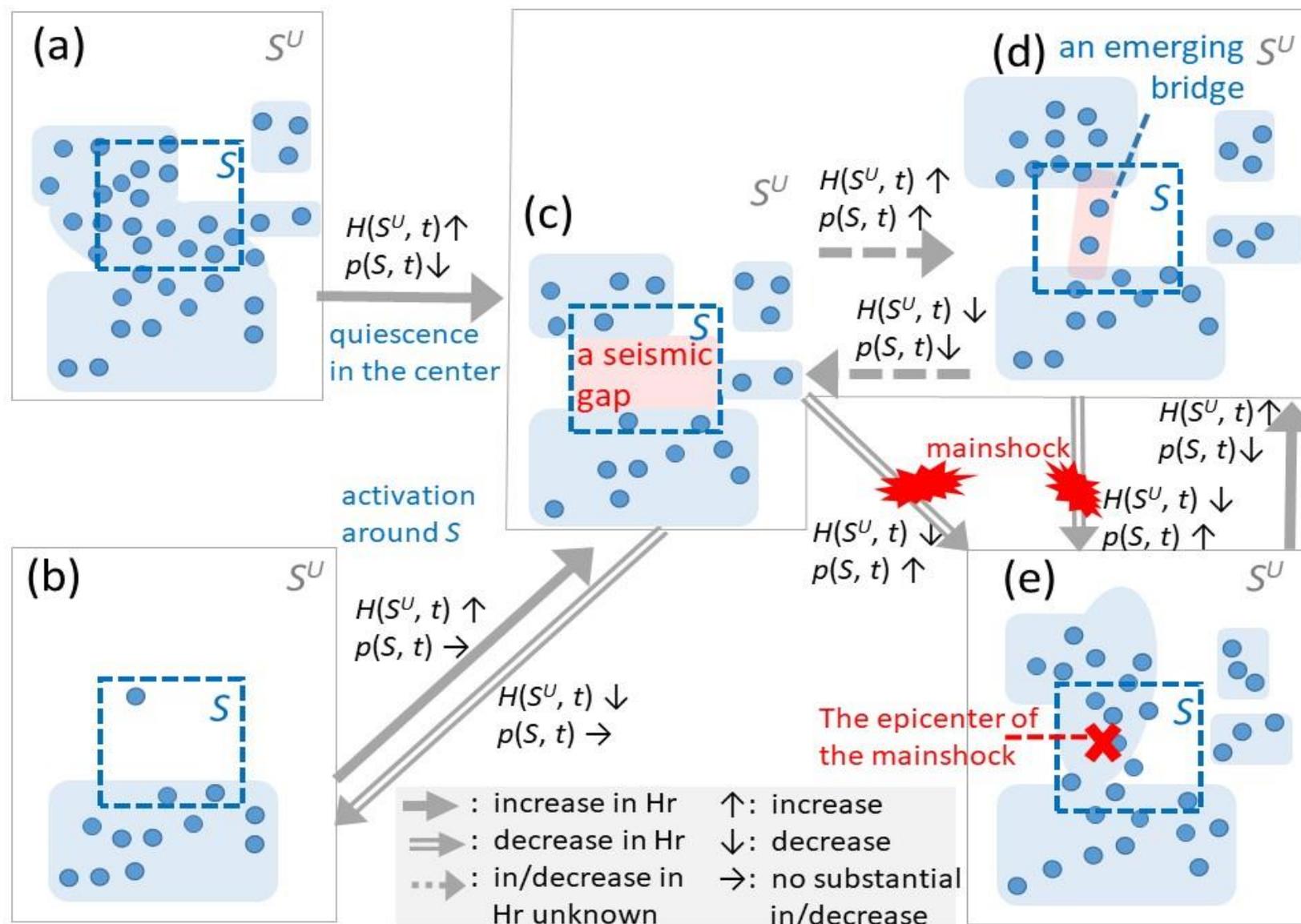
満たした要求

電子書籍の普及

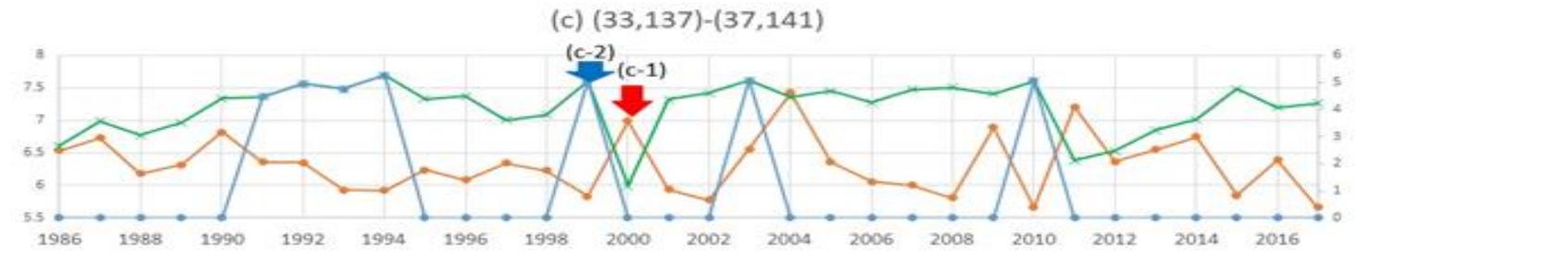
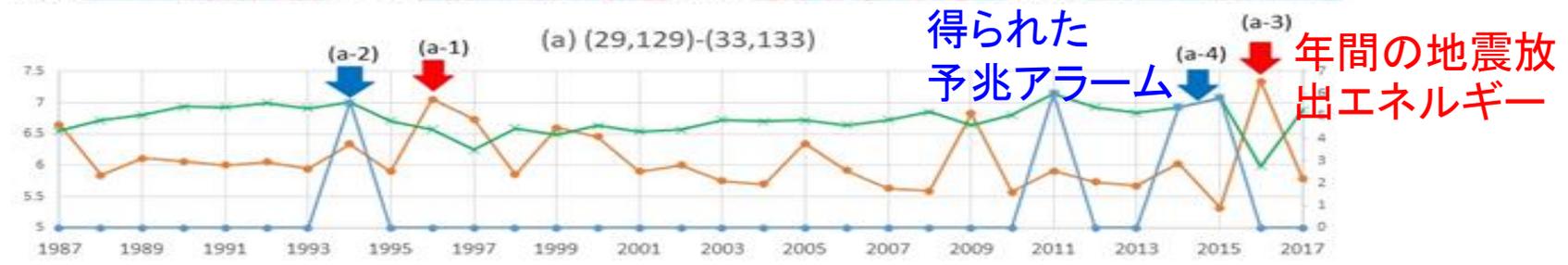
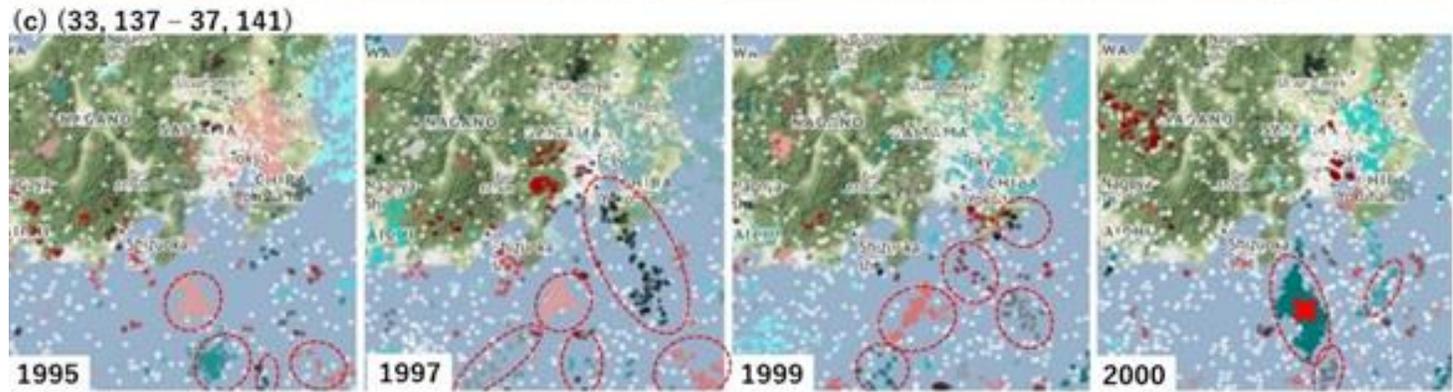
READ MORE

HOW TO USE

RESI = $-\sum_c p(C) \log p(C) - \log p(S)$ の変化で活性化をアラート
 ※ c はクラスタ, S は当該地域。先の GBE と比較してみてください



震央分布についての
クラスタ構造変化



Stay with your community: Bridges between clusters trigger expansion of COVID-19

Yukio Ohsawa, Masaharu Tsubokura

Published: December 3, 2020 • <https://doi.org/10.1371/journal.pone.0242766>

Article	Authors	Metrics	Comments	Media Coverage
Abstract				

Abstract

Abstract

そもそも、GoToとコロナ感染の関連性は前から指摘されてきた。「第三波」が来始めた十一月中旬、日本医師会の中川俊男会長はGoToについて「エビデンス（証拠）がなかなかはっきりしないが、感染拡大のきっかけになったのでは間違いない」と強調した。

十一月七日には、東京大学や米UCLAなどの研究者

学的思考



4日の記者会見でGoTo継続に強い意欲を示す菅首相＝首相官邸で

GoToトラベル → コロナ感染拡大

因果関係「証拠」

「専門家が行ったとは考えられない分析手法が含まれており、「因果関係は確認できず」という見解を出した科学的根拠としては認めがたい。そう語るのは、ビッグデータの解析やデータ市場設計などで知られる東京大の大沢幸生教授（知能情報学、やり玉に挙げたのは内閣官房と内閣府が連名で作成した資料だ。この資料は十一月十九日の厚生労働省の専門家会議で示された。GoToトラベルが始まった七月以降の



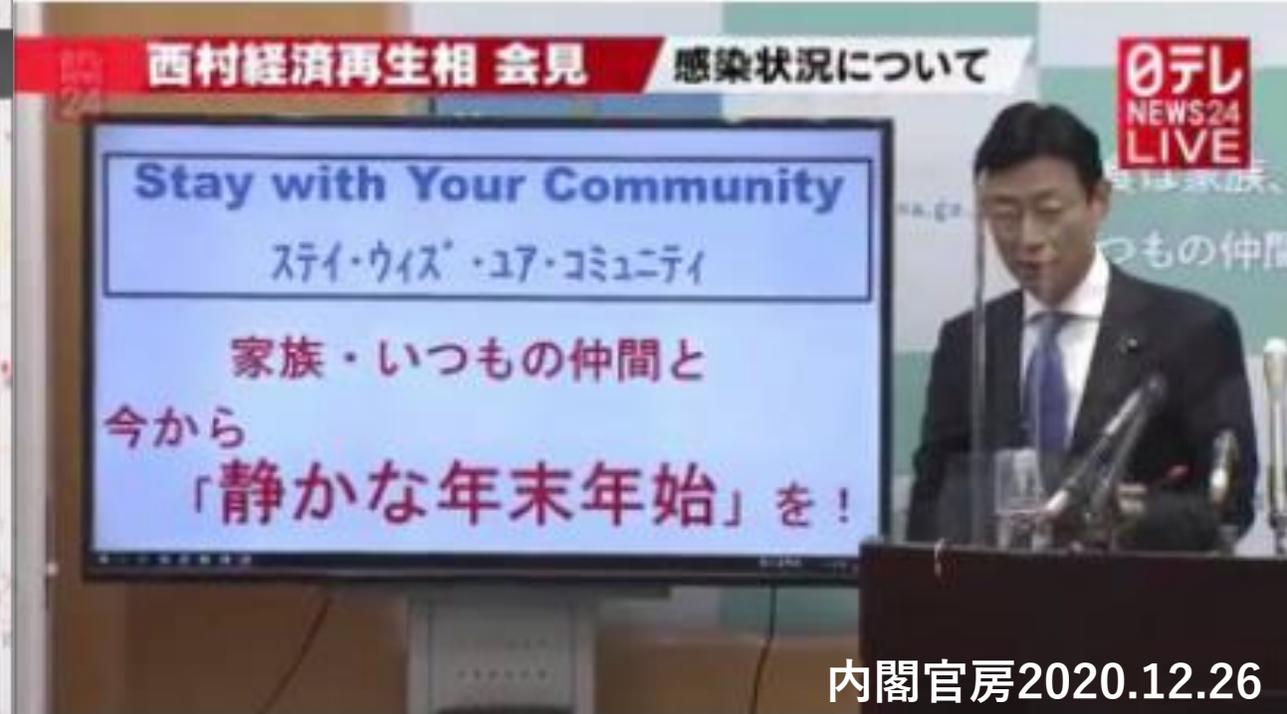
大沢幸生 東京大教授

拡大が止まらない新型コロナ第二波、逼迫した各地の医療機関が上がる。その要因として疑われるのが「GoToトラベル」。義偉内閣は「因果関係を示す証拠はない」として来年六月末まで針を決めた。だが、当の内閣官房からの資金で感染拡大モデルのいる研究者から、「証拠はある」という声が上がった。証拠とはその証拠に基づいた政策判断がされないのか。（橋原亮）石



東京新聞2020.12.13

ステイ・ウィズ・コミュニティです



西村経済再生相 会見

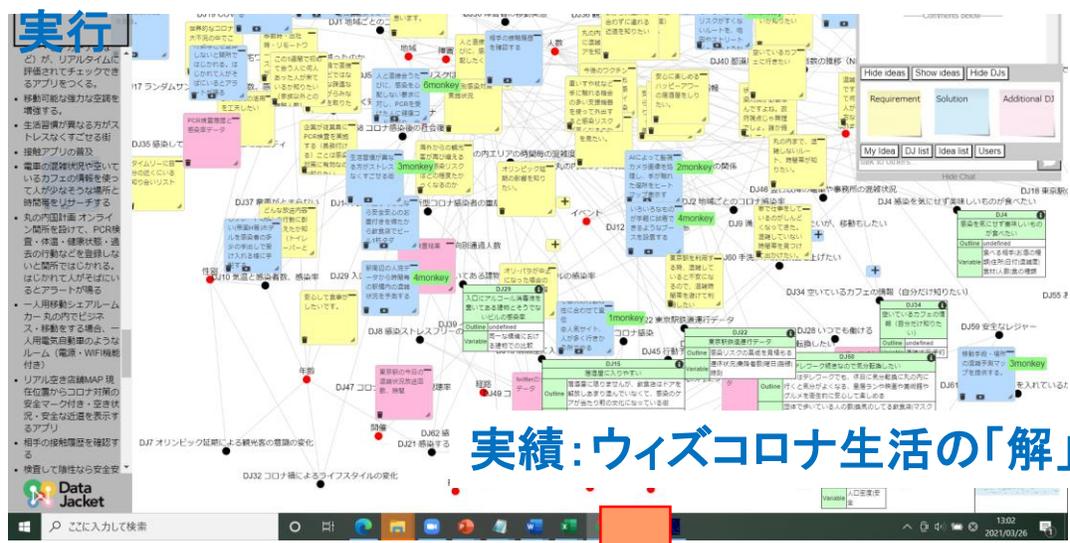
感染状況について

Stay with Your Community
ステイ・ウィズ・ユア・コミュニティ

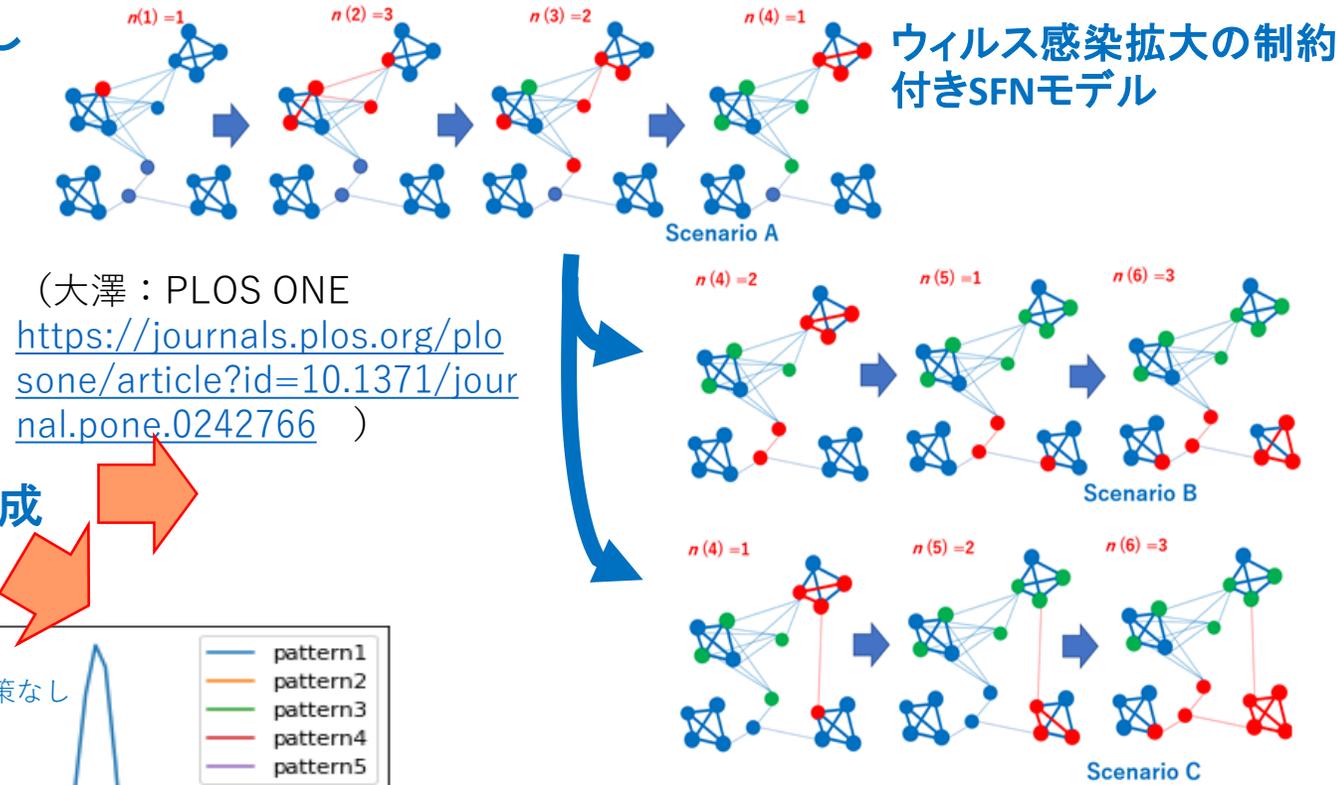
家族・いつもの仲間と
今から
「静かな年末年始」を！

内閣官房2020.12.26

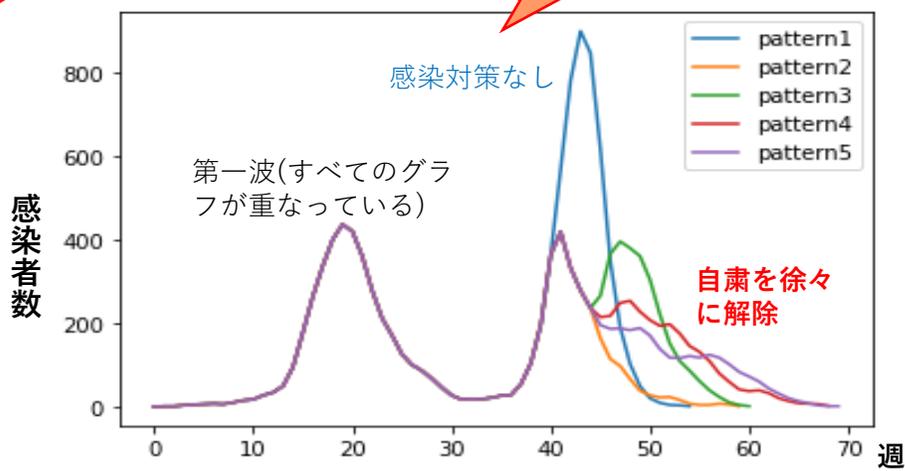
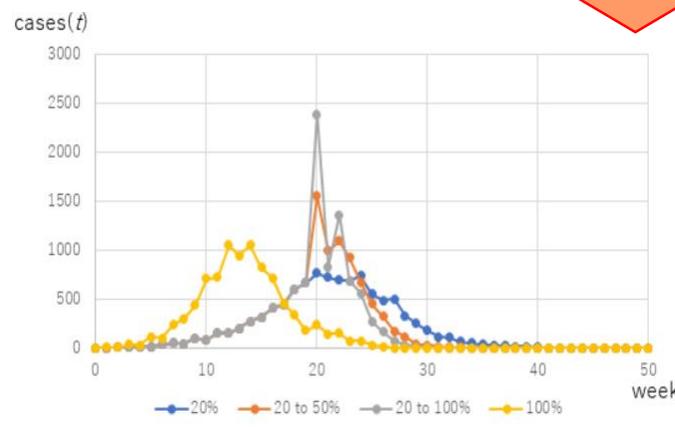
大澤研究室の方法: 地域・企業に民産官学の組織を構成し、データジャケット(大澤2013)データ活用戦略を立案し



実績: ウィズコロナ生活の「解」を生成



(大澤: PLOS ONE <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0242766>)



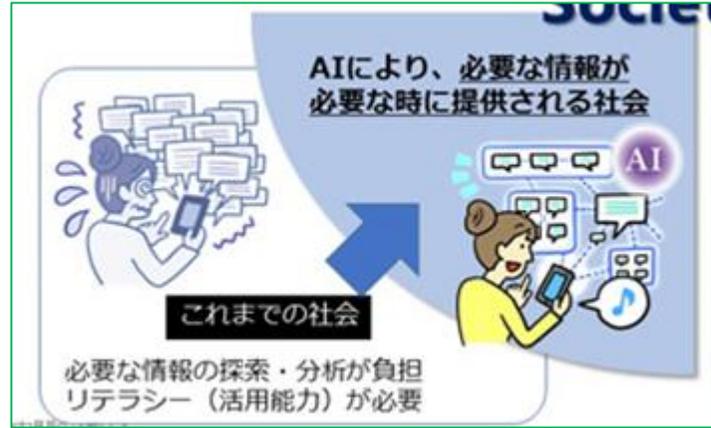
どこに集まると危ない? (NHK、フジTV等で周知)

	F20-39	M20-39	F40-59	M40-59	F60-79	M60-79	平均
学校	1.05	1.29	0.00	0.00	0.00	0.00	0.34
オフィス1	0.00	2.77	3.32	1.66	0.00	2.36	1.71
オフィス2	0.24	1.60	0.00	0.00	0.00	0.00	0.27
病院	0.26	3.64	1.43	1.26	0.81	2.23	1.58
ジム	0.00	4.20	1.62	0.15	3.58	0.00	1.59
飲食店	5.29	3.97	3.09	4.97	1.85	0.00	3.15
居酒屋	2.31	0.00	0.01	0.74	8.67	1.99	2.34
劇場	0.00	0.00	5.30	0.00	0.67	2.90	1.57
スーパー	0.84	5.99	5.18	8.41	5.20	3.94	5.11
デパート	0.22	2.07	1.17	0.20	0.00	2.60	1.01
テーマパーク	0.00	16.61	6.51	0.00	0.00	0.00	3.63
宿泊	0.00	0.00	0.00	1.20	12.89	6.34	3.57
電車	11.35	6.42	4.16	4.94	4.36	0.40	5.06
バス	6.03	0.00	0.45	5.80	0.00	4.80	2.79

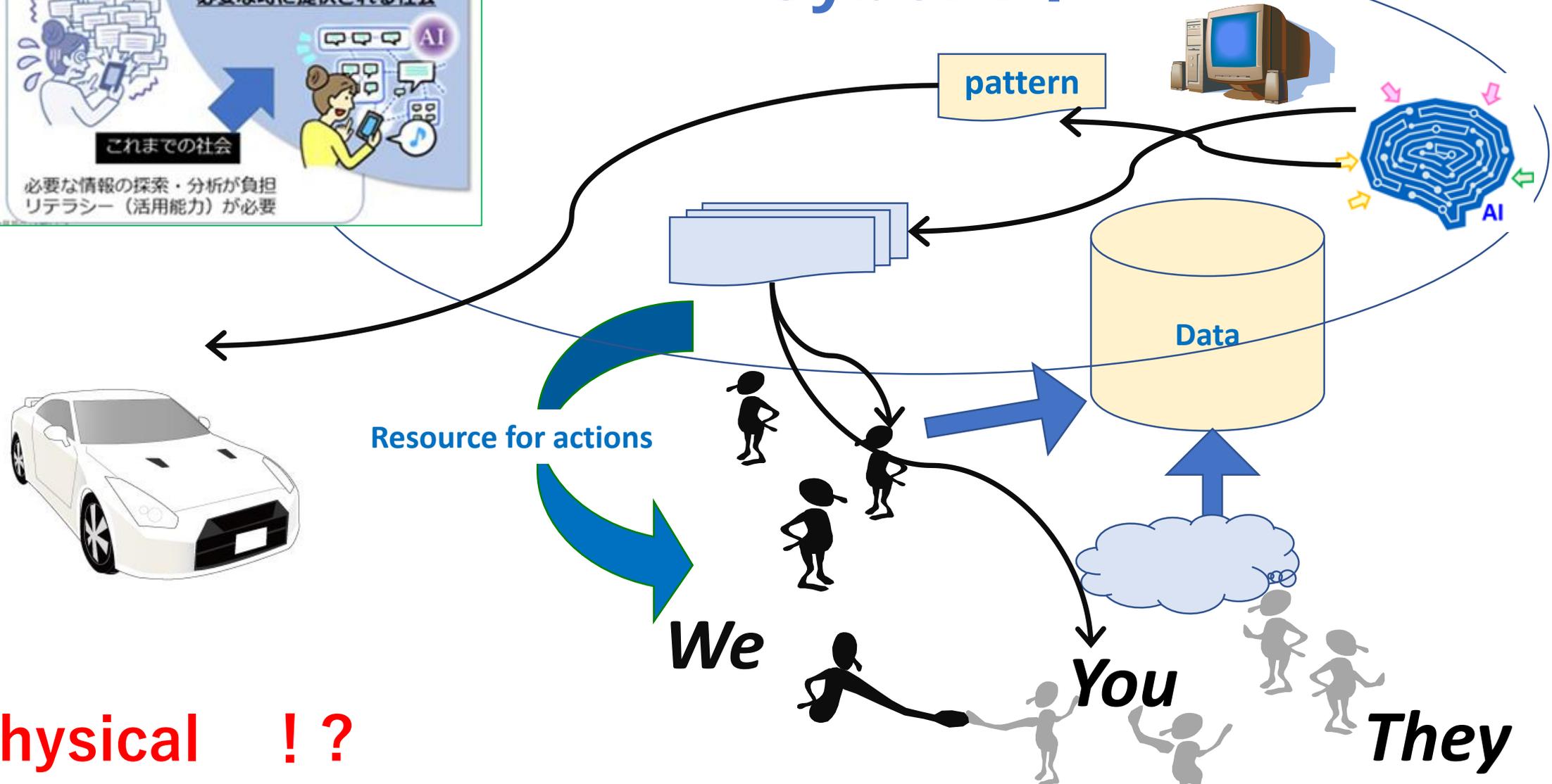
何故「緊急事態」後に大波?
波が終わった後で、ネットワーク構造の急変があると感染対策なしの場合よりも爆発的増加
※遠距離移動やワクチン接種戦略も内閣官房HPなどから発信

緊急事態からの戻し方 (元は w_{10} で、39週目まで自粛 w_1)
Ptrn1: 常に接触率 w_0 (対策無し)
Ptrn2: 上昇中 (第39週) に元に戻ってしまう場合
Ptrn3: 43週から w_0 に戻す場合
Ptrn4: 43週から友達1, 友達2, 知人, その他の順で w_0 に ↓
Ptrn5: さらに友達1との接触は w_0 に戻す前に一度中間値に

Society 5.0 (?)



Cyber ! ?

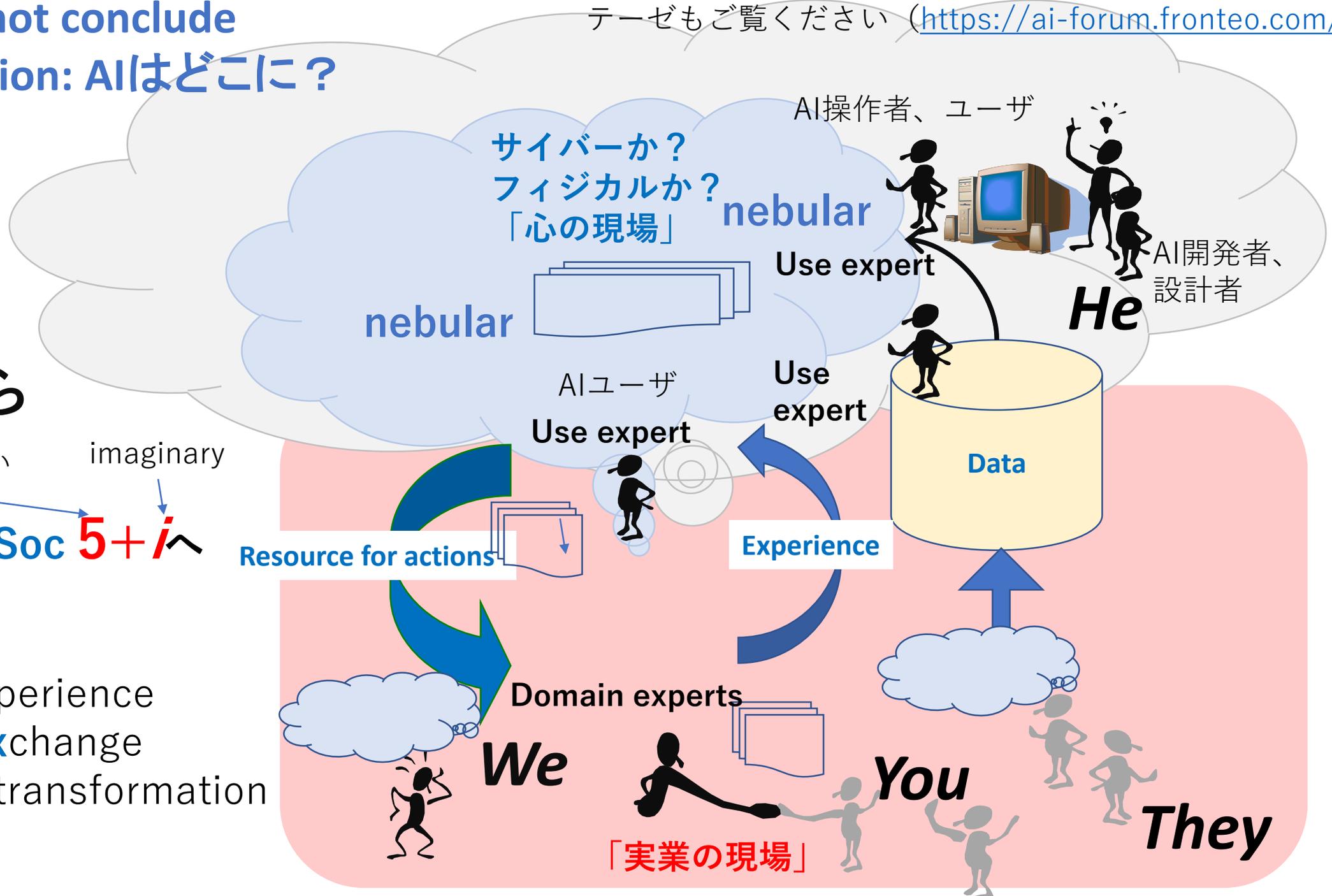


Physical ! ?

Let's not conclude

Question: AIはどこに？

テーゼもご覧ください (<https://ai-forum.fronteo.com/>)



めざすなら

小数点につけない

imaginary

3点セットでSoc 5+*i*へ

Resource for actions

Experience

DX Triple =

- × Dual Experience
- × Data Exchange
- × Digital transformation

We

You

They

「実業の現場」