

オープンリサーチデータ推進センター（H28から）

- 適切なメタデータを付与して、データの検索性、再利用性を高める。
- 研究者や図書館員がオンラインで共同タグ付け（キュレーションすることで**知識が蓄積するデータ評価プラットフォーム**を構築（国文学研究資料館）
- **タグの推薦機能**（メタデータや過去のタグ付け機能と知識基盤を組み合わせ、関連タグの推薦により作業効率を高める。）
- **画像の解析機能**（画像データの内容分析に基づき、メタデータにならないタグを推薦）
- データ評価プラットフォームの横展開
 - ✓ 人文科学から自然科学まで、分野の特性を踏まえたキュレーションが行えるプラットフォームを構築
 - ✓ データのオープン化の便益の見える化

21

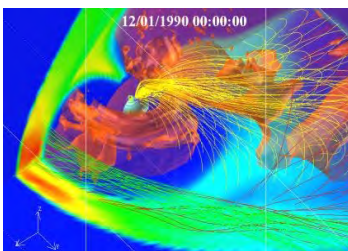
データ解析基盤

22

データ解析の方法開発・ツール提供

- 可視化と構造探索
- モデリングと知識発展, ベイズ推論
- 機械学習, 計算アルゴリズム
- シミュレーション・データ同化・意思決定
- e-サイエンス支援システム

Visualization of high-dimensional data and analysis results

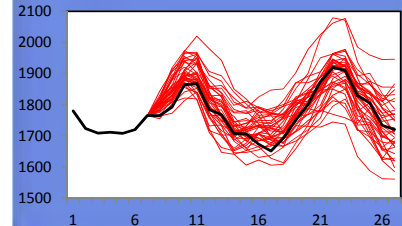


データ同化研究開発センター, ISM

Spiral of knowledge development



Ensemble simulation



23

データ同化：モデルとデータの統合

- 気象学・海洋学の分野で開発
- 数値シミュレーションと観測データを統合

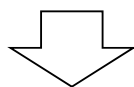
✓ シミュレーションモデル

複雑な現象を完全には表現できない

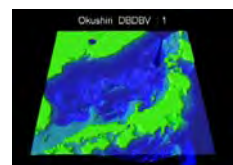
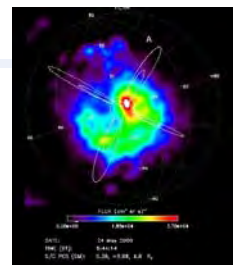
- 厳密な初期条件
- モデル不確実性 (境界条件, パラメータ, 構造)

✓ 観測データ は有限 (スパース)

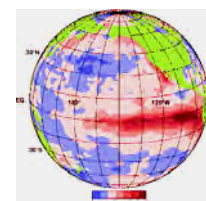
- 物理的または予算上の条件



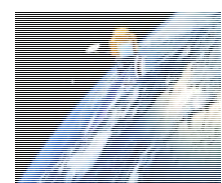
データ同化 = データを使ったシミュレーションモデルと結果の改良

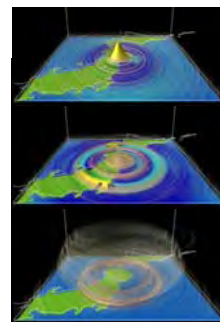


津波データ同化



エルニーニョ予測

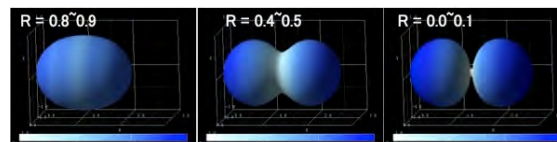




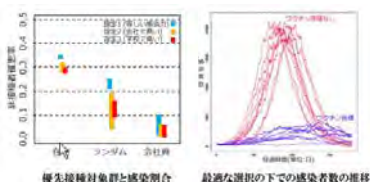
大地震に伴う大気変動のデータ同化

研究開発内容:

- アンサンブル・カルマンフィルタ、粒子フィルタ等のデータ同化の基本技術の研究開発
- データ同化手法の応用
 - ・ 大洋-大気系連成モデル
 - ・ 津波・微気圧データ同化（津波予測）
 - ・ 惑星磁気圏3D構造（宇宙天気予報）
 - ・ バイオインフォマティクス、細胞分裂
 - ・ 感染症拡散予測・制御
 - ・ 化学プラント

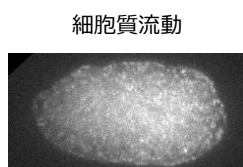


細胞質分裂時の細胞表面シミュレーション



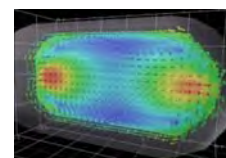
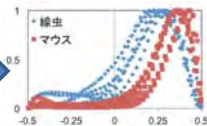
優先候補対象群と感染割合

最適な選択の下での感染者数の推移



細胞質流動

せん断力分布の推定



細胞内流動の可視化

ゲノム解析支援センター（H28より）

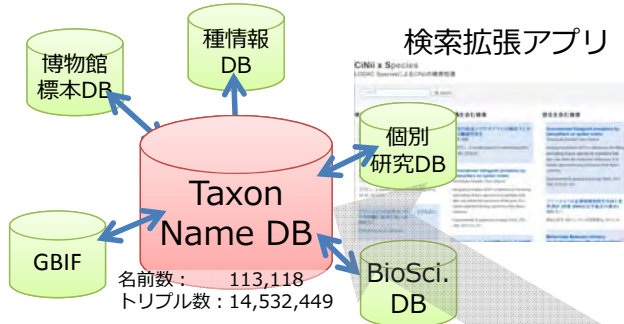
- 大学等の研究者が**次世代大規模ゲノム解析**により革新的な生物学的知見を得るために必要な支援を行う。
- 企業等が提供する解析パイプラインやプログラムの単純な組み合わせでは解析が困難なデータセットを対象に、遺伝研の先端ゲノミクス推進センター、DDBJ, DBCLSとも連携し、高度なゲノム解析支援を総合的に実施する。

これまでの実績：

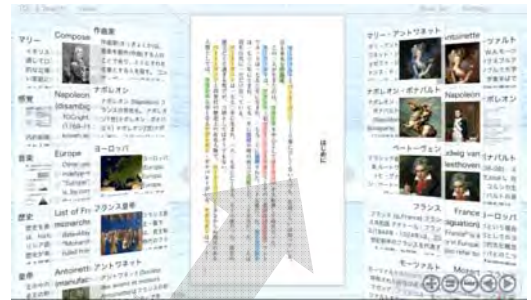
科研費新学術領域研究 **ゲノム支援**（ゲノム科学の総合的推進に向けた大規模ゲノム情報生産・高度情報解析支援）

e-サイエンス基盤構築

オープン・ソーシャル・セマンティックWeb基盤



連想情報処理基盤



情報推薦の基盤技術



- ◆ 論文の意味解析技術
- ◆ 論文の役割解析に基づく推薦指標判定技術
- ◆ 情報資源の同定・検索基盤技術
- ◆ 情報推薦の数学的モデル



e-サイエンス基盤技術，データ中心科学リサーチコモンズ，ROIS

27

従来指標では捉えられない共同利用・共同研究の成果や異分野融合の進展状況を可視化する方法の開発及び活用

IRの現状

- 海外のベンチマーキングでは活用されない日本語の論文や産学官連携活動、社会貢献活動などの軽視 → 日本の学会 活動の停滞

日本の持続的発展への大学等の貢献度を測る指標の開発とその普及が急務

計画の概要

- 戦略企画本部で我が国のIRに対するニーズを把握し、IRのハブ的機能を担う。
- ROISが開発したresearchmapデータ・他を元に、研究情報を自動収集する仕組みを開発し、コミュニティに提供
- 多様な研究情報の分析・可視化の手法を研究開発し、機構の共同利用・共同研究の推進体制等の改善に活用
- 各コミュニティに分析・可視化ツールを広める

戦略性

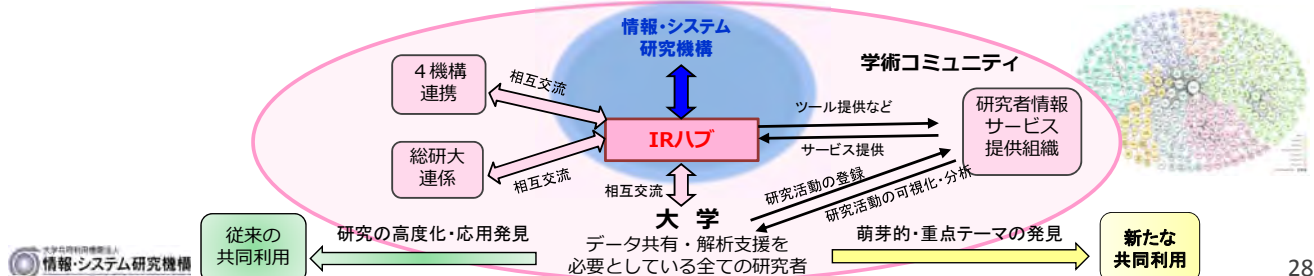
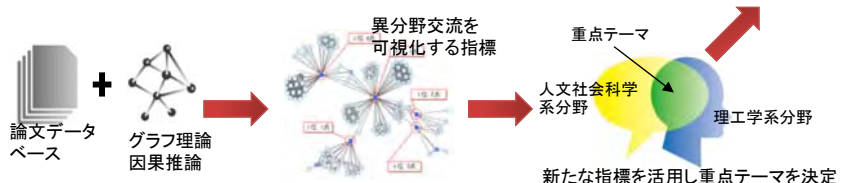
- researchmapは24万人の研究者をカバー。商用サービスにはないデータ（特許・競争的資金・日本語論文・年齢・性別・職種等）を数千万件保有
- 日本の科学技術戦略の立案には、自ら研究情報を所有し、分析する手法を確立することが必須

データ活用支援ツールの研究開発

- IR専門家と情報学の連携構築
- researchmapデータや他の研究者情報から、各機関が自己の研究情報を選別し、IRに活用するためのツール（API）を設計・開発
- 研究情報分析ツールを研究・開発
 - (1) 生産性の高い組織の在り方の分析支援
 - (2) 領域の萌芽や関連研究者の発見支援
 - (3) 基礎と応用研究、社会実装への橋渡し

異分野融合や新分野創成

- 異分野融合の進展や効果を可視化する指標を研究し成果を公開
- 将来的な新たな研究分野創成につながる各分野の動向予測



28

● イメージデータ解析

医用や生体、社会科学などの幅広い分野における複雑な現象を視覚的に表し、そのイメージデータを統計的手法により深く高度に解析して、複雑な現象を視覚的に表し、人間の直感を働かせやすくすることにより、その現象の本質に迫り、意思決定支援に役立てる基盤技術を確立。

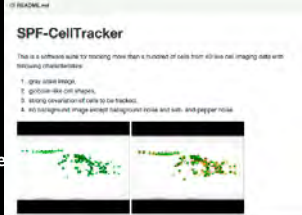
- ① バイオイメージング・インフォーマティクス
- ② イメージングデータを用いた脳の動作原理の解明
- ③ 人間・社会の視覚情報データ解析

SPF-CellTracker
Hirose et al., under review

Software Released @
<http://github.com/ohirose/spf>

TRACK (cell tracking)
C language
Real-time tracking
512x256x20, 1000 frames,
200 cells, about 30 sec
Windows, Mac, Linux
(Ubuntu)

VIEW (visualization software)
OpenGL
Windows, Mac, Linux



● 自然言語処理(メタ知識構造の言語的・統計的モデリング)

学術論文中の明示的・非明示的に表現された因果関係、理由、目的といったメタ知識構造を自動認識し、それに基づき構造化された知識を自然言語テキストから自動抽出する手法の研究。

- 自動学術論文を対象として、メタ知識構造を解析することを目指す
生データには存在せず、テキスト化（＝知識化）された時に初めて現れる「メタ知識構造」に着目
- コーパス分析に基づき、概念や関係の表現方法を策定
- 構造予測モデルによる認識

データ中心科学リサーチコモンズ, ROIS 29

人材育成