

第8回 産業日本語研究会・シンポジウム
第四部 文書データの効果的な利用に向けて



shaping tomorrow with you

文書管理システムとナレッジマネジメント

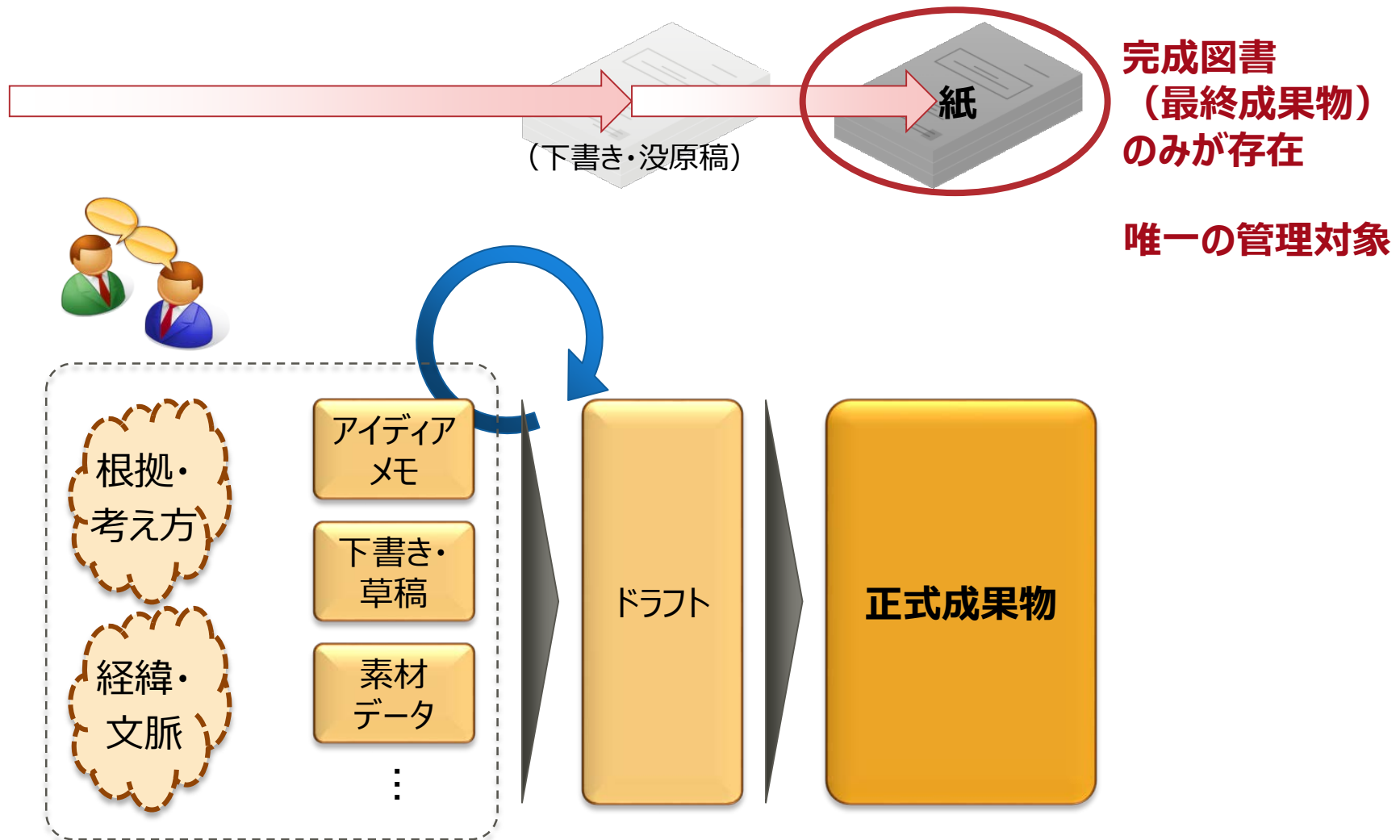
2017年 3月 8日

株式会社富士通総研
第一コンサルティング本部 シニアマネジングコンサルタント
小林 潔

文書管理システムの生い立ち

「紙」文書の発生プロセス

「紙」文書の時代



「紙」文書の時代

(実質的に) 存在していたのは、**最終的な成果物のみ**



文書管理

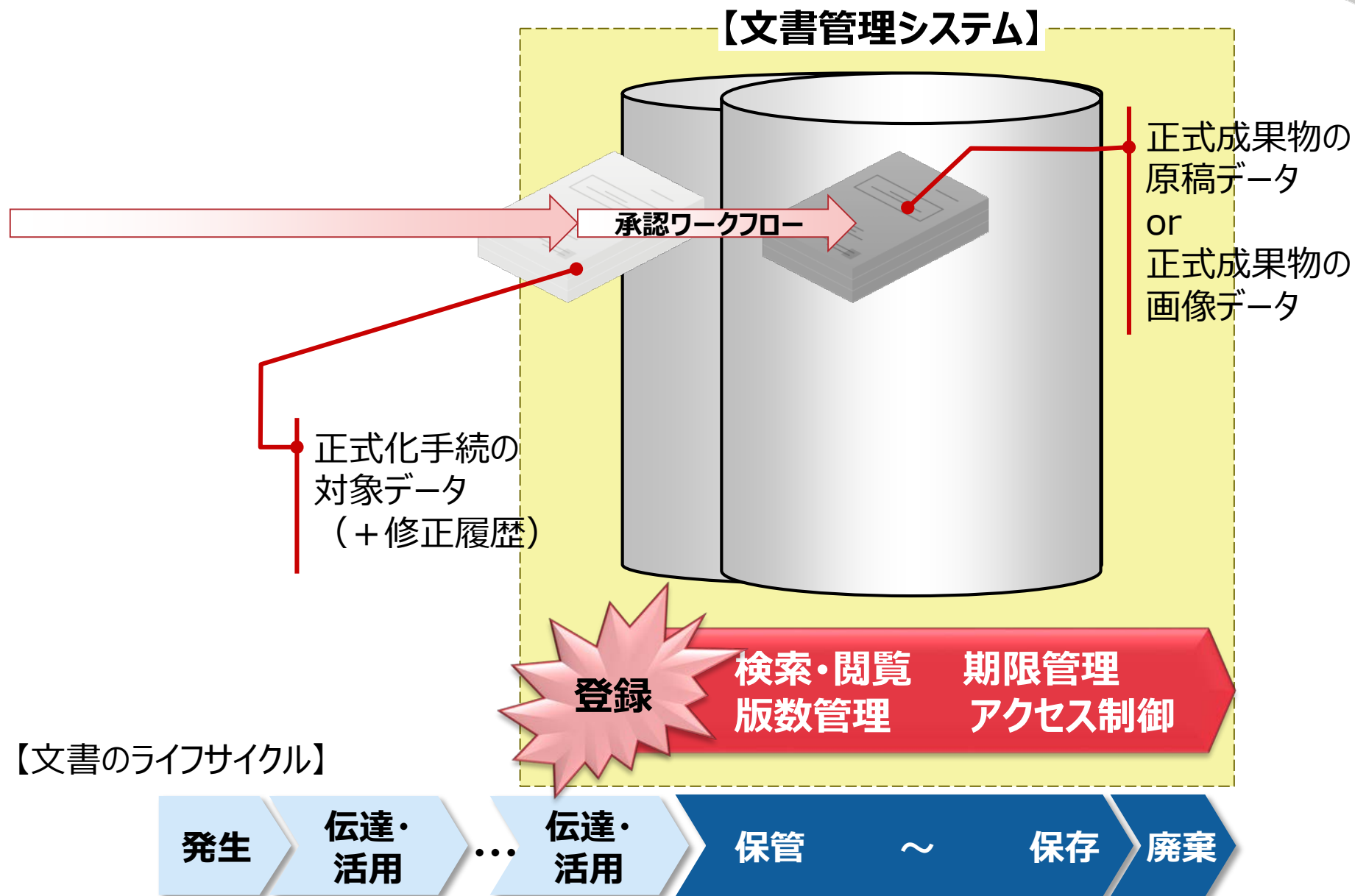


記録管理



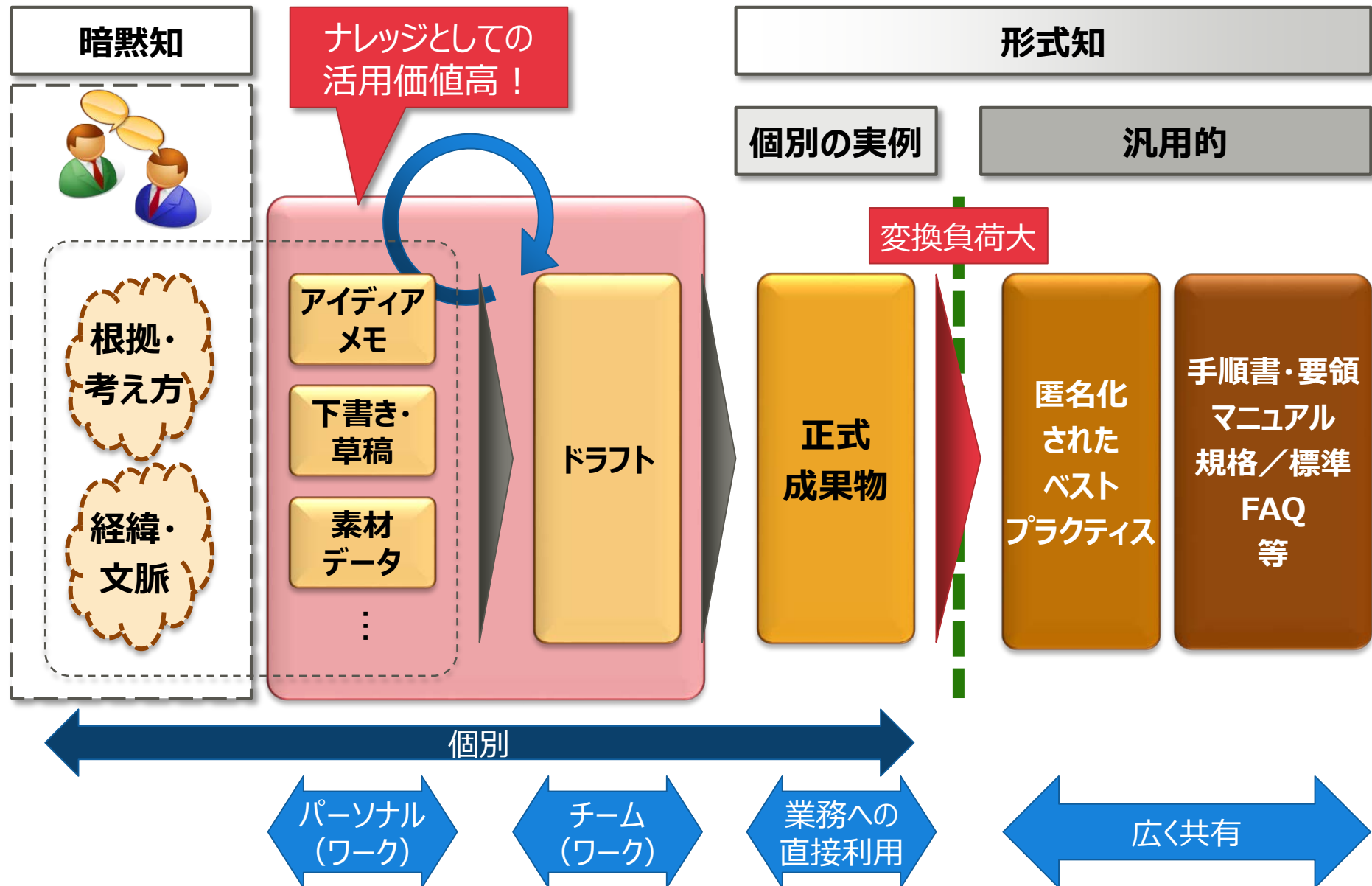
完成図書管理

- 文書 : document
- 記録 : records



ナレッジマネジメントへの応用

文書の発生プロセスとナレッジ



文書管理システムを
正式成果物以外の
ナレッジコンテンツの
格納庫として活用

検索機能を活用

形式知

個別の実例

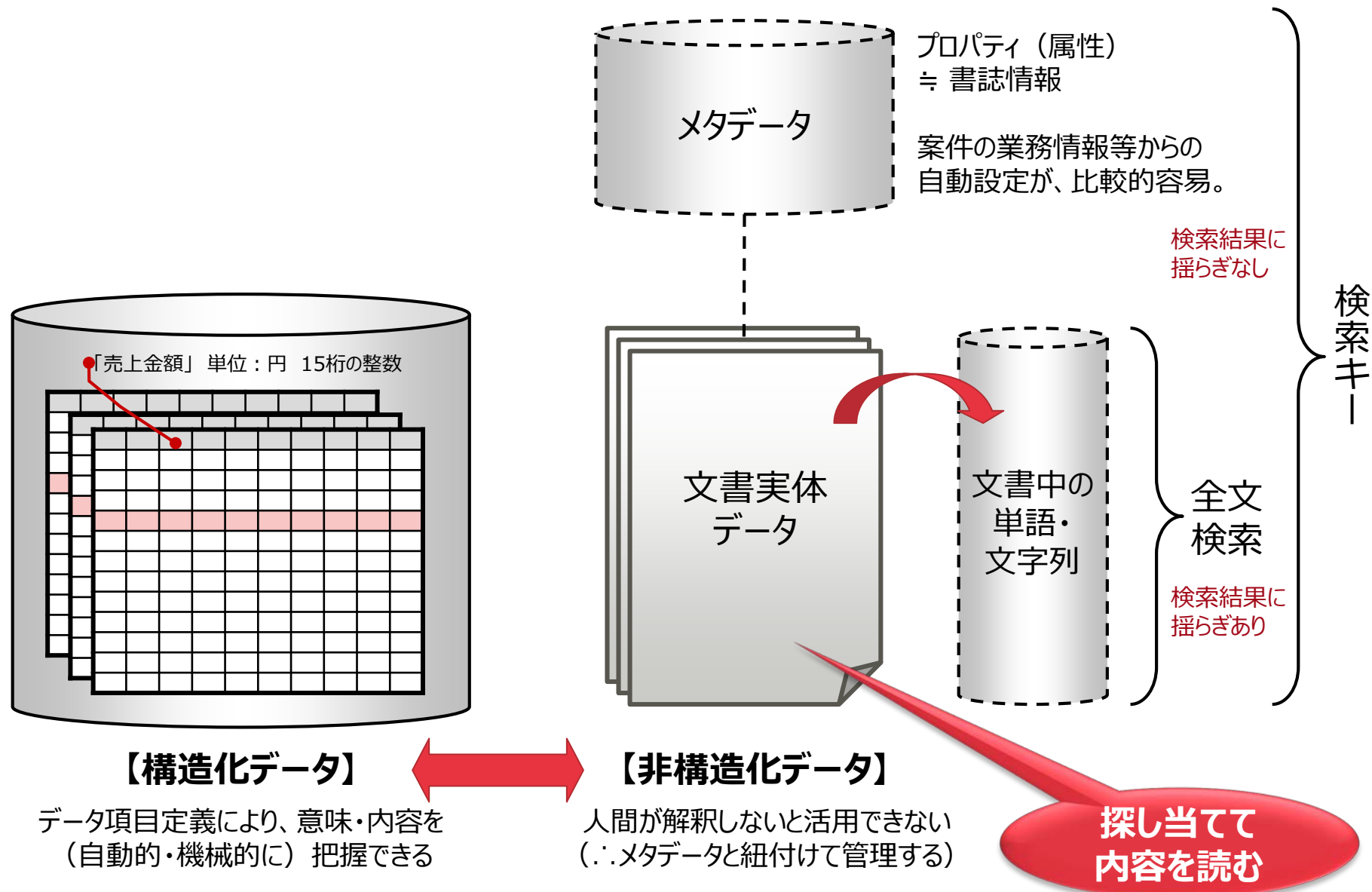
汎用的

正式
成果物

匿名化
された
ベスト
プラクティス

手順書・要領
マニュアル
規格／標準
FAQ
等

文書管理システム

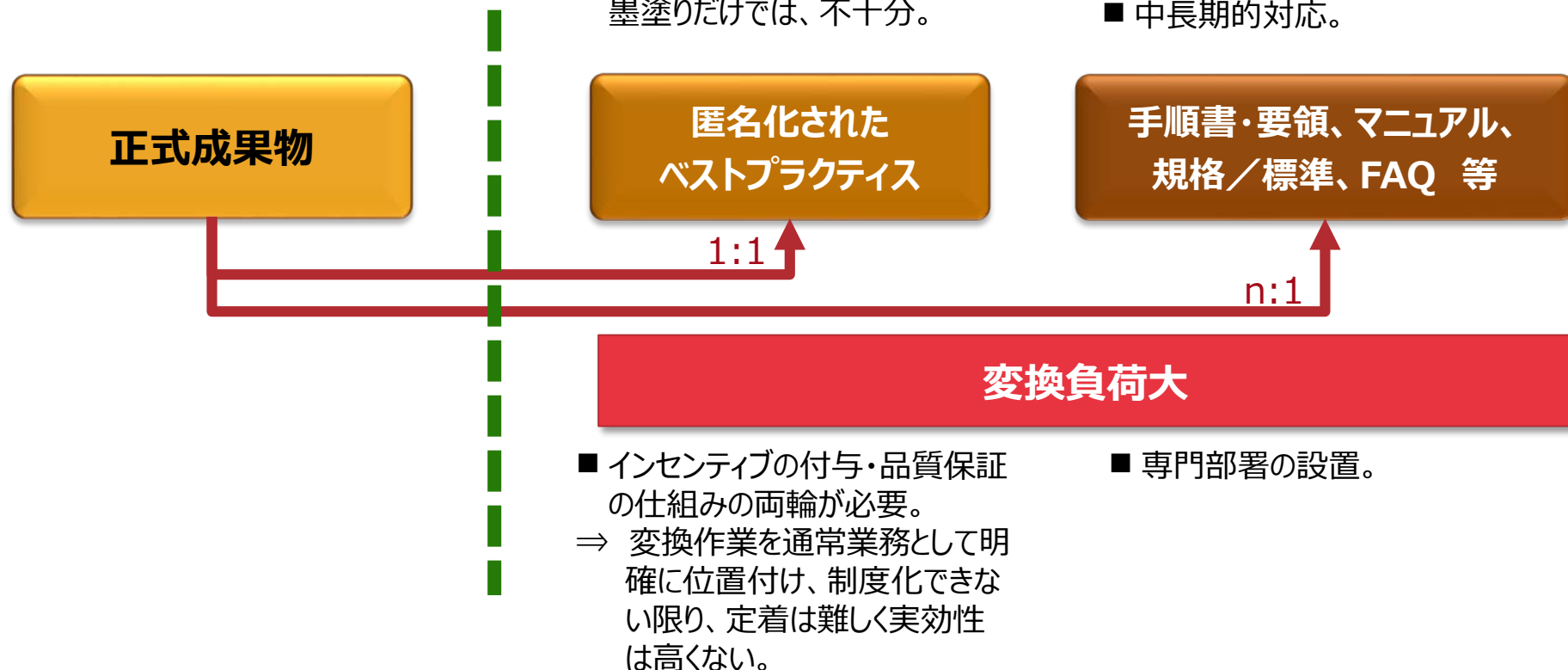


ナレッジコンテンツ

- 純粋な社内情報の場合は、成果物自体がナレッジ、再利用対象物となりうる。

- 他社秘密情報が含まれる場合には、匿名化加工が必要。
- 一般的に、固有名詞の単純な墨塗りだけでは、不十分。

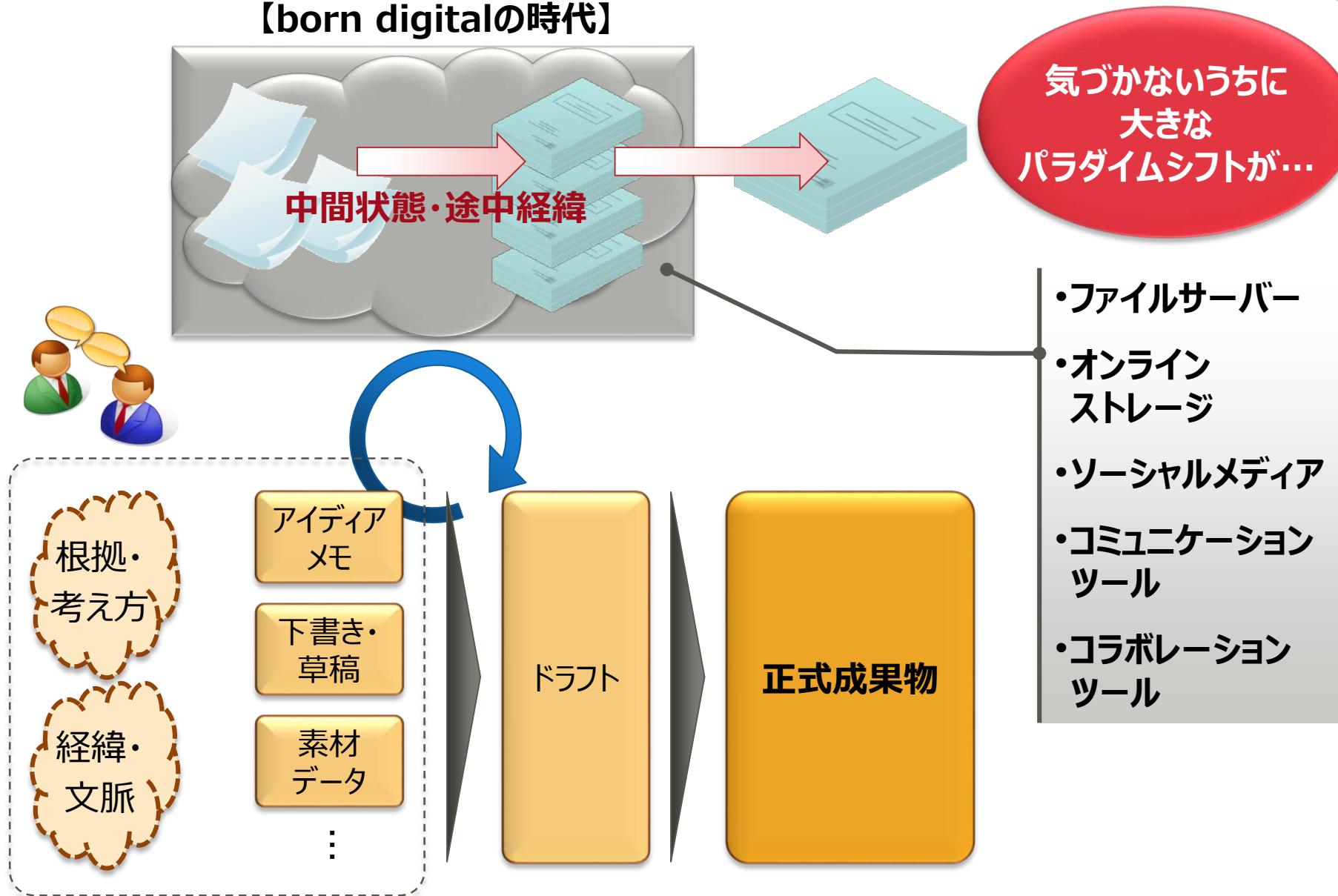
- 複数の実例からナレッジを抽出し、標準化すること（形式知化）が必要。
- 中長期的対応。



それでもなお、他の案件での再利用は難しい...

一方、今日の業務環境は…

【born digitalの時代】



現在のトレンドと到達点

- 課題：より有用性の高いナレッジは試行錯誤の経緯（ex. N Gだった実験パターン、失敗事例、…）



最終成果物にいたるまでの活動記録のデジタル化と管理

【事例】 製薬業界における電子「実験ノート」システム

- 課題：最終成果物に至った経緯や背景がわからないと、活用できない



文書管理システムのコミュニケーションツール／コラボレーションツールとの連携・一体化

- 途中の検討経緯・議論の状況を、最終成果物へと至るタイムラインに紐付けて一括管理
- 文書の執筆者への直接コンタクト。モノ（文書）ではなくヒト（のプロフィール）の検索

- 課題：文書からのナレッジの抽出負荷、検索精度など、非構造データとしての課題あり



セマンティック技術の発展 ⇒ AIの活用

- ビジネスの観点からの事例の出現
ex. XBRL (eXtensible Business Reporting Language)
⇒ 財務報告等のXMLによる国際標準化
- 事物に対し、万人が共通の解釈をできる「意味」を与える技術の進展
オントロジー（概念体系）、L(O)D (Linked (Open) Data)

製薬業における R & D 部門の「実験ノート」（研究日誌）のシステム化

user01 - EN1103 E-Notebook Instance

File Edit View Tools Help

Back Forward Demo_Notebook-001 Browse Search Batch Explorer Work Offline Help

user01

- User Configuration
- Offline
- Demo_Notebook
- Demo_Notebook-001

New Section

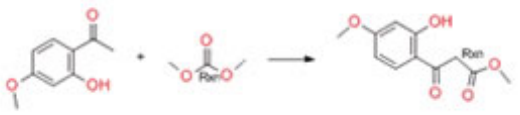
Batch Explorer

Add

Character Map

Reaction

Reaction



Add... Quick Add... Define... Delete

Reactant	MF	Limit?	MW	Eq	Moles	Sample Mass	Vol	Molarity	d
1	C9H10O3	☒	166.174	1.000	25 mmol	4.15 g			
2	C3H6O3	☐	90.078	1.000	27 mmol	2.432 g			

Product	RegSys	MF	Actual Mass	Actual Mol	Yield	Purity	MW	Eq	Theo Mo
1	OK	C11H12O5	5.45 g	23.58 mmol	94 %	97 %	224.210	1.000	25.00 mm

Preparation

Reaction

In a 100 mL pear flask (10g) was 1 (4.15 g, 25 mmol) and [REACTANTS2] in THF ([VOLUME]) to give a [COLOR] [SOLN/SUSP].

Solvents

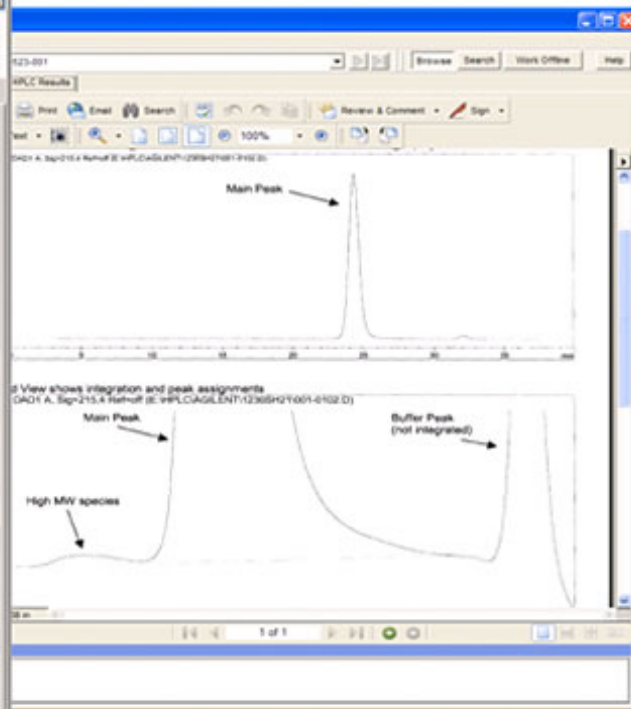
Name	Ratio	Volume
1 THF		

Reaction Conditions

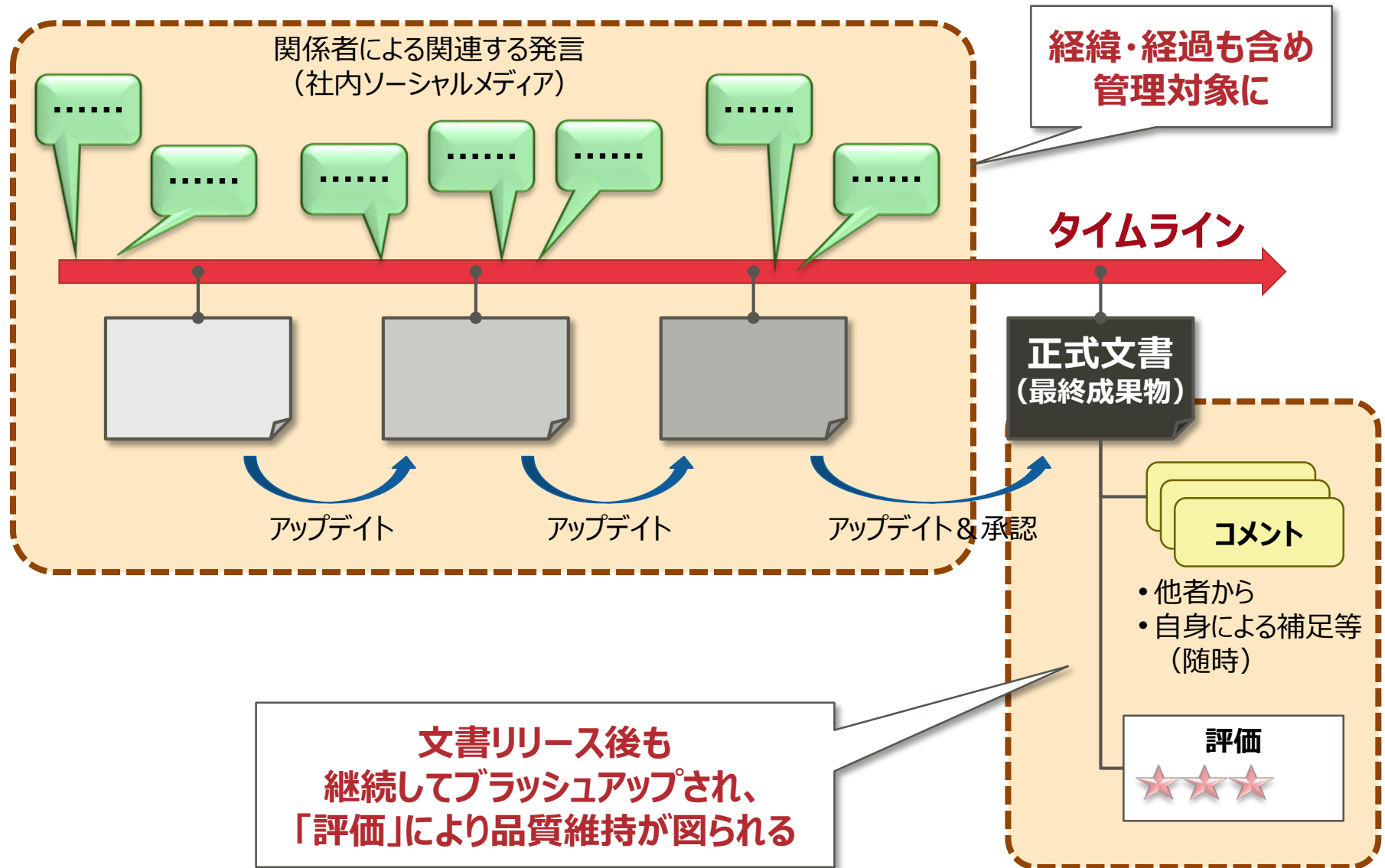
Reaction Molarity

Pressure

Temperature



中間状態や検討経緯のナレッジとしての管理



PRESS RELEASE (技術)

2016年10月6日
国立大学法人京都大学
富士通株式会社

平成28年度 日本医療研究開発機構「臨床ゲノム情報統合データベース整備事業」に採択 ゲノム・遺伝子多型の臨床解釈を推定するAI技術に関する研究開発を開始

国立大学法人京都大学（以下、京都大学）は、国立研究開発法人日本医療研究開発機構（以下、日本医療研究開発機構）が進める「臨床ゲノム情報統合データベース整備事業」に、京都大学大学院医学研究科の奥野恭史教授を研究代表とする研究開発課題「ゲノム医療を促進する臨床ゲノム情報知識基盤の構築」が採択されたことをお知らせします。本採択事業において富士通株式会社（以下、富士通）は、医師などが行っている臨床解釈の付与を支援する機械学習・AI（人工知能）技術の開発を担当します。本研究課題は、遺伝子情報を病気の診断や治療などに活用するゲノム医療を実現・普及させることを目的として、既に日本医療研究開発機構が整備を進めている、がんなどの疾患領域ごとの臨床情報やゲノム情報のデータベースと、各種公開データベースを統合し、信頼度・精度の高い臨床解釈を付与して公開する「臨床ゲノム統合データベース」を構築することを目指すものです。なお、実施期間は5年間（平成32年度まで）を予定しております。

【背景と目的】

遺伝子情報を病気の診断や治療などに活用するゲノム医療の実現に向けて、広く臨床現場で活用できる知識基盤としての「臨床ゲノム情報統合データベース」の構築は喫緊の課題となっています。米国の事例では、疾患に関連する遺伝子多型とその臨床解釈を集約する枠組みとして、専門機関が自由にアクセスできるデータベースをいち早く構築して公開しています。しかし、データの信頼度や精度が、臨床解釈を提供した臨床現場・組織に依存することが課題となっており、専門家による評価の仕組みを体系化するなど信頼度や精度改善のための取り組みが進められています。国内においても、データベース基盤の整備を進めていく中で、データの集約と臨床解釈の付与についての標準化・体系化の取り組みが求められています。

【本事業の取り組み】

今回、京都大学と富士通は共同で、これまで専門家が行ってきた人手によるデータ集約作業の過程を定式化・アルゴリズム化し、専門家によるデータ集約作業の支援を行うキュレーションシステムを開発します。本キュレーションシステムでは、独自の機械学習・AI 技術を開発することで、高精度かつ高速な臨床解釈付与の支援を目指します。

本キュレーションシステムの開発において、富士通は、現時点で疾患との関連性が明らかではない遺伝子多型に対する臨床解釈を推定する以下の技術についての研究開発に取り組みます。

株式会社富士通研究所がこれまで開発してきたLOD（Linked Open Data）活用基盤を利用し、医療分野における学術文献や公共DBを集約した知識ベースを構築します。

構築した知識ベースを元に、富士通のAI技術「Human Centric AI Zinrai（ジンライ）」を構成する独自の機械学習技術を適用し、臨床解釈の推定と、その根拠となるエビデンスおよび治療薬候補などを出力するシステムの構築を目指します。

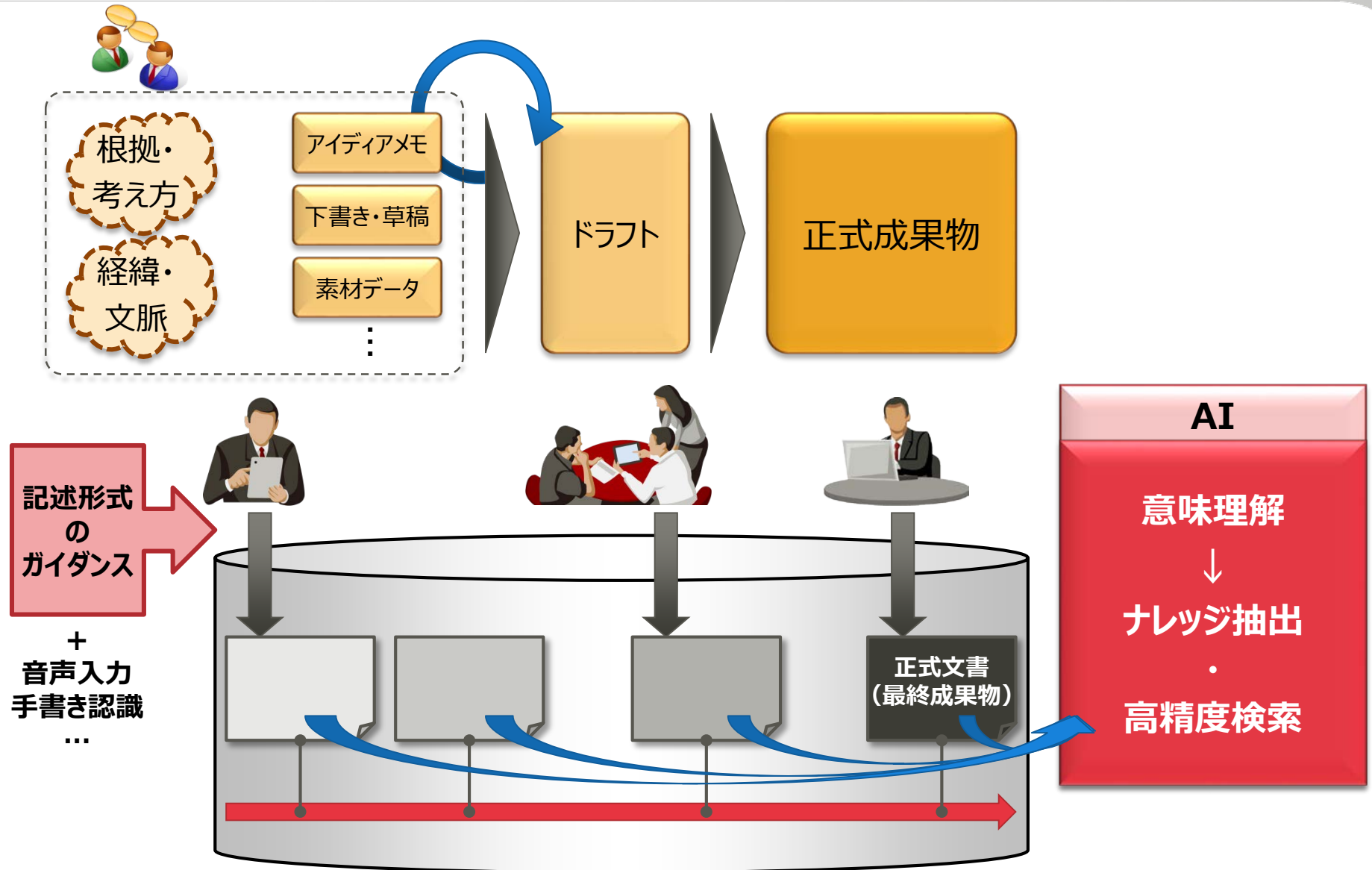
これらの技術を発展させ、医療だけでなく様々な分野における企業や大学の研究開発を支えるAI技術として富士通のビジネスへも展開していく予定です。

【商標について】

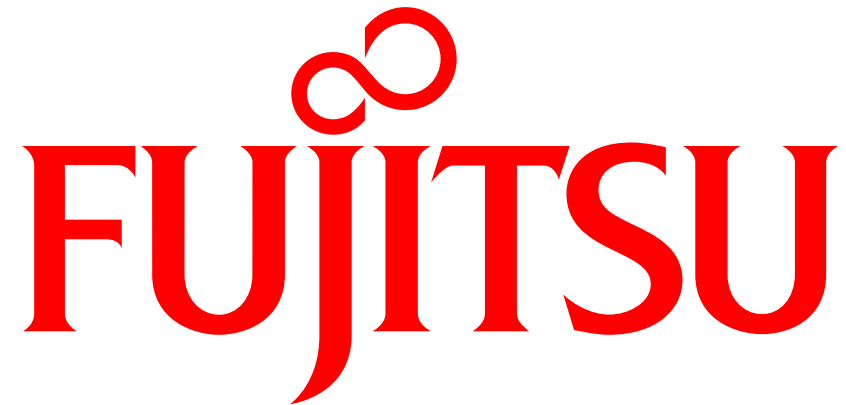
記載されている製品名などの固有名詞は、各社の商標または登録商標です。

以上

未来予想図



文書管理システム ⇒ Enterprise Content Management ⇒ その先へ



shaping tomorrow with you