

分散PDSを活用した 教育ビッグデータの蓄積・活用

2022-11-09 橋田 浩一



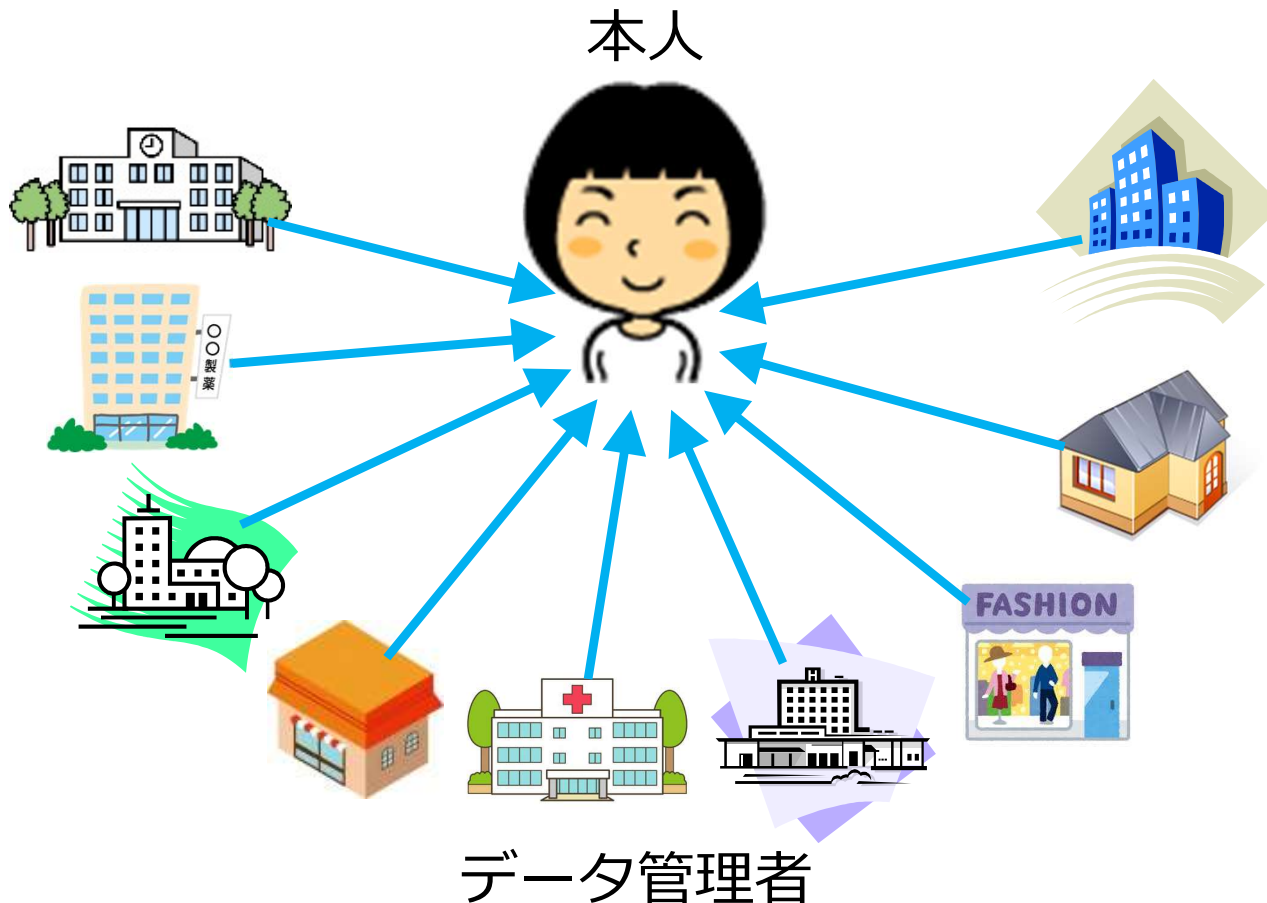
東京大学大学院情報理工学系研究科
ソーシャルICT研究センター



理化学研究所
革新知能統合研究センター

パーソナルデータ(PD)の分散管理

- データ管理者の間で共通のIDは不要
- 人権リスクなしにPDを名寄せして価値向上



PDの分散管理 → 価値の最大化

PDを本人に集約(名寄せ) ⇒ 利便性

1次利用: 本人に直接関わる利用

- 名寄せされたPDを本人のためにフル活用

2次利用: 多数の人々のPDの統計分析

- 名寄せされたPDを容易に収集・分析
- パーソナルAI (PAI)が同意を支援

他者による集中管理なし ⇒ 安全性

- 大規模な不正使用等なし

PDの分散管理(補足)

- ある種の公共目的等には集中管理が必須
 - ◆ 教育(選別など)、課税、防犯、公衆衛生、他
- 本人も他者もPDを正しい目的にしか使わない
 - ✗ 縁談の釣書の家系図など
 - ◆ データ最小化(data minimization)
 - ◆ 技術と制度・文化による制限

パーソナルAI (PAI)

PLR (Personal Life Repository)

オントロジーでデータを正規化

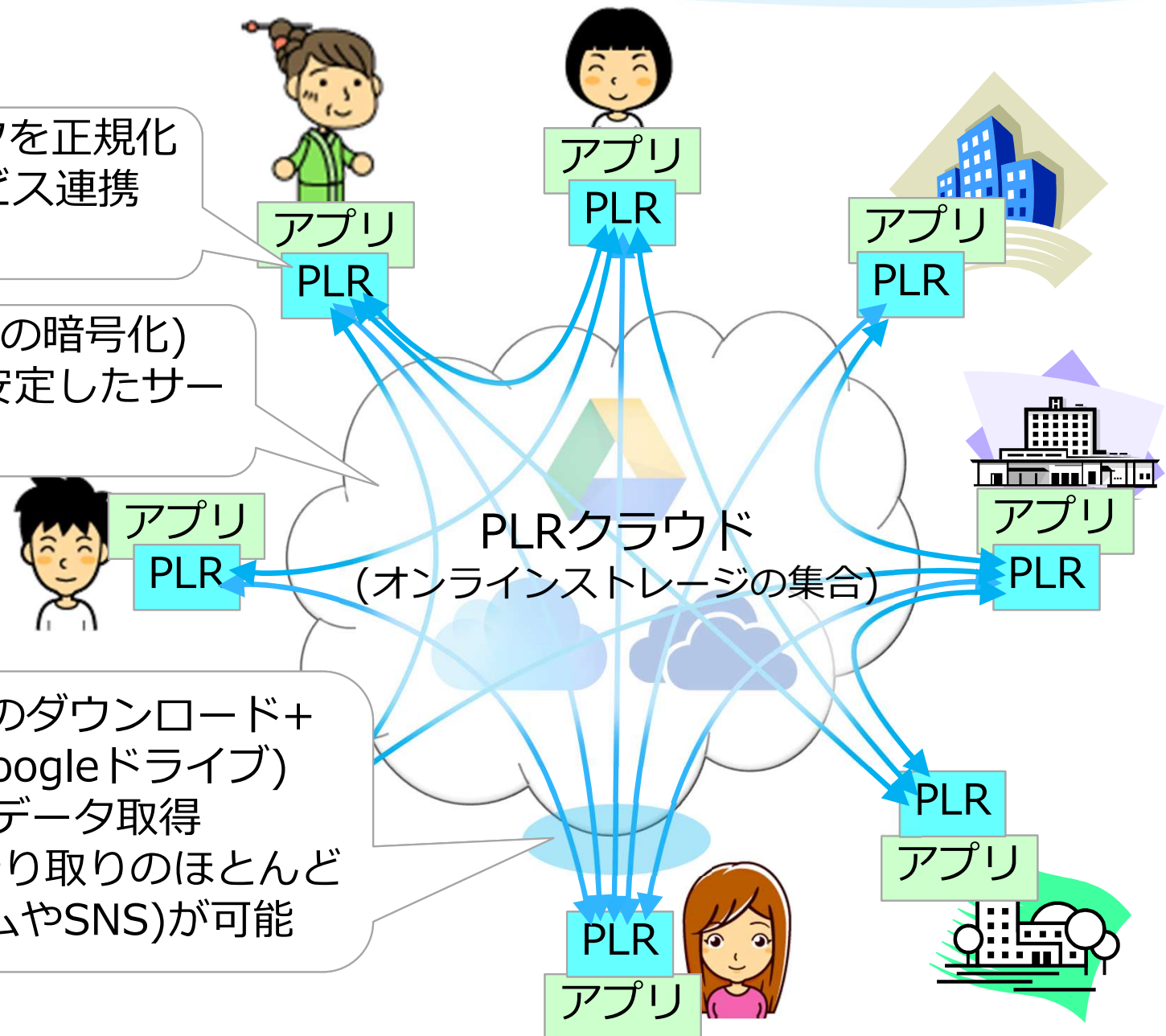
- データ連携・サービス連携が容易

- 分散管理(end2endの暗号化)

- 数十億の利用者に安定したサービスを安価に提供

2回/秒ちょっとのダウンロード+アップロード(Googleドライブ)

- 18万回/日のデータ取得
- 人間同士のやり取りのほとんど(業務システムやSNS)が可能





検索

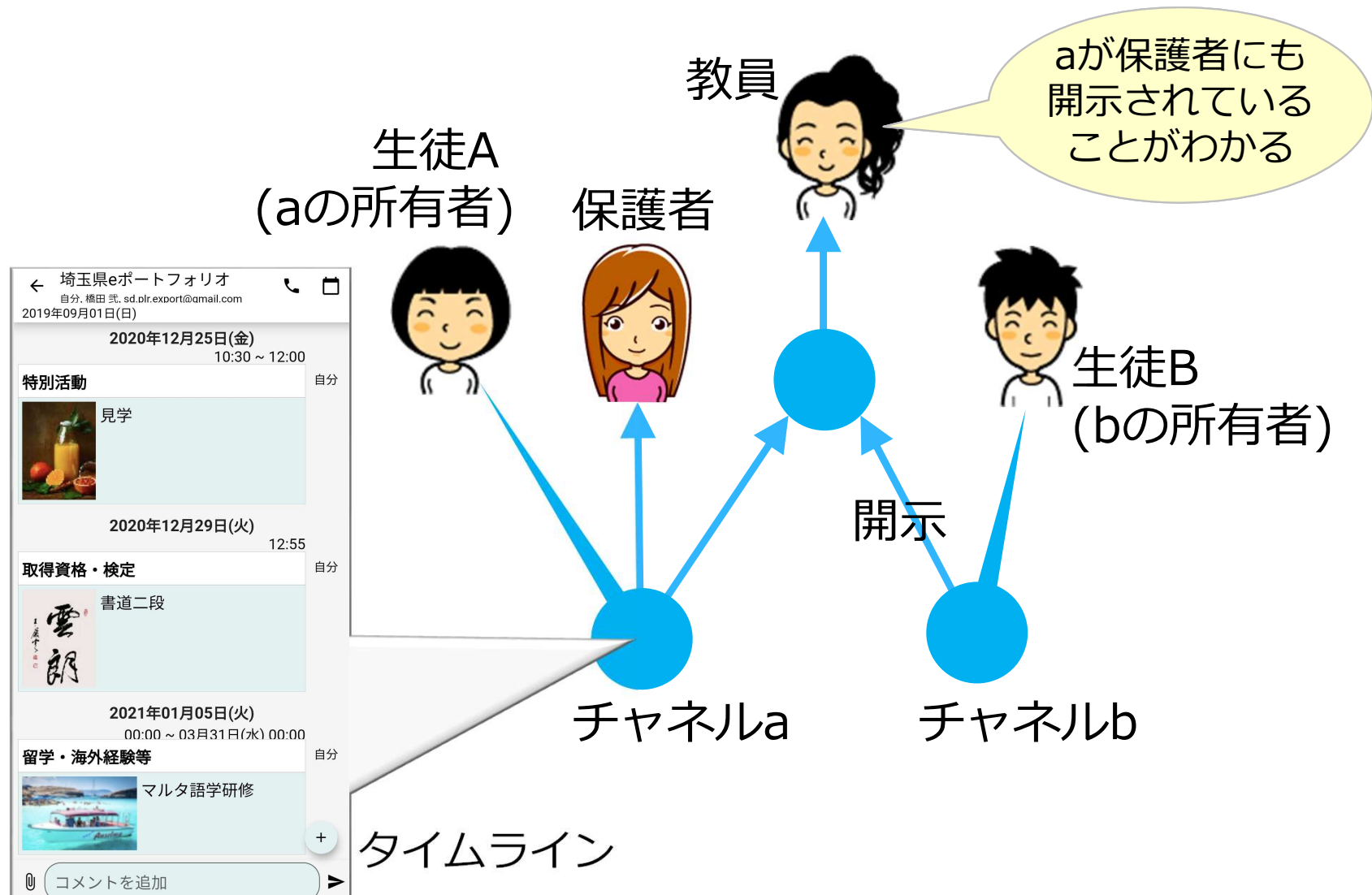
Personary (PLR標準アプリ)

- Android、iOS、デスクトップ
- ◆ Flutter/Dart
- データ作成・活用
- ◆ 生活録や問診
- ◆ 友達/グループ
- カスタマイズ
- ◆ オントロジー
- ◆ スタイルシート
- ◆ 他システムとの連携



チャンネル

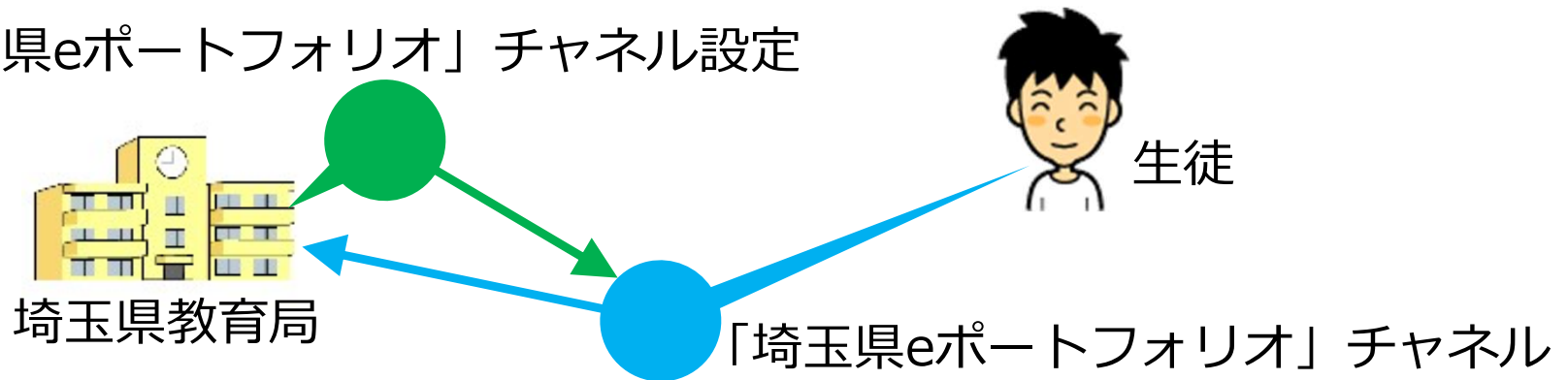
- PLR利用者(個人や事業者)の間で共有されるデータのまとまりの単位
- 各利用者は自分がアクセスできるチャンネルが他のどのチャンネルと利用者に開示されているかがわかる



チャンネル設定

- チャンネルの型の定義
 - ◆ オントロジーとスタイルシート
 - ◆ チャンネル設定の埋め込み
- 埼玉の県立高校の生徒が埼玉県教育局を友達登録すると自分の「埼玉県eポートフォリオ」チャンネルが自動生成されて埼玉県教育局に開示される

「埼玉県eポートフォリオ」チャンネル設定



チャンネルのカスタマイズ

0:42

< 種別: COVID-19スクリーニング

報告者: 自分

開始日時 2020.10.17 01:54 終了日時

体温 35.7°C

呼吸 25回/分

脈拍 68回/分

経皮的酸素飽和度 94.0%

息切れ(Borg) 0.5

24時間以内の咳 頻繁にあった

倦怠感 ない

咽頭痛 頻繁にある

頭痛 ときどきある

胸痛 ほとんどいつもある

腹痛 頻繁にある

24時間以内のくしゃみ 頻繁

24時間以内の鼻水 ほとんど

味覚 少しだけ感じる

嗅覚 感じない

24時間以内の下痢 2回

睡眠 まあまあ眠れた

キャンセル × OK ✓

新規保存 ↓

1:34

< COVID-19テレナーシング

自分, 橋田 浩一 UT

2021年06月29日(火) 一般

2021年06月29日(火)

01:31 自分

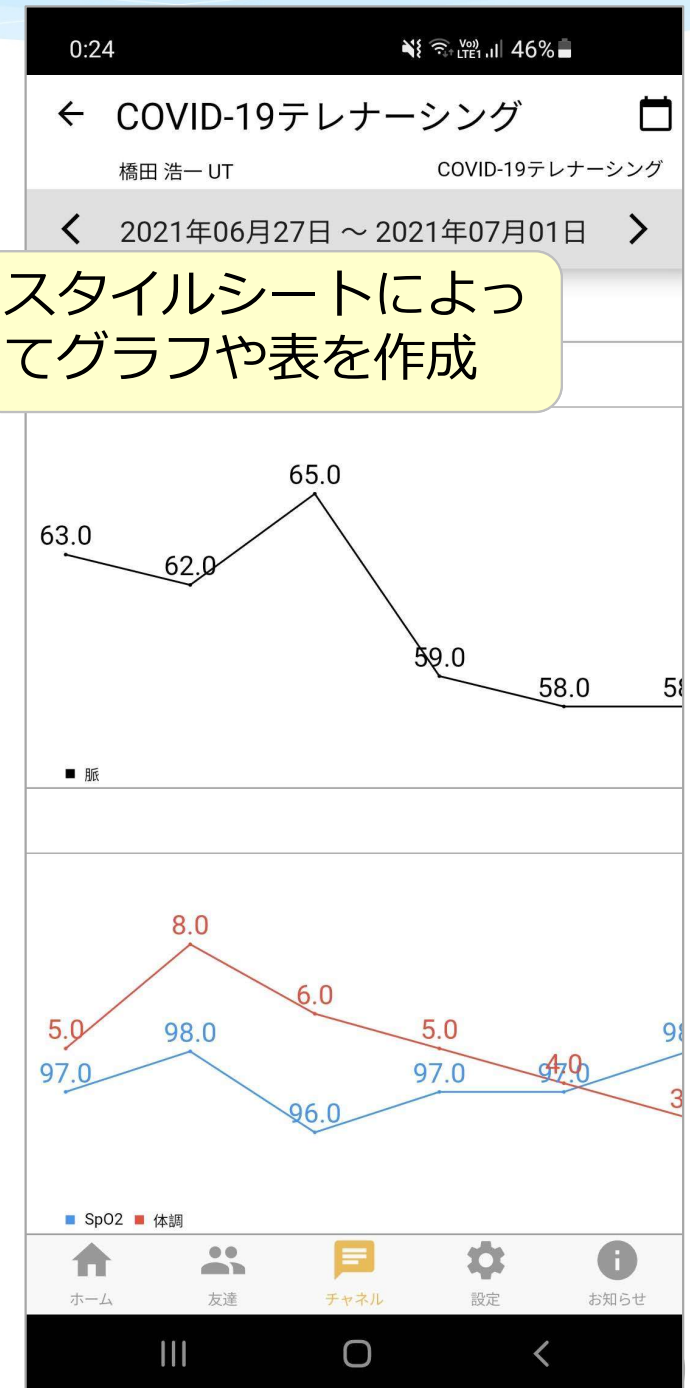
初期登録 身長168.0cm 体重52.0kg BMI21.0 先発医薬品アレセンサカプセル150mg 先発医薬品アバスチン点滴静注用400mg/16mL 喫煙歴

01:32 自分

COVID-19スクリーニング 体温36.0°C 呼吸21回/分 脈拍61回/分 酸素飽和度97.0% Borg7 咳ときどきあった 倦怠感頻繁にある 咽頭痛ほとんどない 頭痛ほとんどいつもある 胸痛頻繁にある 腹痛ときどきある くしゃみほとんどなかった 鼻水ほとんどいつもあった 味ほとんど感じない 嗅いいつも通り 下痢1回 まあまあ眠れた 体調3 宿泊療養中

コメントを追加

ホーム 友達 チャンネル 設定 通知



オントロジーでデータ形式を定義するとUIが自動生成される

設定ファイル(オントロジー)

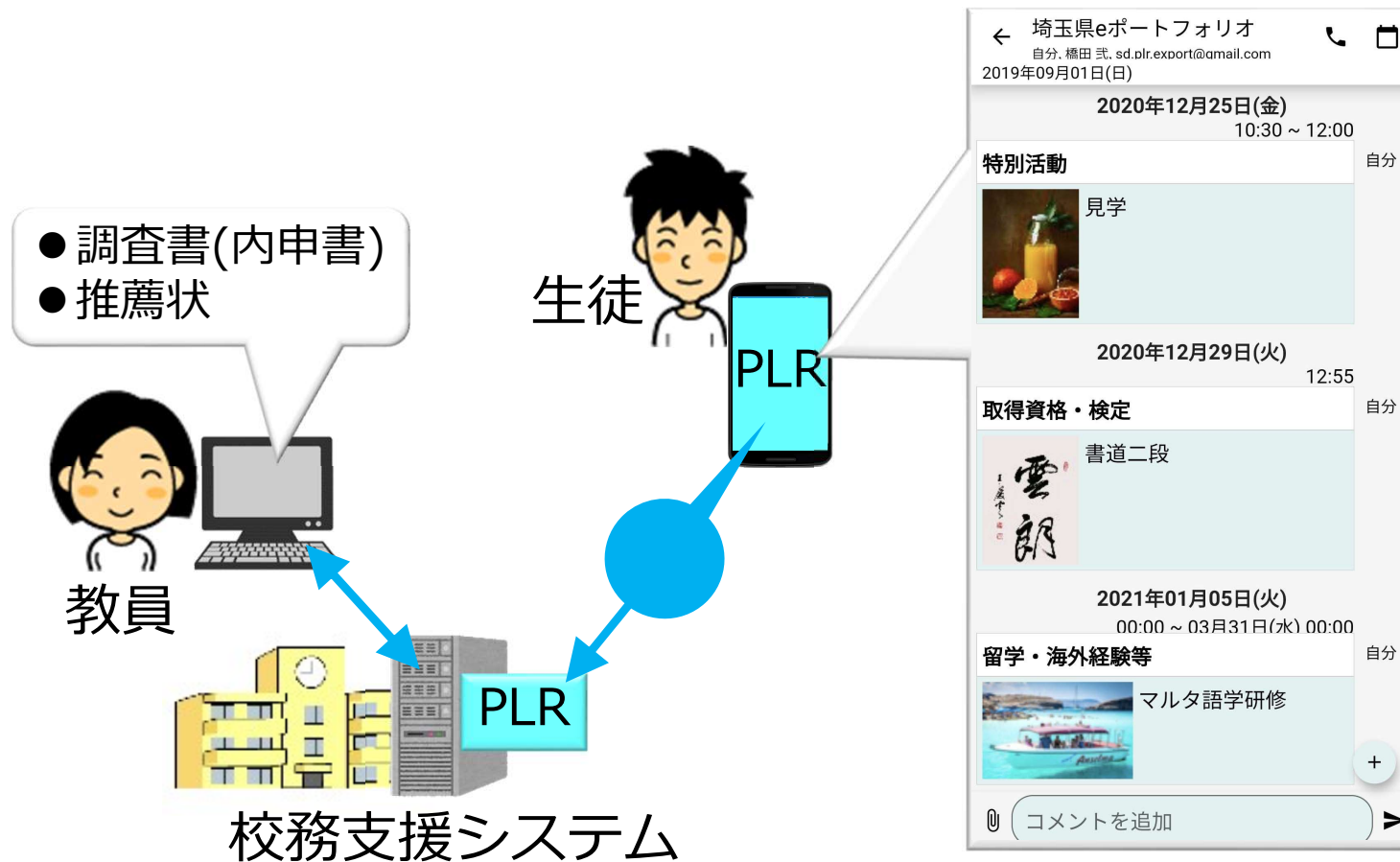
Event

TNSS0 スクリーニング

tnss0=1 ja:体温;body temperature 体温 []°C	decimal)25<50	
tnss1=1 呼吸 呼吸 []回/分	integer)10(150	
tnss2=1 脈拍 脈拍計 回/分	integer)30(150	
tnss3)0(1経皮的酸素飽和度 酸素飽和度 []%	decimal)0(100	
	_tnss410 10 非常に強い	
	_tnss49 9	
	_tnss48 8	
	_tnss47 7 とても強い	
	_tnss46 6	
	_tnss45 5 強い	
	_tnss44 4 多少強い	
	_tnss43 3	
	_tnss42 2 弱い	
	_tnss41 1 やや弱い	
	_tnss405 0.5 非常に弱い	
	_tnss40 0 感じない	
tnss4=1 ja:息苦しさ(Borg);Short Breath Borg		
tnss5a=1 24時間以内の咳 咳 空咳と湿性咳	_tnss5 咳など	_tnss54 なかった
		_tnss53 ほとんどなかった
		_tnss52 ときどきあった
		_tnss51 頻繁にあった
		_tnss50 ほとんどいつもあった
tnss5j=1 痰 痰	_tnssA 普段より多い	
	_tnssB 普段と同じ	
	_tnssC 出ない	
tnss5k=1 日常生活動作(昨日と比べて)	_tnssD 支障あり	
	_tnssE 動きづらい	
	_tnssF 可能	
tnss5b=1 倦怠感 倦怠感	_tnss6 痛みなど	_tnss64 ない
		_tnss63 ほとんどない
		_tnss62 ときどきある
		_tnss61 頻繁にある
		_tnss60 ほとんどいつもある
tnss5c=1 咽頭痛 咽頭痛	_tnss6	
tnss5d=1 頭痛 頭痛	_tnss6	
tnss5d=1 腰痛 腰痛	_tnss6	
tnss5i=1 関節痛 関節痛	_tnss6	
tnss5e=1 胸痛 胸痛	_tnss6	
tnss5f=1 腹痛 腹痛	_tnss6	
tnss5g=1 ja:24時間以内のくしゃみ;sneeze in 24h くしゃみ	_tnss5	
tnss5h=1 ja:24時間以内の鼻水;runny nose in 24h 鼻水	_tnss5	
	_tnssD0 感じない	
	_tnssD1 ほとんど感じない	
	_tnssD2 あまり感じない	
	_tnssD3 少しだけ感じる	
	_tnssD4 いつも通り	
tnssD=1 味覚 味	_tnssD0 感じない	
	_tnssD1 ほとんど感じない	

分散eポートフォリオ

- 埼玉県教育庁が2020年度から実運用
 - ◆ 生徒がPersonaryで入力した課外活動のデータを教員が調査書や推薦状の作成に活用
 - ◆ 最悪のUX: データ活用が本人に見えない

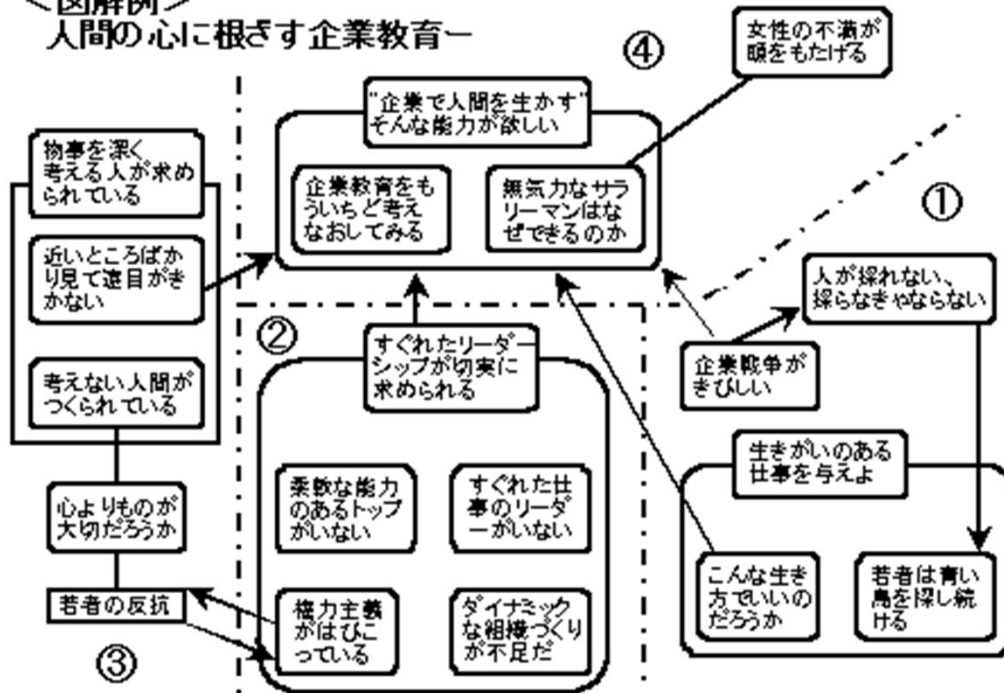


グラフによる意味の表現

KJ法A型図式

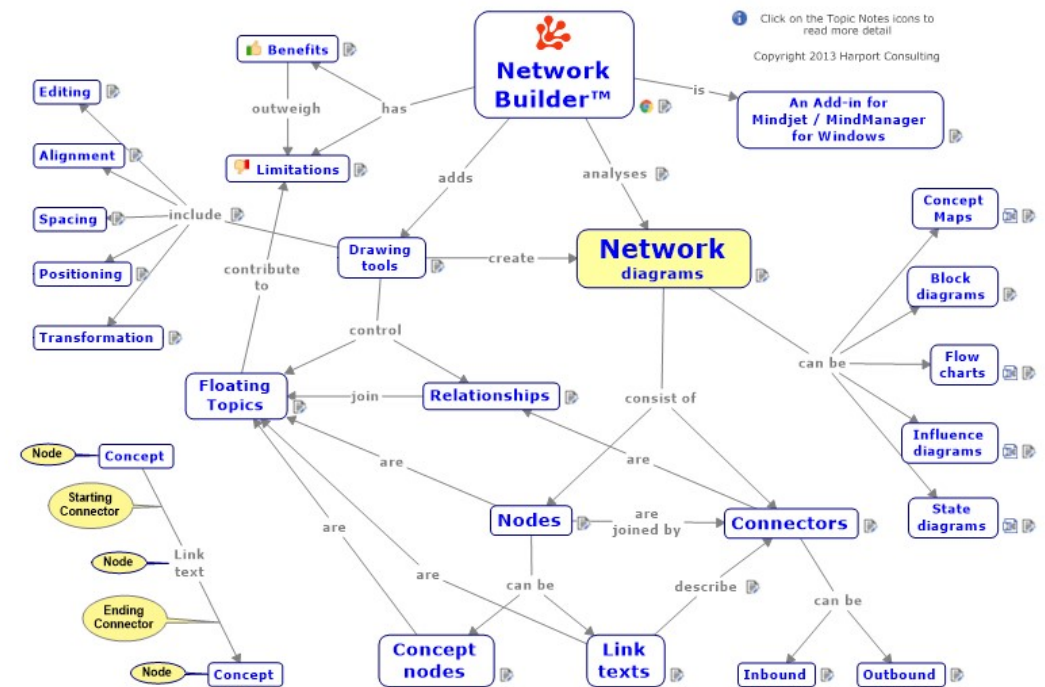
<図解例>

人間の心に根ざす企業教育—



関係を明示しないので意味を明確に表示できない

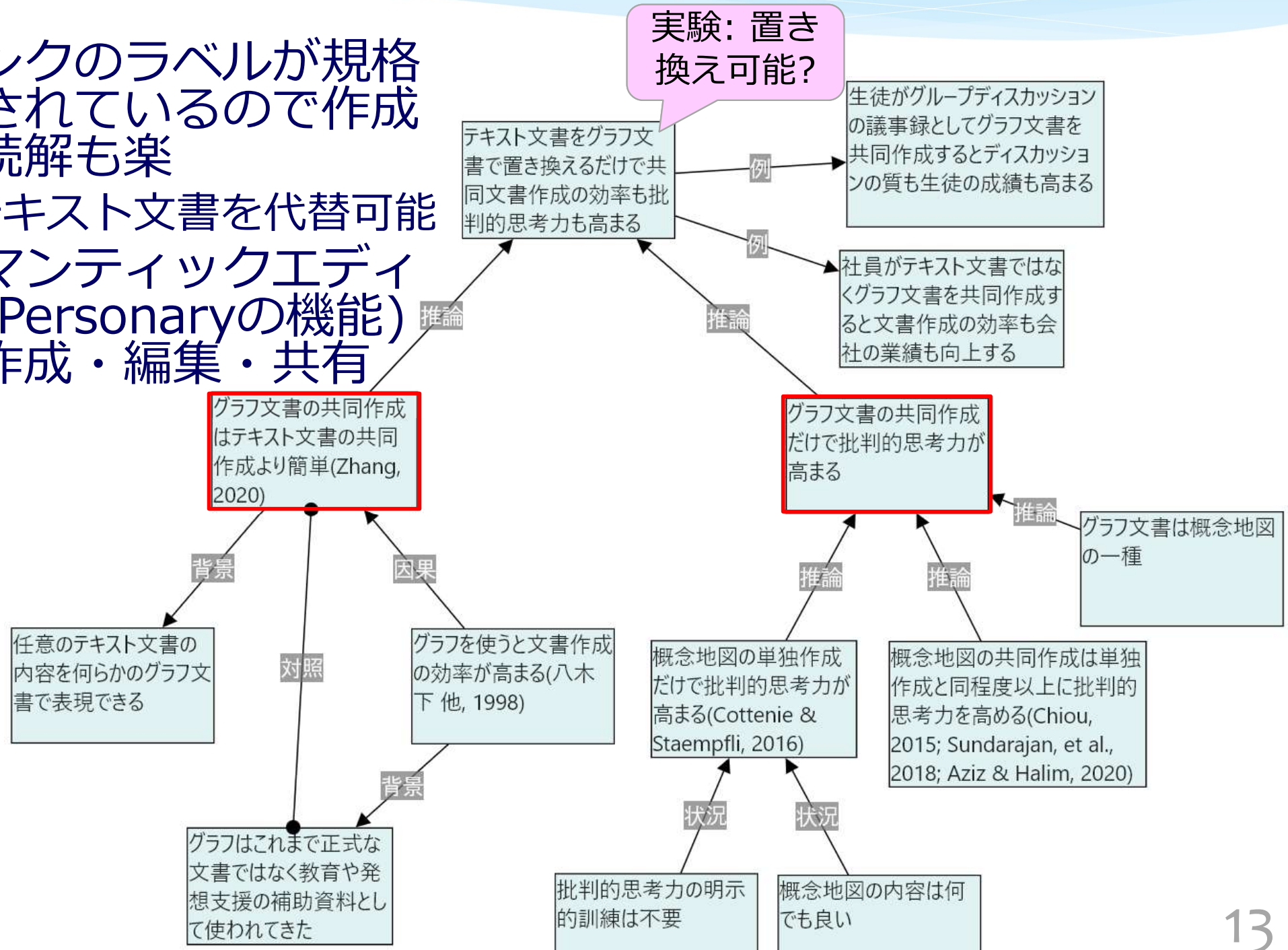
概念地図(Concept Map)



関係をテキスト入力するので作成が面倒

グラフ文書

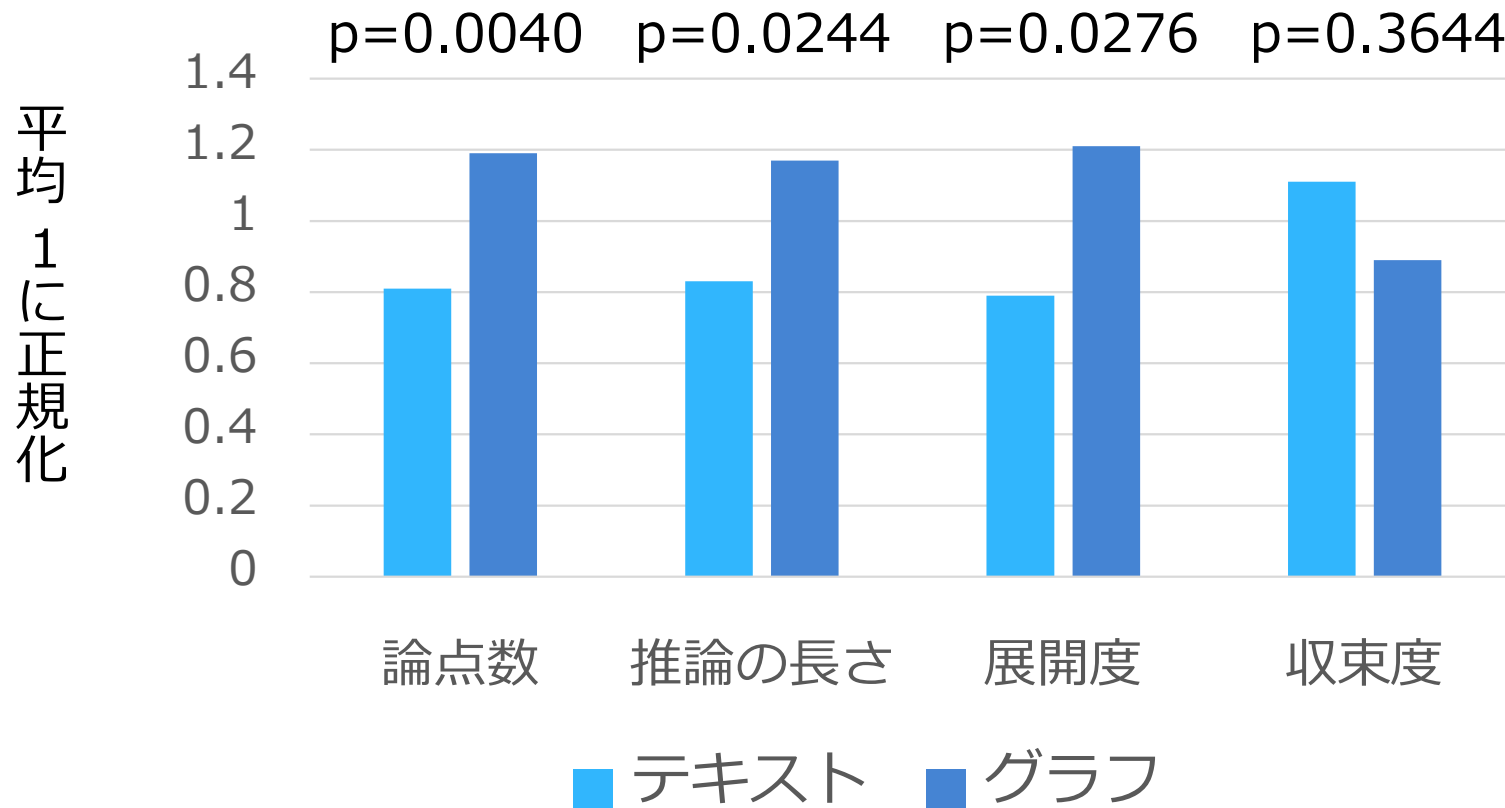
- リンクのラベルが規格化されているので作成も読解も楽
- ◆ テキスト文書を代替可能
- セマンティックエディタ(Personaryの機能)で作成・編集・共有



同期的共同作業の効率(Zhang, 2020)

- 2人組9グループ+タスク9種

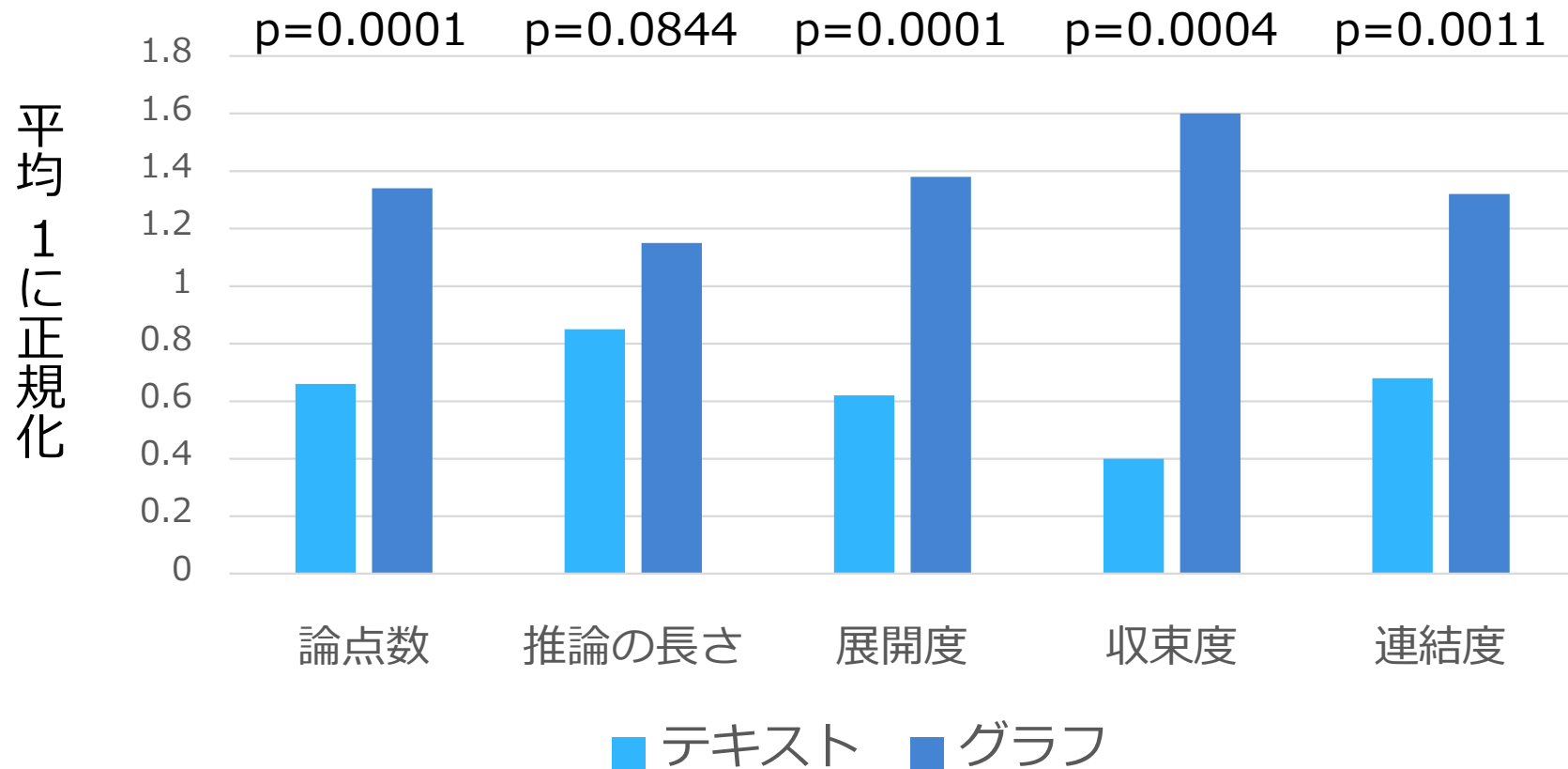
- ◆ 各グループの各セッションで2人が対面で会話しながら各々のPCで共有テキスト文書または共有グラフ文書を編集
- ◆ 9種のタスクの各々について、1グループがテキスト文書、別の1グループがグラフ文書を作成



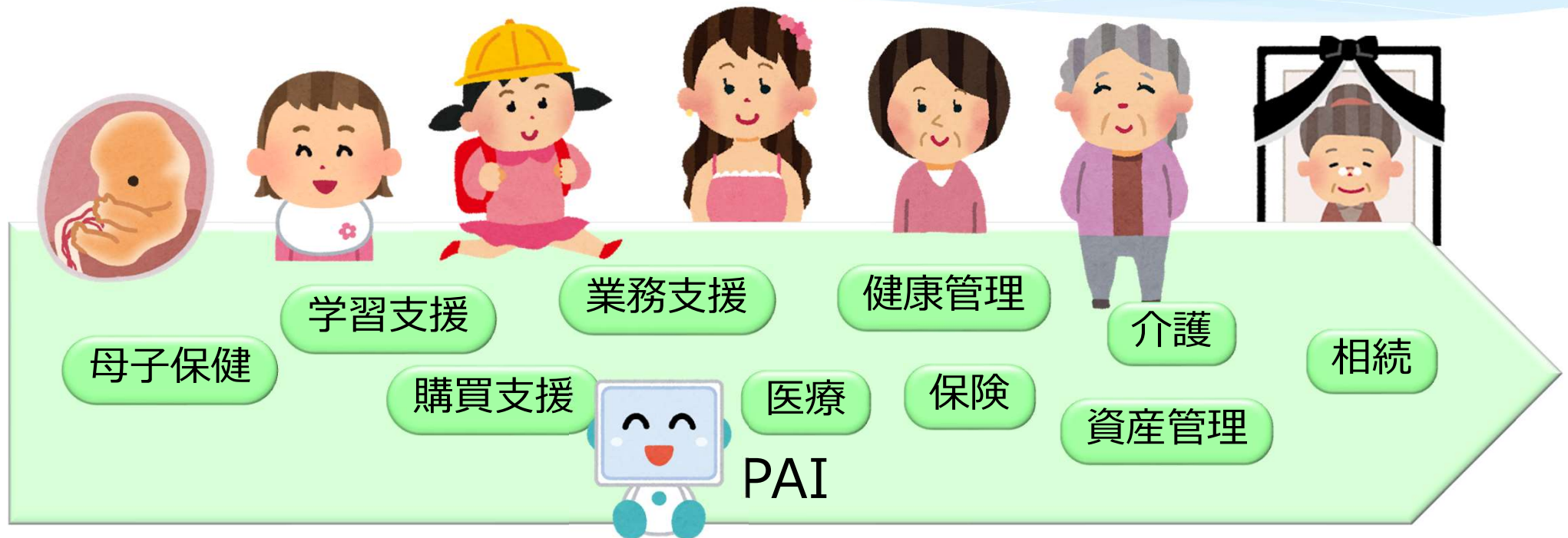
非同期的共同作業の効率(Zhang, 2020)

- 実験参加者10人+タスク10種

- ◆ 各タスクにつきテキスト文書とグラフ文書を1つずつ作成
- ◆ 各文書はまず1人の参加者が半ば作成した後にもう1人の参加者が完成させる
- ◆ 文書を共有する以外のコミュニケーションなし



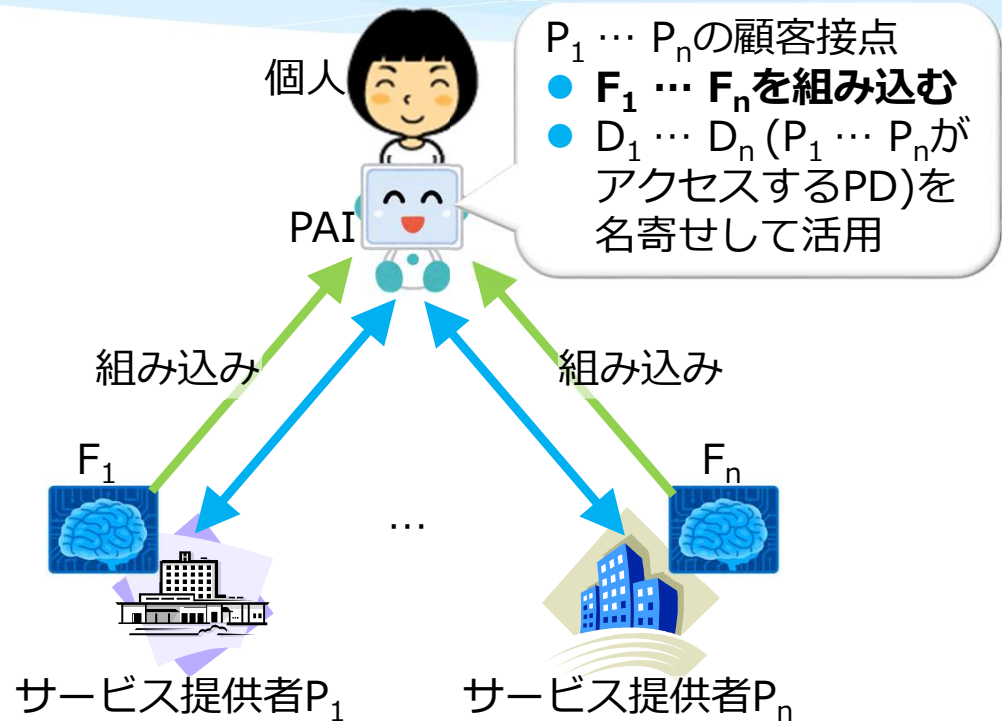
パーソナルAI (PAI)



- 各個人に専属しパーソナルデータ(PD)の管理を代行
 - ◆ 原則としてPDを他者に開示せず
 - ◆ デジタルツインの管理運用
- **PDをフル活用**して本人に深くキメ細かく介入
 - ◆ 意思決定と行動(変容)を支援
 - * 商品・サービス・職業等の選択、生活習慣の改善、ヘルスリテラシーの向上、PD開示、...
- 中央集権AI (CAI)をはるかに上回る付加価値: GDPの25%
- リスクもCAIより大きいので厳格なガバナンスが必要

付加価値: PAI >> CAI

- サービス提供者 $P_1 \dots P_n$ がデジタルな顧客接点としてCAI ($F_1 \dots F_n$)の代わりにPAIを共用
- 価値が n に関して非線型に増大
- P_i は F_j と D_j にアクセスしない($i \neq j$)



	F_1	...	F_n	他のAI	PAIの機能
D_1	P_1 によるサービス		$D_1 + F_n$ によるサービス		採寸データを使って健康管理
$\vdots$		$\dots$			
D_n	$D_n + F_1$ によるサービス		P_n によるサービス		業務のデータを使って健康管理
他のPD					生活習慣のデータを使って学習指導

PAIが名寄せして活用するPD PAIによる最大 $(n+1)^2$ 種類のサービス

n=2

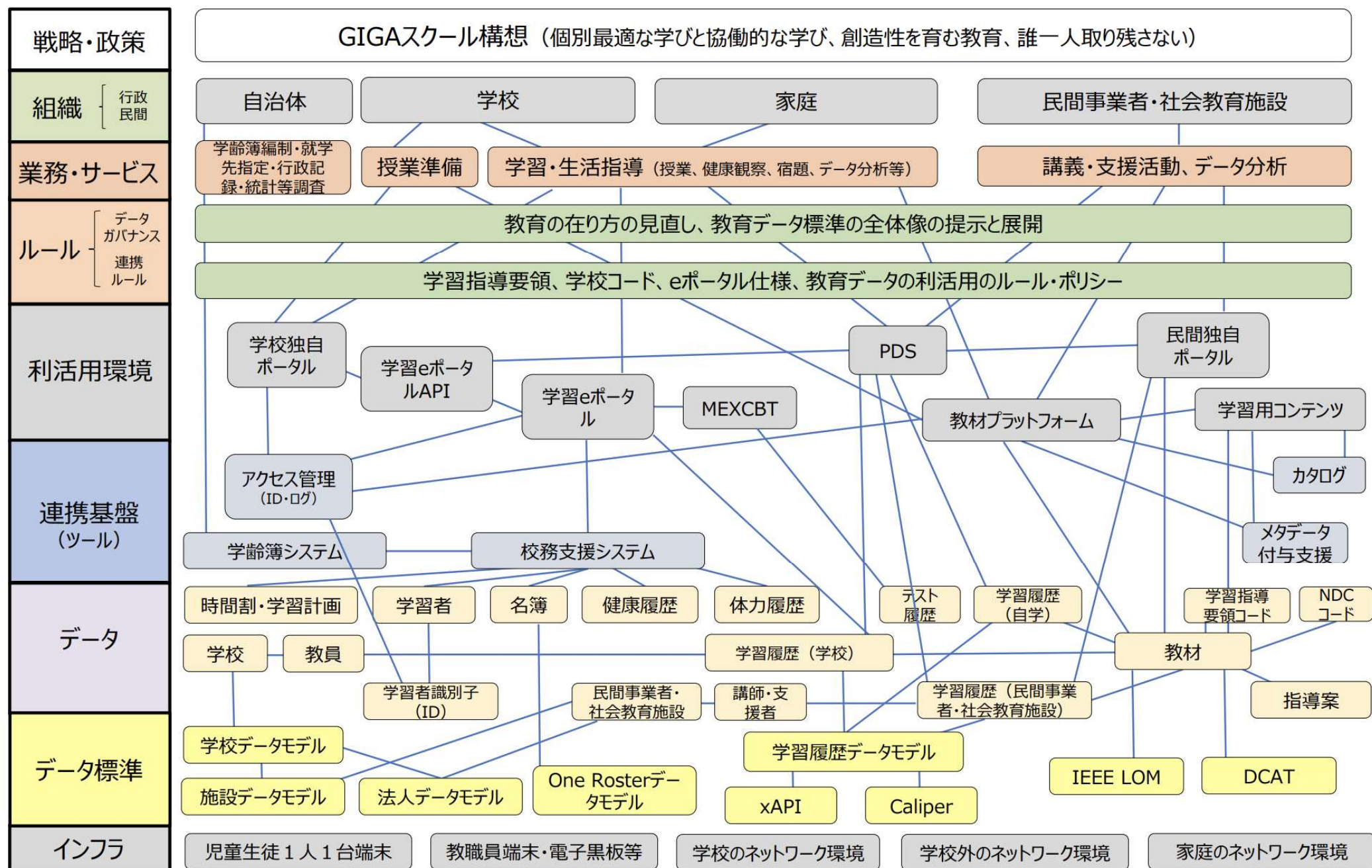
- 買い物や食事や運動などの生活行動を支援
 - ◆ cf. Amazonあんしんメール



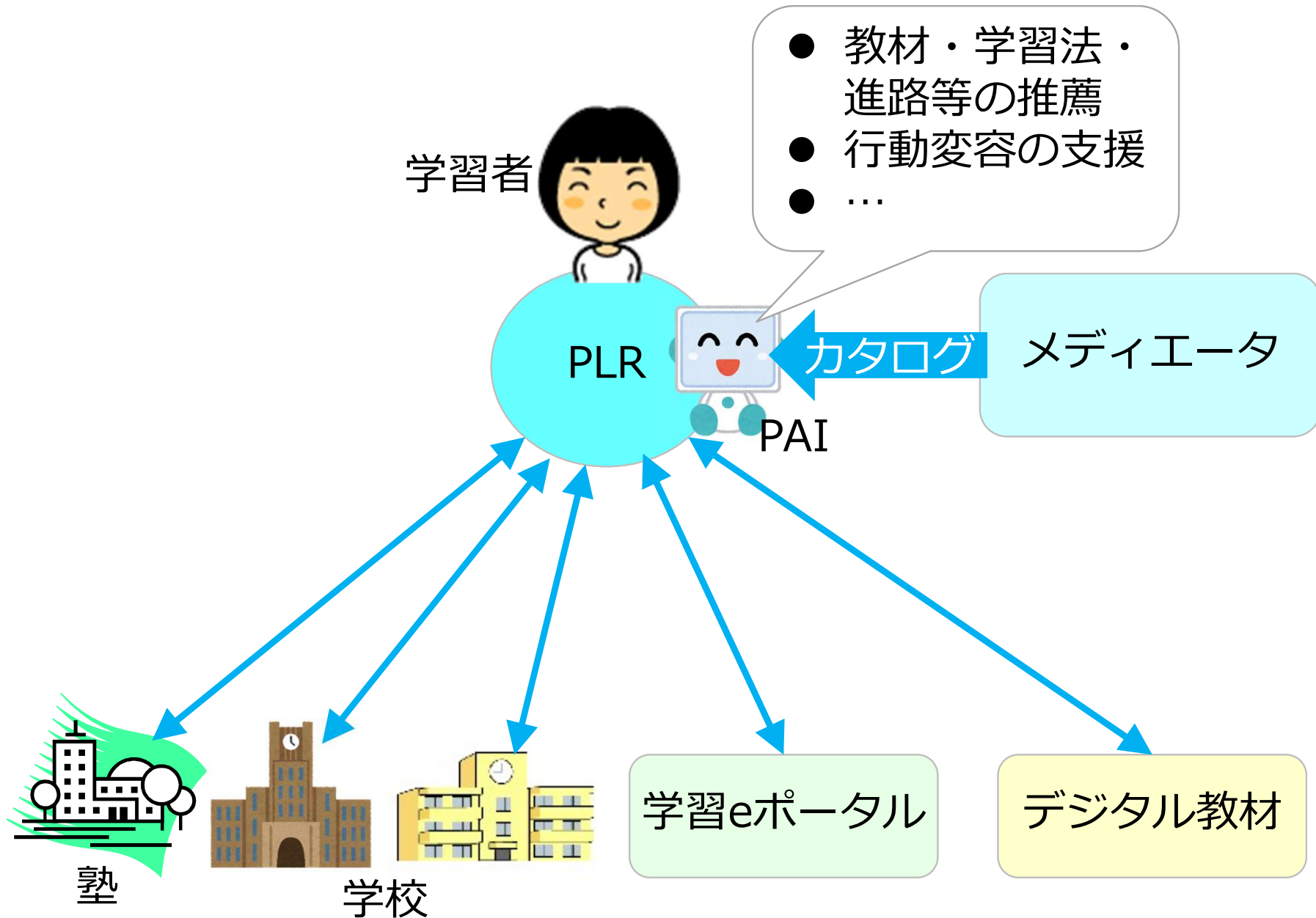
事業者がPAIの普及を主導する

- 一人勝ち
 - ◆ データの囲い込みによる顧客の囲い込み
- 協調
 - ◆ データの「共有」による顧客の共有
- 企業は可能なら一人勝ちしたいと考えがちだが実際には協調し合うことも多い
 - ◆ カルテルやトラスト等の協調は違法
 - ◆ 顧客接点としてのPAIの共有は適法
- CAIからPAIへの移行により収益も評判も向上

教育データの蓄積と流通の将来イメージ（アーキテクチャ：初中教育）

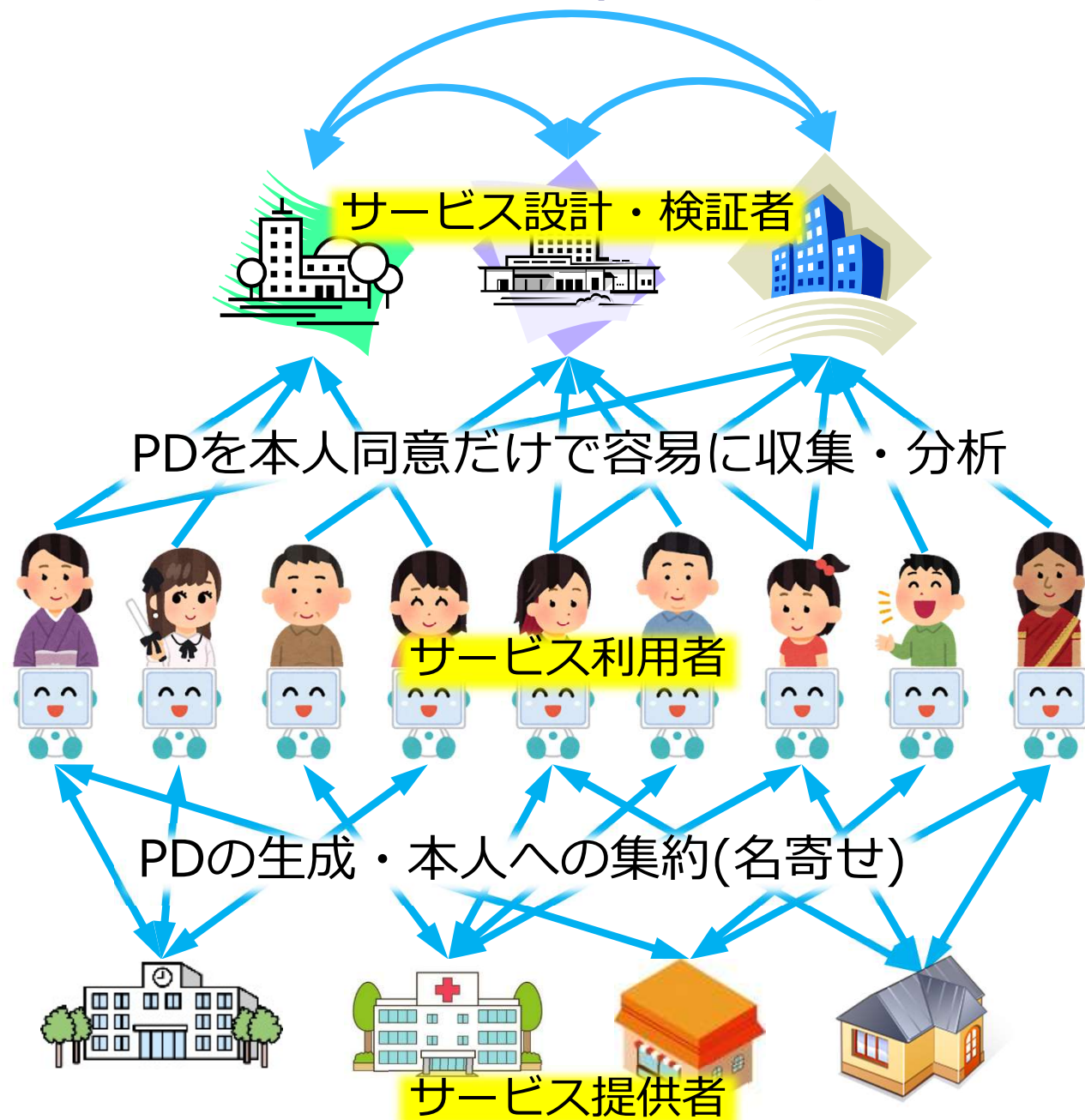


教育・学習の場合

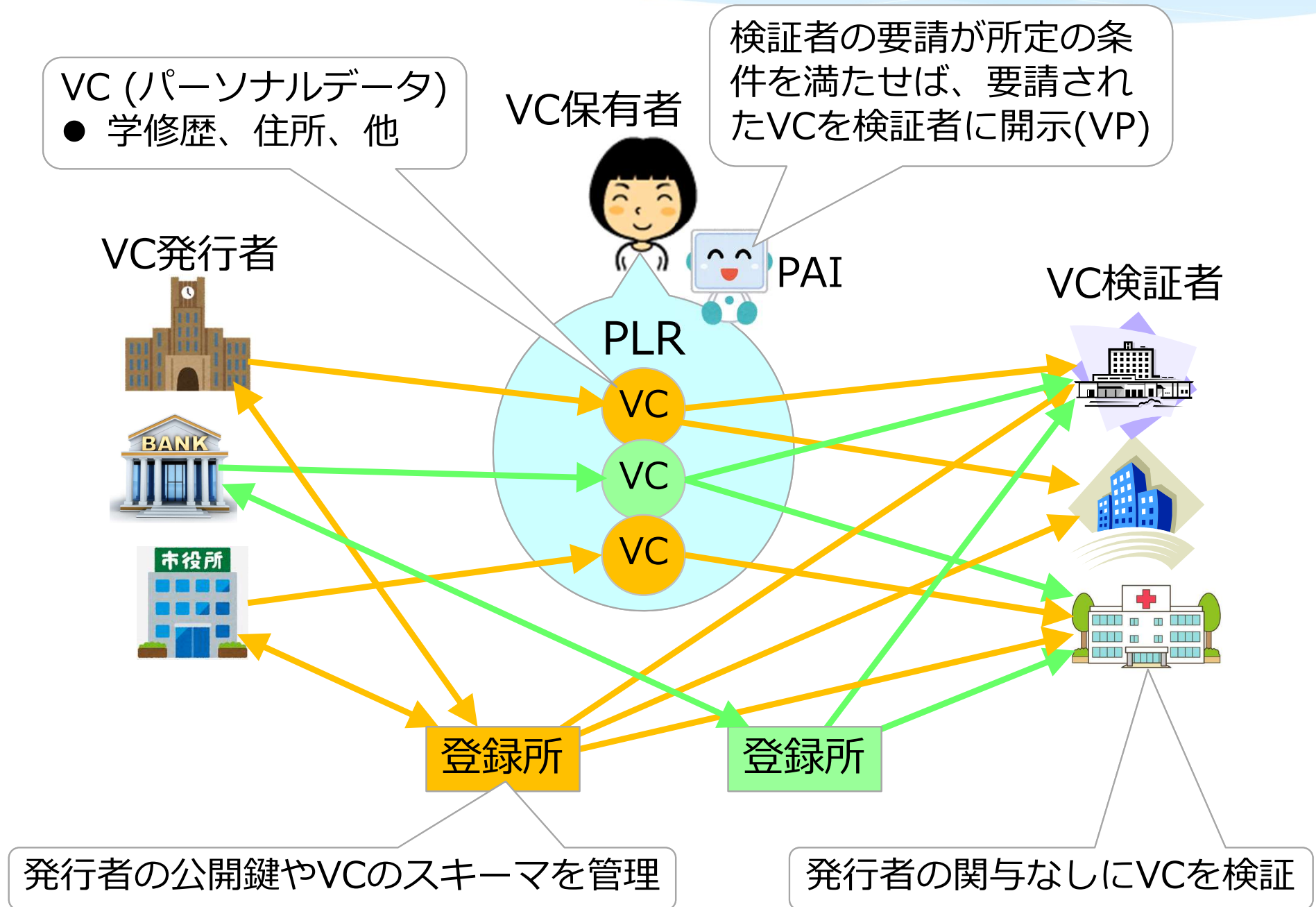


拡張版：オープン市民科学

民主政治・産業・学術・文化等に共通の基盤



検証可能属性証明(Verifiable Credential)の運用



行動変容

- 本人に行動変容を促す
 - ◆ サービスの利用
 - * 受診など
 - ◆ 生活習慣の改善
 - ◆ ヘルスリテラシーの向上
 - ◆ PDの開示
 - ◆ 臓器提供
 - ◆ ...
- 過剰な誘導のリスクも



来週火曜日は定例会がなくなったので健康診断に行きませんか？

でもその日は別の用事が入りそうだから。



じゃあ再来週の水曜日の朝はどうですか。

そうねえ。



この近所だと山本クリニックで受診できますよ。予約しときました。

あ、ありがとう。



メタバーズ→実世界での行動変容

- PAI: 本人を知り尽くしたパートナー
 - ◆ 本人(のペルソナ)に適応したナッジ
 - ◆ 実世界でもメタバーズでも有効
- アバター: メタバーズでの本人のペルソナ
 - ◆ ペルソナが含む望ましい行動特性等の発現を促進
 - ◆ 学習が成立すればその結果は実世界でも有効
- 実世界とメタバーズでの体験等に関する機微なPDを他者に開示せずPAIがフル活用
 - ◆ メタバーズでの行動変容を実世界に転移
 - * ペルソナを纏うことで改善する能力や行動特性をPAIが強化して実世界で活用
 - ◆ メタバーズで積極的に失敗して学習…旅の恥は掻き捨て
 - * 複数の場面にわたりPDを名寄せできるは本人だけ

まとめ

- PDの分散管理 → 価値の最大化
 - ◆ PDの名寄せと安全で安価なフル活用
- データポータビリティ
 - ◆ 学習eポータル等とPLRの連携
- グラフ文書
 - ◆ 知的共同作業の質と批判的思考力の向上
- パーソナルAI
 - ◆ 個人と事業者のメリット
- 行動変容とメタバース
 - ◆ メタバースの恥はかき捨て