

「バイオデジタルコンテンツ整備における 京速計算機の役割」

理化学研究所バイオリソースセンター
情報解析技術室 深海 薫

*Foundation for Discoveries
and
Access to the Future*



バイオデジタルコンテンツ

- ライフサイエンス研究における新たな研究基盤 -

ゲノムプロジェクトの進展により
生命科学が抱える
情報量は飛躍的に増大



コンピュータによる情報処理・管理を行うため
生命科学情報をデジタルデータとして持つ必要



「バイオデジタルコンテンツ」は
ライフサイエンス研究の遂行に必要な
情報源として、雑誌情報に匹敵
あるいはそれを上回る
重要な役割を持つようになる



「バイオデジタルコンテンツ」を 有用な研究基盤とするために

研究コミュニティに散在するコンテンツを
特定の目的・視点の下に収集・整備する必要

生命情報

デジタルデータ

データベース
ハイパーテキスト

統合データベース

バイオリソースの
所在情報
特性情報

電子化

ウェブカタログ
(データベース)

他データベースとの
統合

ライフサイエンスの研究用材料(バイオリソース)においても
バイオデジタルコンテンツの整備が進行中

ナショナルバイオリソースプロジェクト(NBRP)

特性情報：実験材料に関する知識・情報

あるマウス系統に
関する知識・情報として
どのようなものがあるか？



外見

形態

行動

生理学的特性

遺伝子型

発現パターン

ゲノム情報

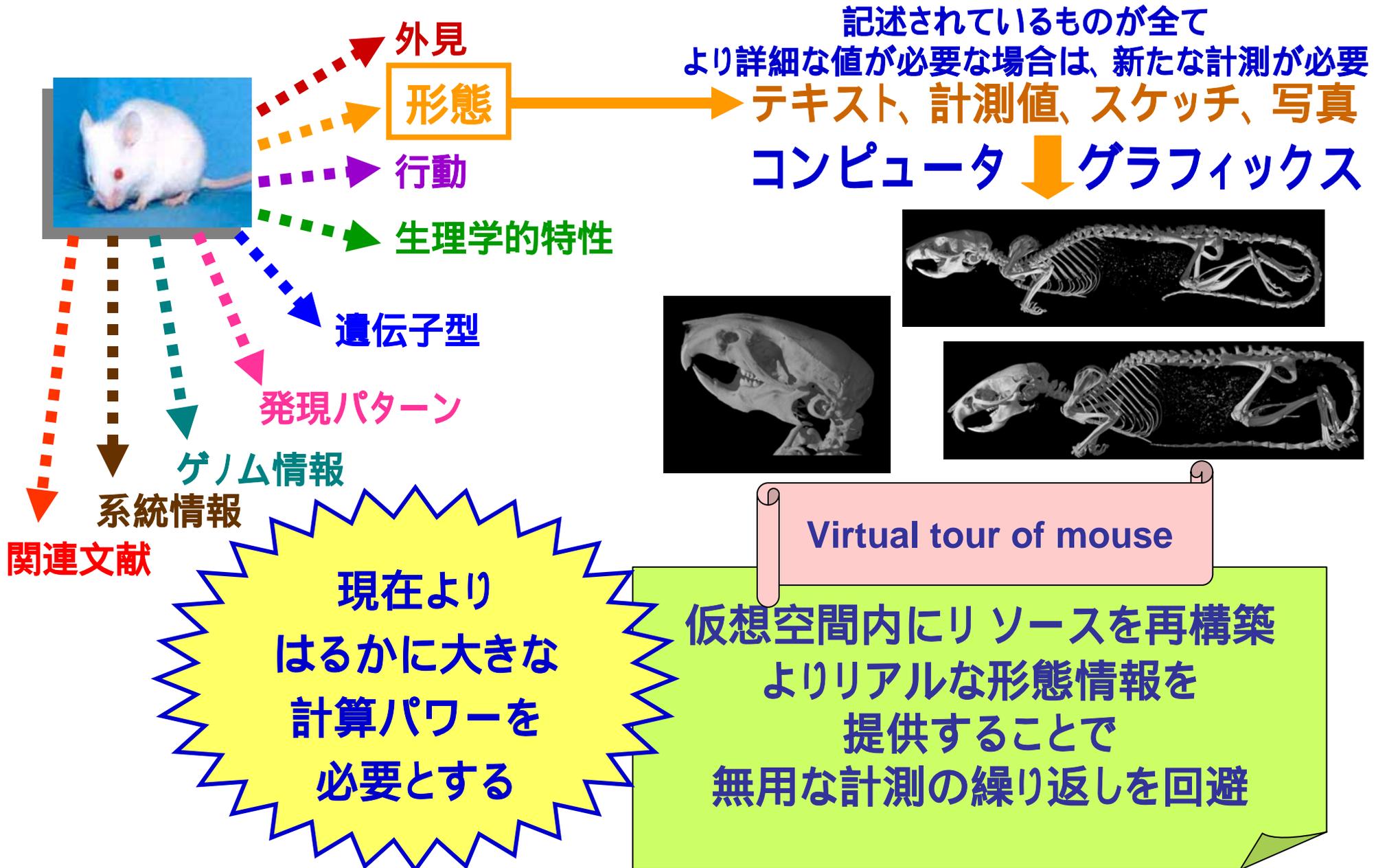
系統情報

関連文献

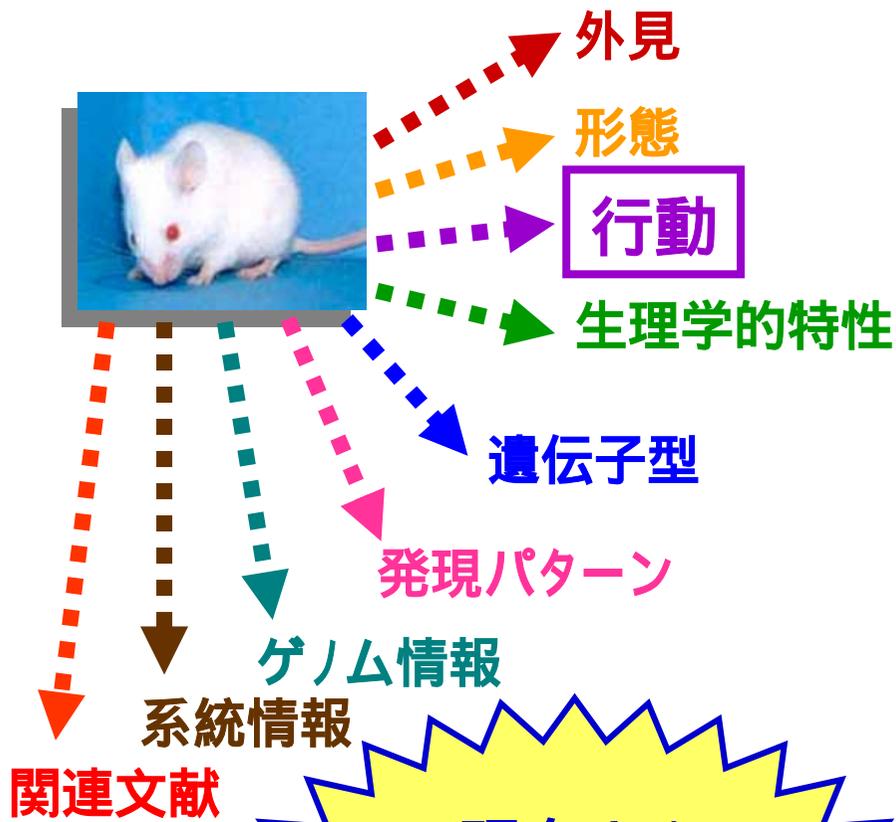
こうした情報の質・量が
研究成果の質・量を左右する
大きな要因となる

既知の特性情報を
アーカイブ化し
研究コミュニティに
提供することで
同じ実験の繰り返しを
回避できる

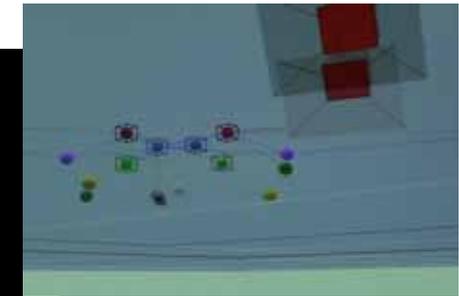
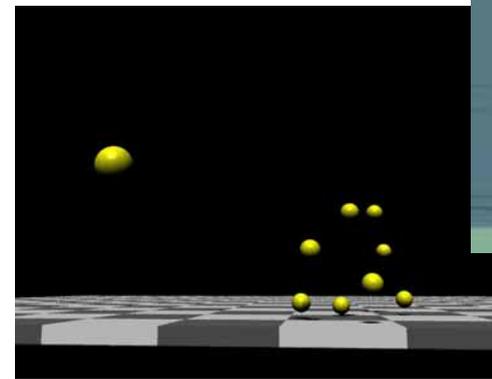
デジタルコンテンツを駆使した特性情報提供



デジタルコンテンツを駆使した特性情報提供



テキスト
計測値
ビデオ



Digital mouse

現在より
はるかに大きな
計算パワーを
必要とする

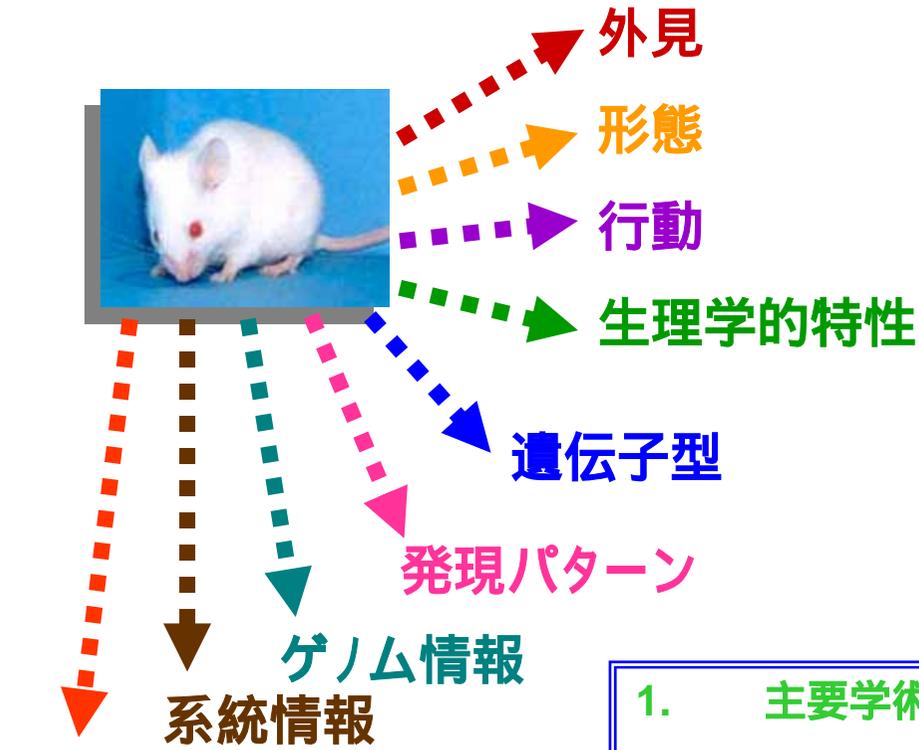
大量の行動データを
客観的に解析できる

デジタルコンテンツを駆使した特性情報提供

BRCが提供したリソースから生み出された
文献成果は...

1. BRCリソースの有用性を示す具体例となる
2. BRCリソースの特性情報となる
 - リソースデータベース充実
 - リソースの付加価値上昇
 - 知的資産の蓄積

文献情報から知識への変換



関連文献

情報量の大きさ
内容の豊かさでは
ダントツ

1. 主要学術雑誌の選定
2. 各雑誌の対象論文
3. 絞り込み
4. キーワード
5. 専門家による
6. BRCリソースを用いて

現在の文献検索・解析の
精度は、計算機パワーに
制約されている
(十分な解析を行うためには
0.1ペタフロップス程度必要)

まとめ

- バイオデジタルコンテンツは、ライフサイエンス研究における新たな研究基盤である
- 有用性を高めるため、特定の目的・視点の下に収集・整備する必要がある
- ライフサイエンスの研究用材料(バイオリソース)においてもバイオデジタルコンテンツの整備が進行中である
- デジタルコンテンツを高度に駆使して特性情報を提供するためには、現在よりはるかに大きな計算パワーを必要とする