



# 計算生命学概論

## 2016年10月4日

神戸大学大学院 科学技術イノベーション研究科  
先端医療学講座  
森一郎

# 計算生命科学の基礎Ⅲ

生命科学と理工学の融合による  
生命理解と医療・創薬への応用

# 本日の講義内容

1. 遠隔講義「計算生命科学の基礎」について
2. 計算生命科学とは
3. ノーベル賞と計算生命科学
4. 「生命科学研究への情報科学技術のインパクト」から
  - ① データサイエンス
  - ② 構造ベース創薬 (タンパク質構造データベース PDB)
  - ③ 人工知能 (AI)
5. 2016年度シラバス紹介
6. まとめ Q&A

# 講義の趣旨・対象者

## 講義の目的

- 生命科学の近年の急速な進歩は計算機科学、統計学、シミュレーション科学等によって大きく支えられてる。
- 生命科学と理工学の学際的研究領域である「計算生命科学」の基礎と展望を学ぶ機会を全国に提供
- 「計算生命科学」のさまざまな分野で活躍する人材の育成を目指す

## 講義内容の範囲

- 遺伝子情報～医療応用まで
- 情報科学、シミュレーション科学、人工知能技術

## 受講者対象(レベル)

計算生命科学に興味をお持ちの方全て

大学院生～企業・アカデミアの研究者（高校生・大学生、一般）

# 2016年度スタッフ紹介

## 企画・コーディネート担当

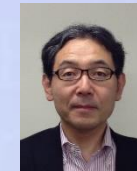
◆木下 賢吾 (東北大学大学院 情報科学研究科)



◆江口 至洋 (神戸大学 連携創造本部\*)



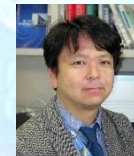
◆田中 成典 (神戸大学大学院 システム情報学研究科)



◆森 一郎 (神戸大学大学院 科学技術イノベーション研究科)



◆鶴田 宏樹 (神戸大学 連携創造本部\*)



### 講義配信担当

◆渡邊 博文 (神戸大学 連携創造本部\*)

◆近藤 洋隆 (神戸大学 計算科学教育センター)

\*10/1以後: 学術・産業イノベーション創造本部

### 講義運営担当

◆横川三津夫(神戸大学 計算科学教育センター)

◆橋本賀津子(神戸大学 計算科学教育センター)

◆土井 陽子(理化学研究所 生命システム研究センター)

## 企画協力: CBI学会、日本バイオインフォマティクス学会




計算科学教育センター  
Education Center on Computational Science and Engineering


神戸大学  
ENGLISH

概要・目的

遠隔講義

シミュレーションスクール・セミナー

n-computer

n-VizStudio

n-CAVE

遠隔講義

TOP > 遠隔講義

遠隔講義一覧

講義名をクリックすると詳細がご覧になります。

| 講義名                                  | 日程                                                    | 場所                                                                                                                                                                                          | 申込            |
|--------------------------------------|-------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|
| <a href="#">計算生命科学の基礎Ⅲ</a>           | 2016/10/4 (火) ～<br>2017/1/24 (火) 毎週火曜日<br>17:00～18:30 | 会場：神戸大学計算科学教育センター セミナー室208 (神戸大学統合研究拠点2階)<br>会場は定員30名です。講師へ直接質問できる機会もございますので、ぜひ会場でもご受講ください！<br>会場以外からはWebEX (インターネット) による受講が可能です。事務局からの招待状メールの指示に従って、PC・モバイルから記載URLにログインするだけで、どこからでも参加可能です) | <div>申込</div> |
| <a href="#">計算生命科学の基礎Ⅱ (一終了しました)</a> | 2015/10/7 (水) ～<br>2016/2/3 (水) 毎週水曜日<br>17:00～18:30  | メイン会場：神戸大学工学部本館1階 C3-101(創造工学スタジオ2) ※神戸大学(六甲台キャンパス)は、立地場所的に11月頃よりかなり寒くなりますので、講義会場で受講される方は暖かい服装でお越しください。<br><br>中継会場(東京)：産業技術総合研究所 臨海副都心センター 別館10階 会議室<br>(学校からはWebEXによる受講が可能です。事            | <div>—</div>  |

遠隔講義のご案内



シミュレーションスクール・セミナーのご案内



π-computer  
バイ・コンピュータ



π-VizStudio  
バイ・ビズスタジオ



π-CAVE  
バイ・ケイブ



## 計算生命科学の基礎Ⅱ 2015

銅谷 賢治（沖縄科学技術大学院大学神経計算ユニット 教授）

▼はじめに「計算生命科学の概要：田中 成典 《10月7日(水)》 [【講義資料PDF】](#) ※二次配布禁止 [→アーカイブ動画](#)

コンピュータやIT技術、情報科学の進歩とともに近年急速に進展してきた計算生命科学が、いかに生物学、医学・薬学、農学などの生命科学の研究を促進してきたのか、その理論的背景は何であったのかを、核酸やタンパク質の分子レベルから細胞、組織、臓器レベル、さらには個体間や環境との相互作用（生態系）レベルの階層性を意識しながら示し、もって本講義全体の導入的紹介とする。

～第1編 ゲノムから見る生命科学～ [参考図書] 1. 柳田敏雄他「計算と生命」岩波書店（2012）  
2. Uri Alon（著）、倉田博之・宮野 悟（訳）「システム生物学入門 -生物回路の設計原理-」共立出版（2008）  
3. Gregory N. Stephanopoulos, Jens Nielsen, Aristos A. Aristidou（著）、清水 浩・塩谷捨明（訳）「代謝工学—原理と方法論」東京電機大学出版局（2002）

▼1.1 「ヒトの病気はヒトの研究で～生命情報の統合によるヒト生物学と先制医療」：松田 文彦  
《10月14日(水)》 [【※公開講義資料なし】](#)

実験動物や細胞でヒトの病気を研究するのではなく、ヒトが極めて多様な集団であることを意識しつつ、各個人から得られる様々な生命情報を統合した生命ビッグデータを解析することで、さまざまな疾患の発症機構の解明および疾患の超早期診断や予後予測を試みる「ヒト生物学」の目指すところとその方法論について紹介する。

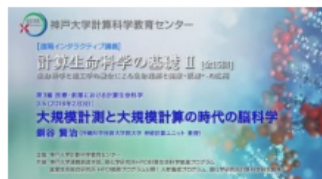
▼1.2 「生物システムの設計：システム生物学から合成生物学へ」：荒木 通啓 《10月21日(水)》 [【講義資料PDF】](#) ※二次配布禁止 [→アーカイブ動画](#)

人工物の設計にCAD（Computer-Aided Design）が利用されるように、生物システムの設計にも

# eラーニングアーカイブ (2015年度講義ビデオ)

## 理化学研究所 計算科学研究機構

### 計算生命科学の基礎



- |            |    |      |                                                           |
|------------|----|------|-----------------------------------------------------------|
| 2016.02.03 | 動画 | レベル2 | <a href="#">大規模計測と大規模計算の時代の脳科学</a>                        |
| 2016.01.13 | 動画 | レベル2 | <a href="#">フラグメント分子軌道法によるタンパク質-リガンド相互作用の高精度解析と創薬への応用</a> |
| 2015.12.09 | 動画 | レベル2 | <a href="#">QM/MMシミュレーションによるタンパク質機能解析</a>                 |
| 2015.12.02 | 動画 | レベル2 | <a href="#">分子動力学計算を活用したインシリコ創薬</a>                       |
| 2015.11.25 | 動画 | レベル2 | <a href="#">分子動力学計算によるタンパク質の機能解析</a>                      |
| 2015.11.18 | 動画 | レベル2 | <a href="#">タンパク質の量子化学計算</a>                              |
| 2015.11.04 | 動画 | レベル2 | <a href="#">到来する大規模生命情報の解析に備えて</a>                        |
| 2015.10.21 | 動画 | レベル2 | <a href="#">生物システムの設計：システム生物学から合成生物学へ</a>                 |
| 2015.10.07 | 動画 | レベル2 | <a href="#">はじめに：計算生命科学の概要</a>                            |

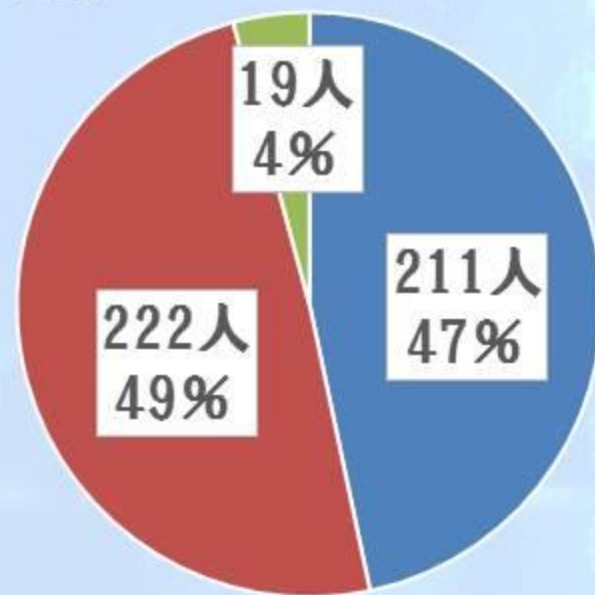
アーカイブなら  
いつでもどこでも  
学習可能です。

Copyright © 理化学研究所 計算科学研究機構 <http://www.aics.riken.jp/jp/> 神戸大学計算科学教育センター  
<http://www.eccse.kobe-u.ac.jp/>

<http://www.aics.riken.jp/jp/course/course-base>

# 遠隔講義受講者登録状況

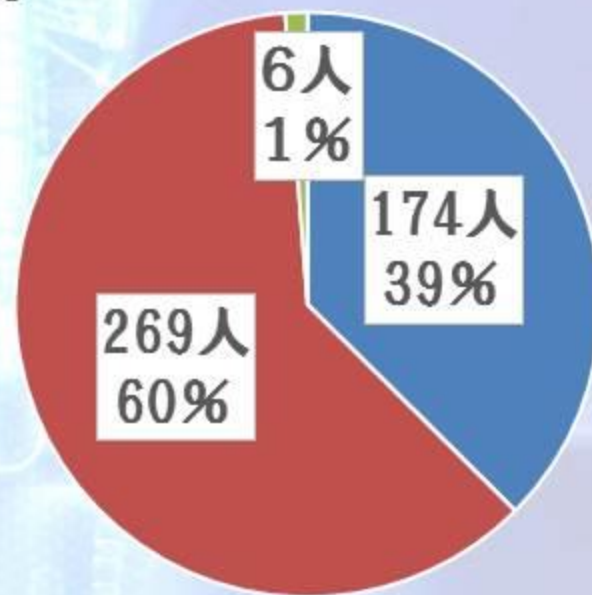
2015年度



■ 大学 ■ 企業・研究所等 ■ その他

2015年受講登録者  
(452人 最終)

2016年度



■ 大学 ■ 企業・研究所等 ■ その他

2016年受講登録者  
(449人 10/3現在)

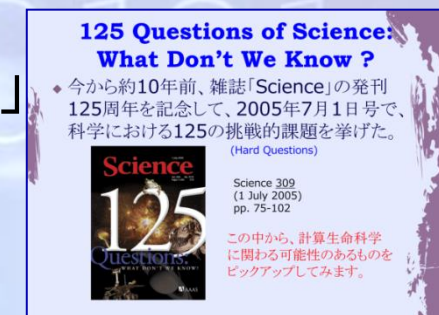
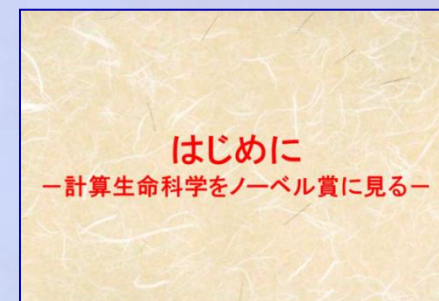


# 「計算生命科学概論」

「計算生命科学の基礎 III」は 昨年度 の講師を9名入れ替えている

## 「計算生命科学概論」での話題提供

- 2014年 理化学研究所 江口至洋先生 PW不要  
「計算生命科学をノーベル賞に見る」
- 2015年 神戸大学 田中成典先生 PWは事務局へお問合せ下さい  
「125 Questions of Science: What Don't We Know ?」  
科学における 125 の挑戦的課題からピックアップ  
Science 309 (1 July 2005) pp.75-102
- 2016年  
「生命科学研究への情報科学技術のインパクト」  
2016年3月 八尾徹 (理化学研究所)  
(ライフサイエンス振興財団 「ライフサイエンスに係る研究開発の将来動向調査」からダウンロード)



←「創薬化学者からの視点」

# 本日の講義内容

1. 遠隔講義「計算生命科学の基礎」について
2. 計算生命科学とは
3. ノーベル賞と計算生命科学
4. 「生命科学研究への情報科学技術のインパクト」から
  - ① データサイエンス
  - ② 構造ベース創薬 (タンパク質構造データベース PDB)
  - ③ 人工知能 (AI)
5. 2016年度シラバス紹介
6. まとめ Q&A

# 「計算生命科学」とは？ ＝ 生命科学 X 計算科学 ？

**生命科学＝ 「主に観察と実験に基礎を置いた学問領域」**

生命現象を生物学を中心に化学・物理学等の基本的な面と、  
医学・薬学・農学・工学・心理学等の応用面とから総合的に研究しようとする学問

ジャック・モノー

1965年にノーベル生理学・医学賞を受賞

実験研究者であり、直観よりも論理的思考を重視

**“立派な理論に支えられていない実験なんか、  
決して信じてはいけないよ”**

**計算科学＝ 理論と実験が協調的に進められる中間点**

「理論に基づく計算を行い、実験を支える。

実験から得られるデータを計算し、解析し、理論を構築する。」

# 「計算生命科学」とは？



HPCI 戦略プログラム 分野1

「予測する生命科学・医療および創薬基盤」

**計算生命科学**は、スーパーコンピュータを積極的に活用することにより、革新的な計測技術から生み出されるヘテロなビッグ・データを効率的に解析し、動的な多階層システムを整合的につなぎ、生命システムを総体として理解することを目的としています。

この結果、生命科学ははじめて予測可能性と制御可能性を獲得することが出来るのです。

理化学研究所 HPCI戦略プログラム分野1 WEBページより抜粋 <http://www.scls.riken.jp/research/>

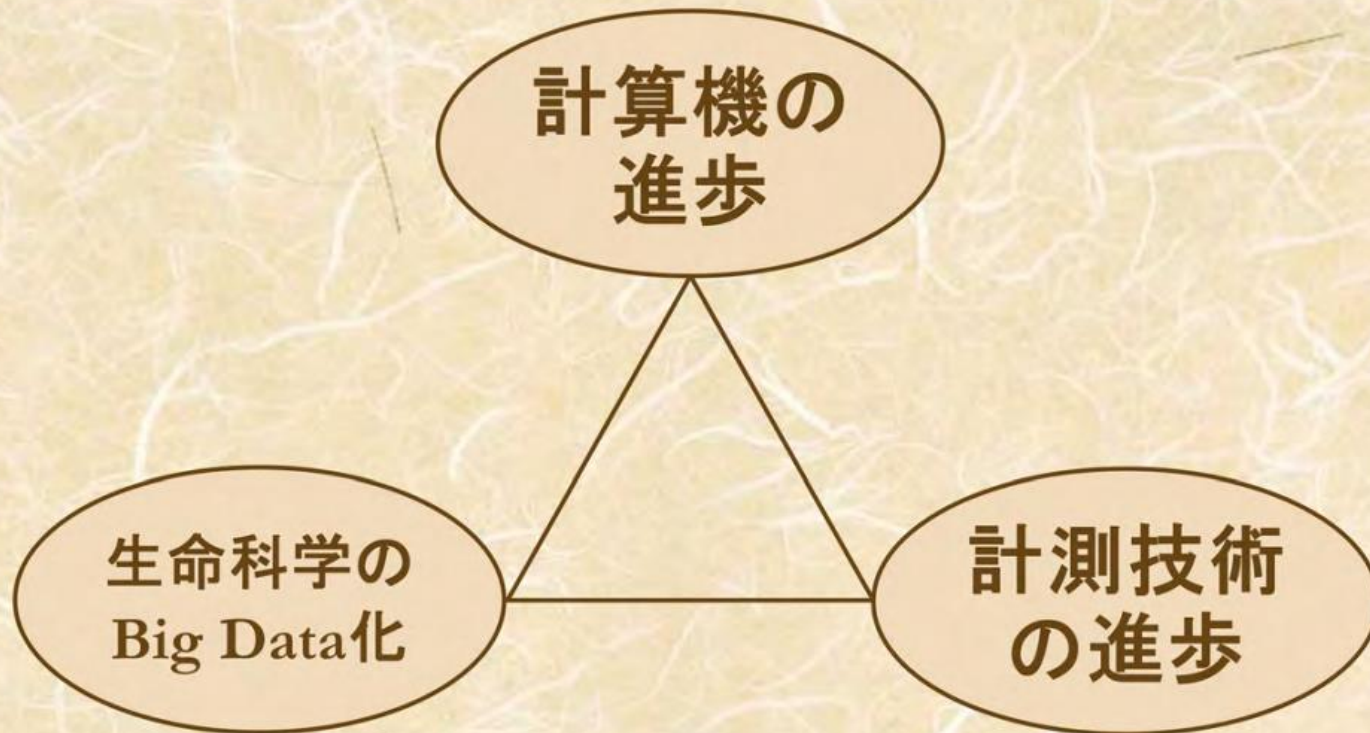
HPCI戦略プログラムは2016年3月31日に終了

→ フラッグシップ2020プロジェクト(ポスト「京」開発プロジェクト)における計算生命科学分野([重点課題1](#)と[重点課題2](#))の研究開発に期待



# 21世紀に計算生命科学が求められる理由

## 3つの要因



# 「計算生命科学」 理学、工学、生物学などの枠を超えた協力関係

この講義は実験生命科学 及び 広く理工系の方々を対象

**生命科学を志す皆さんには、**

意識するとしなにかかわらず、研究室にある実験・計測装置から生み出される大量のデータや、世界的に公開されている膨大な生命科学データベースを基盤に研究を進めることが求められてくると思います。

**理学、工学を志す皆さんには、**

その基礎的な理論を生命科学に生かしていく場がますます広がってきているとともに、生命科学の方々から連携を強く求められてくると思います。

この講義により理学、工学、生物学などの枠を超えて、「理論、計算、観察・実験の協力関係がより一層深まって行くことを願っています。」

# 本日の講義内容

1. 遠隔講義「計算生命科学の基礎」について
2. 計算生命科学とは
3. ノーベル賞と計算生命科学
4. 「生命科学研究への情報科学技術のインパクト」から
  - ① データサイエンス
  - ② 構造ベース創薬 (タンパク質構造データベース PDB)
  - ③ 人工知能 (AI)
5. 2016年度シラバス紹介
6. まとめ Q&A



# ノーベル賞発表週 今週！

## 10月3日～10日

### 計算生命科学をノーベル賞に見る

- ◆ 1963年ノーベル生理学・医学賞：「神経細胞膜の末梢および中枢部分における興奮と抑制に関するイオン機構の発見」(A.L. Hodgkin, A.F. Huxley)；膜の電気回路モデルの提案
- ◆ 2011年ノーベル生理学・医学賞：「自然免疫の活性化に関する発見」(B. Beutler, J.A. Hoffmann)；配列データベース、相同性検索ソフトの活用
- ◆ 2012年ノーベル生理学・医学賞：「成熟した細胞に対してリプログラミングにより多能性(分化万能性)を持たせられることの発見」(J. Gurdon, S. Yamanaka)；遺伝子発現データベース・解析プログラムの活用
- ◆ 2013年ノーベル化学賞：「複雑な化学系のためのマルチスケールモデルの開発」



# 2013年のノーベル化学賞は 計算生命科学に

## The Nobel Prize in Chemistry 2013



Photo: A. Mahmoud  
**Martin Karplus**  
Prize share: 1/3



Photo: A. Mahmoud  
**Michael Levitt**  
Prize share: 1/3

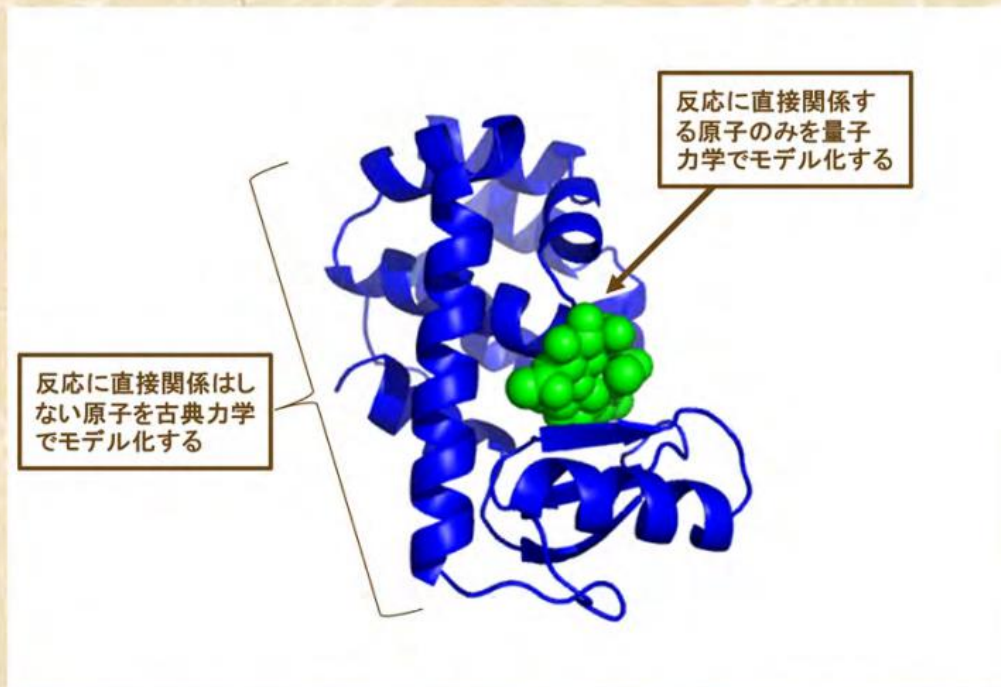


Photo: A. Mahmoud  
**Arieh Warshel**  
Prize share: 1/3

The Nobel Prize in Chemistry 2013 was awarded jointly to Martin Karplus, Michael Levitt and Arieh Warshel *"for the development of multiscale models for complex chemical systems"*.

# 2013年のノーベル化学賞は 計算生命科学に

古典力学と量子力学を統合した研究  
「複雑な化学システムのためのマルチスケール・モデル」



PDB:148Lを用いて作成した。

29

イメージ図： 青色は酵素で、緑色は酵素反応の基質だとします。「複雑な化学システムのためのマルチスケール・モデル」では、反応に直接関係する原子（複数の原子からなりますがそれほど多くはありません）のみを量子力学でモデル化し、「反応に直接関係はしない原子を古典力学でモデル化する」という意味でマルチスケール化して酵素反応の解析を可能としたのです。

# 2013年のノーベル化学賞は 計算生命科学に

## ➤ ノーベル賞の授賞理由

「マルチスケール・モデルは理論と実験の実りある協力関係を作り上げた」

その協力関係により「他の方法では解けない問題をも解く道を切り拓いた」

➤ 計算生命科学がその協力関係の結節点にしていることを示す例であり、ある意味、理論と手を結び、実験とも手を結んでいる、つまり両手が必要。



# 2015年のノーベル賞



## The Nobel Prize in Physiology or Medicine 2015

William C. Campbell, Satoshi Ōmura, Youyou Tu

Share this: [f](#) [G+](#) [t](#) [+](#) [e](#) 2.5K

## The Nobel Prize in Physiology or Medicine 2015

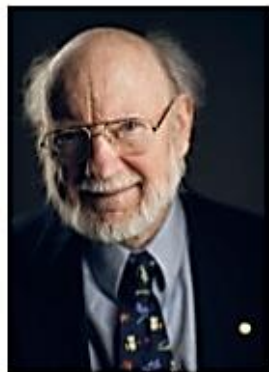


Photo: A. Mahmoud  
**William C. Campbell**  
Prize share: 1/4



Photo: A. Mahmoud  
**Satoshi Ōmura**  
Prize share: 1/4



Photo: A. Mahmoud  
**Youyou Tu**  
Prize share: 1/2

The Nobel Prize in Physiology or Medicine 2015 was divided, one half jointly to William C. Campbell and Satoshi Ōmura *"for their discoveries concerning a novel therapy against infections caused by roundworm parasites"* and the other half to Youyou Tu *"for her discoveries concerning a novel therapy against Malaria"*.



## The Nobel Prize in Chemistry 2015

Tomas Lindahl, Paul Modrich, Aziz Sancar

Share this: [f](#) [G+](#) [t](#) [+](#) [e](#) 2.3K

## The Nobel Prize in Chemistry 2015



Photo: A. Mahmoud  
**Tomas Lindahl**  
Prize share: 1/3



Photo: A. Mahmoud  
**Paul Modrich**  
Prize share: 1/3



Photo: A. Mahmoud  
**Aziz Sancar**  
Prize share: 1/3

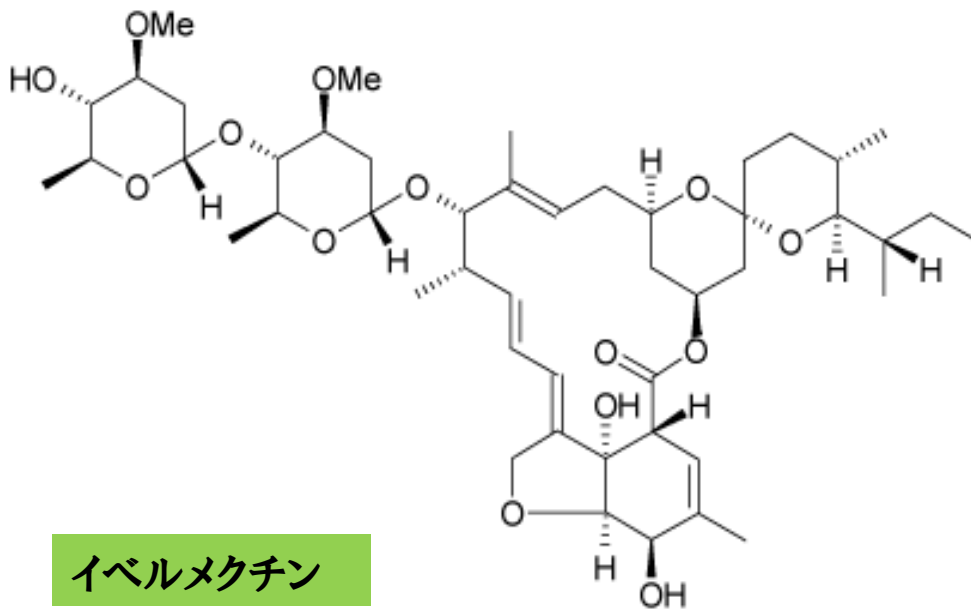
The Nobel Prize in Chemistry 2015 was awarded jointly to Tomas Lindahl, Paul Modrich and Aziz Sancar *"for mechanistic studies of DNA repair"*.



**イベルメクチン  
オンコセルカ症（河川盲目症）治療薬**



大村智 北里大学特別栄誉教授  
2015年ノーベル医学生理学賞



- 世界保健機関はメルク社の協力を得て、アフリカなど寄生虫病に苦しむ地域にイベルメクチンを配布するプログラムを開始。
- メルク社によると、2012年までに延べ10億人以上にイベルメクチンが無償提供された。

# 微生物ゲノムーメタゲノム解析 応用例 菌放線菌 *S.avemectinius* の全ゲノムの情報解析

大村智が、その研究の過程で、抗生物質生産菌放線菌 *S.avemectinius* の全ゲノムの情報解析から、この放線菌が avermectin の他に 37 種の有機化合物を作る遺伝子を持つことを明らかにしています。  
(avermectin → ivermectin)

Complete genome sequence and comparative analysis of the industrial microorganism *Streptomyces avermitilis*.

Ikeda H1, Ishikawa J, Hanamoto A, Shinose M, Kikuchi H, Shiba T, Sakaki Y, Hattori M, Omura S. Nat Biotechnol. 2003 May;21 (5):526-31. Epub 2003 Apr 14.

Genome sequence of an industrial microorganism

*Streptomyces avermitilis*: Deducing the ability of producing secondary metabolites

Satoshi Omura, et al, Yoshiyuki Sakaki, and Masahira Hattori;  
PNAS 98 (21) 12215-12220 Oct.9 2001

# 2016年 ノーベル賞発表スケジュール

*“For the greatest benefit to mankind”*

*Alfred Nobel*

**2016 NOBEL PRIZE  
ANNOUNCEMENTS**

**3-10 October**

<http://www.nobelprize.org/>

## 2016年 ノーベル賞発表スケジュール

(日時はスウェーデン標準時)

生理学・医学:10月3日(月)11:30(日本時間18:30)～

物理:10月4日(火)11:45(日本時間18:45)～

化学:10月5日(水)11:45(日本時間18:45)～

平和:10月7日(金)11:00(日本時間18:00)～

経済学:10月10日(月)11:45(日本時間18:45)～



# トムソン・ロイター引用栄誉賞」 (ノーベル賞予測) 2016年9月21日

2016年の「トムソン・ロイター引用栄誉賞」  
日本人からは3名が受賞

京都大学 本庶佑氏(医学・生理学)  
「プログラム細胞死1 (PD - 1) および  
その経路の解明により、がん免疫療法の発展に貢献」



崇城大学・熊本大学 前田浩氏(化学)  
国立がん研究センター 松村保広氏(化学)  
「がん治療における高分子薬物の  
血管透過性・滞留性亢進(EPR)効果の発見」



トムソン・ロイター引用栄誉賞」について:

2002年より毎年9月の発表が恒例化されている本賞は、学術論文の引用データ分析から、ノーベル賞クラスと目される研究者を選出し、その卓越した研究業績を讃える目的で発表されるものです。15回目となる本年は、日本人研究者3名を含む合計24名が受賞



# 2016年ノーベル生理学・医学賞

発表:10月3日(月)11:30(日本時間18:30)



2016年の受賞者は？

# 本日の講義内容

1. 遠隔講義「計算生命科学の基礎」について
2. 計算生命科学とは
3. ノーベル賞と計算生命科学
4. 「生命科学研究への情報科学技術のインパクト」から
  - ① データサイエンス
  - ② 構造ベース創薬 (タンパク質構造データベース PDB)
  - ③ 人工知能 (AI)
5. 2016年度シラバス紹介
6. まとめ Q&A

## 「生命科学研究への情報科学技術のインパクト」 ーゲノム総合科学、システム生物学、情報科学技術ー

平成 28 年 3 月

八尾 徹

理化学研究所 客員研究員

兼 慶応義塾大学 SFC 研究所上席所員

ライフサイエンス振興財団の依頼により、著者が長年担当してきていたゲノム・遺伝子・タンパク質を起点とする「生命情報科学」の動向を中心に、歴史的・国際的な視点からまとめた報告書(64ページ)。

私見:情報⇔計算



## 報告書にまとめられている課題

1. ゲノム総合科学の展開と情報科学
2. タンパク質構造ゲノミクス・構造生物学の進展と情報科学
3. システムバイオロジー、合成バイオロジーの進展
4. 脳科学の進歩と情報科学、そして人工知能
5. 情報科学の新たな展開－データサイエンス、ビッグデータ

# 21 世紀に入ってから 3つの大きなパラダイムシフト

ゲノム・遺伝子・タンパク質・細胞、生命システムの解明と  
応用に大きな影響を与えつつある要素

- バイオインフォマティクス (生命情報科学)
- 計算生命科学
- システムバイオロジー
- Bio-IT (コンピュータ・インターネット)

これらは生命現象解明の基礎研究のみならず、その健康医療分野の研究と応用についても大きな影響を与えるでしょう。

# 21 世紀に入ってから 3つの大きなパラダイムシフト

## ① ゲノムワイド解析

(ゲノム科学、オミックス、GWAS, タンパク構造ゲノミックス)

## ② 生命システム解明

(システムバイオロジー、マルチスケール解析、合成バイオロジー)

## ③ 情報科学の新展開

(データサイエンス、ビッグデータ、人工知能)



# 情報科学的に見た生命体の特徴

生命体は無機的な物質・物体ではなく、以下の特徴を持つ

- 常時動き、変化していく。個別である。再現性がない。
- 相互作用をしている。集団行動をする。
- 環境の影響を受ける・避ける・働きかける。
- 自己複製・自己修復する。遺伝する——生命体自身が情報を持っている。
- 高等生物・特に人間においては、精神的な情報の扱いが必要である。



# 生命科学研究にとって情報科学は必須

前述のような特徴を持つ生命体の研究に必要なもの

1. 最初に総合的な観察記述情報
2. 次に対象を絞った個々の測定データの収集・処理・解析・蓄積
3. そしてそれらをまとめて総合的な理解に至る

これらのすべての情報を扱い、生命現象の解明を目的とする生命情報科学は、生命科学研究にとって当たり前の存在であり、すべての対象・段階で必須のものとなっています。

# 生命科学研究にとって情報科学は必須

日常的に使うデータベース・ソフトウェア～

非常に大量で複雑な情報を扱う高度な数理解析技術

各種段階での整備・普及促進・先端的な研究開発推進が並行して進める必要性

## 背景となる要因

1. 生命計測技術の進歩によるデータの急増
2. ゲノム解析・ゲノム総合科学の進展
3. 生体部品の解明から生命システムの理解への指向
4. 個人・社会・環境からのデータ収集(自動、高速、大量)に伴うビッグデータの急増

コンピュータの発達・インターネットの普及・オープンサイエンス、クラウドソーシング、データサイエンスの発展が後押し。

これらを踏まえた総合的な施策と人材育成が今後の急務！

# 生命情報科学への期待

## 生命科学と情報科学

### 21世紀 生命科学と情報科学

#### バイオインフォマティクスへの期待

生命情報科学、生命情報学

計算生物学、計算情報生物学

システム生物学、生命情報システム学

「生命現象の解明に情報科学技術を活用する」

「生命現象を情報科学的に理解する」

「生命現象を情報システムとして理解する」

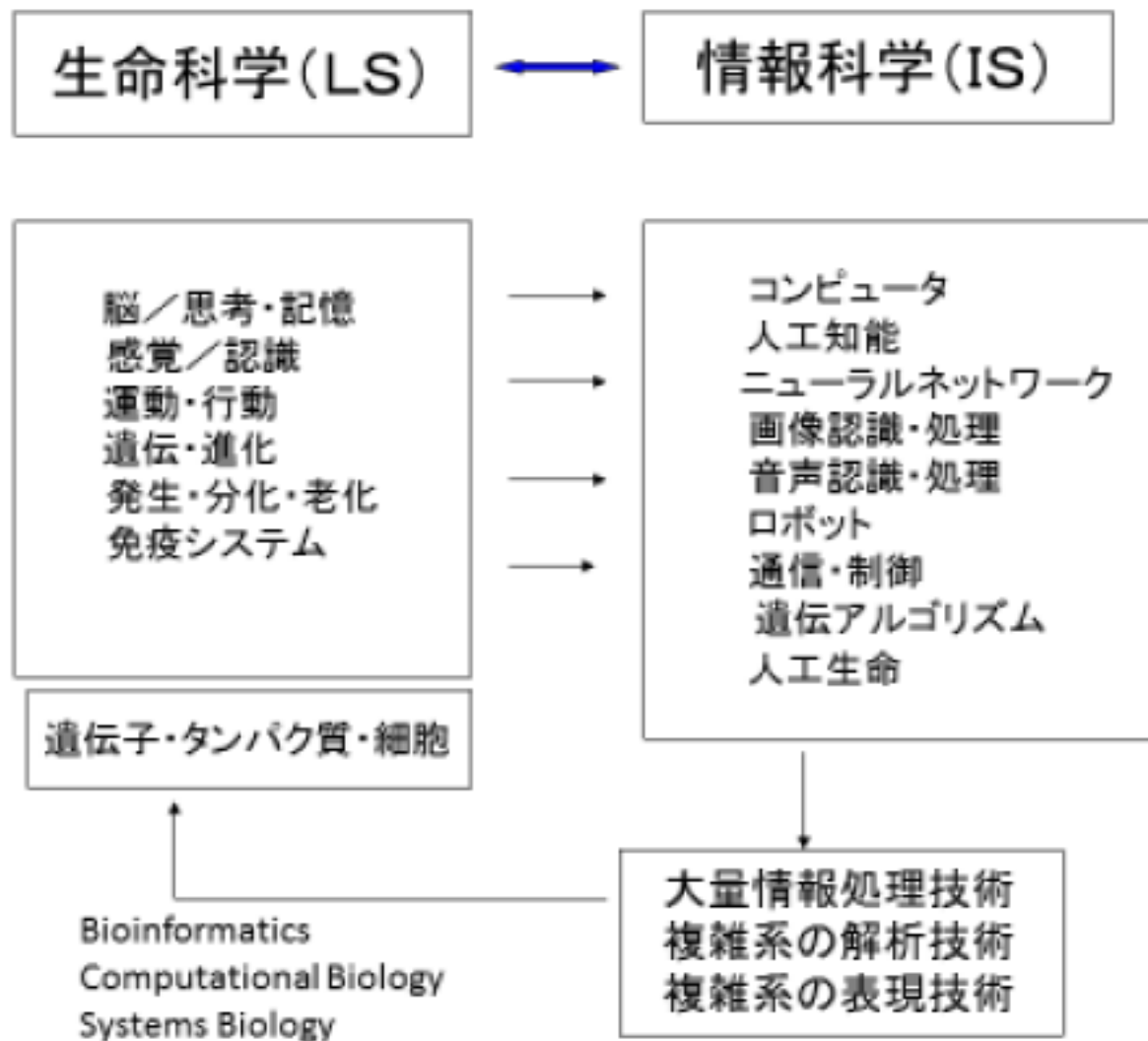
「生命現象から情報科学技術を学ぶ」

「生命体となじむシステムを開発する」

#### 生命現象の

個別的・部分的な解析から総合的・包括的な解析へ

# 生命科学と情報科学の相補関係



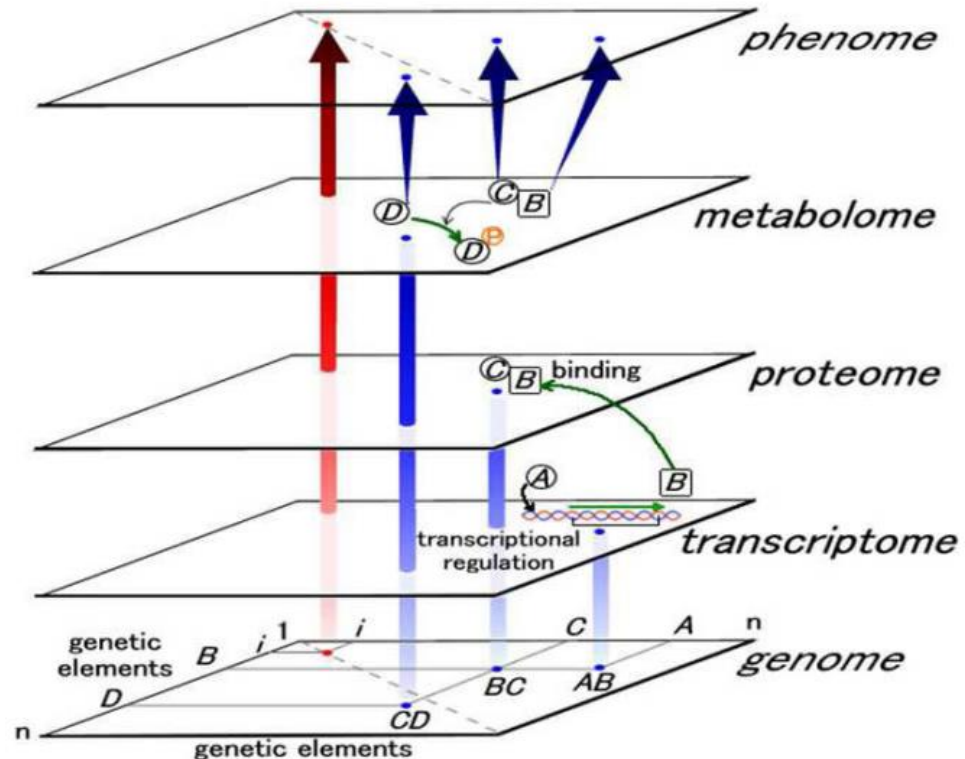


# オミックス平面解析から立体解析へ

■オミックス解析について、重要な概念が2002年頃に理研GSCの和田所長によって提唱。具体的図示化が豊田リーダーによってなされた。

■各オミックス平面の詳細な解析から更に進んで、各平面を縦に貫通するオミックス空間解析が重要というものの。

■2012年以降に東大黒田真也が Trans Omics Analysis という言葉で上記の概念を具体的に発展させ始めています。



理研ホームページ及び東大黒田資料より

オミックス空間解析 トランスオミックス解析

1. ゲノム総合科学の展開と情報科学
2. タンパク質構造ゲノミクス・構造生物学の進展と情報科学
3. システムバイオロジー、合成バイオロジーの進展
4. 脳科学の進歩と情報科学、そして人工知能
5. 情報科学の新たな展開ーデータサイエンス、ビッグデータ

## ① 設立 10 周年を迎えた Google 社の動き。

世界中の膨大な文章情報・画像情報さらには映像情報を蓄え、人工知能を利用した検索サービスをして世界のあらゆる階層の人の知的レベルの向上に役立っていることを示した上で、2018 年までにさらに多様な情報の統合サービスが進むことを多くの識者が述べていました。

Google 社は、その後画像情報のビッグデータ解析によって「ネコ」の顔認識が「教師なし学習法」によってできたことを示しました。

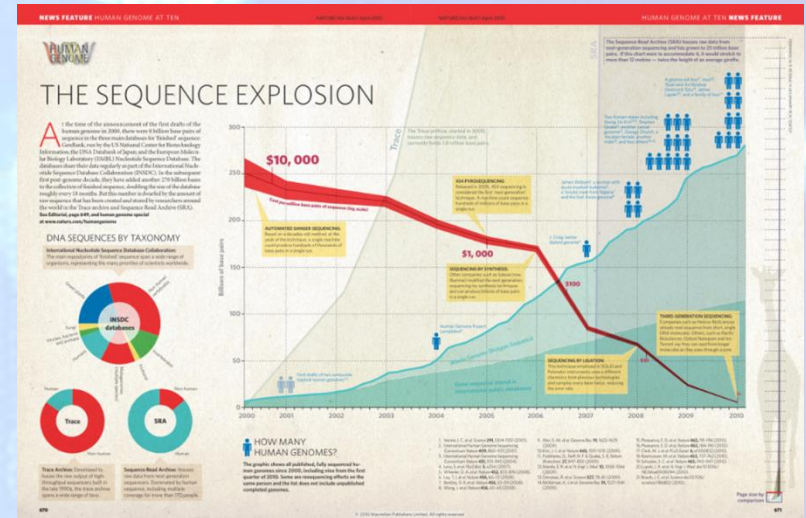
ビッグデータと人工知能技術の融合による成果例。

② ゲノム・DNA の解析技術いわゆる次世代シーケンサーNGS 発展を  
迎えて世界最大のシーケンスセンター Wellcome Trust Sanger  
Institute が ペタセンターになったことを紹介

しかし、その後の NGS の更なる性能向上によって、自身で 20 ペタバイトを保有するだけでなく、ついに外部のクラウドコンピューティングを利用することになっていったこと (Nature June 13, 2013) 同一感染症の国際的大規模疫学のビッグデータによって原因究明の精度が格段に向上した例が紹介されています (Science Nov. 28, 2014)。



ヒトゲノム解読計画がほぼ終了した 2000 年にヒト一人のシンケンシングコストが**\$400M**であったものが、2015 年現在**\$1000**に低下



<https://www.nist.gov/>

これを受けて米国 NIH は 2015 年 1 月に、**100 万人以上を対象とする "Precision Medicine Initiative"** (精密医療または適確医療) を立ち上げました (Obama \$215M)

これは Precision Medicine の実現を目指すためにガンをはじめとする各種疾病の患者の医療記録・遺伝子型・細菌叢・生活環境・生活スタイルのビッグデータを総合的に集めて分析するコホートスタディで世界最大規模です。**これまでの中小コホートスタディ 200 件(数千~1万人対象)を一気に上回るもので、個人別医療の実現に向けて大きく期待されています。**

「生き物の動的な状態の3次元計測とその表示・解析」の急速な進展

静的・平面的な顕微鏡データではなく、**刻々の変化**を捉え、それを直接コンピュータに送ります。その最も基本になるのが、「**細胞**」観察。

単細胞である微生物から多細胞である高等生物まで、細胞研究は大きく変わり始めています。ここでは、**時間変化を含む膨大な4次元データ**が昼夜を問わず集積され、この中から**特徴ある変化**を捉えなければならない。

研究スタイルの大きな変化がおりつつあります。このことは研究発表あるいは教育にも影響を与えています。**(YouTube による発表が急増)**

# 科学における第4のパラダイム データサイエンス

”The Fourth Paradigm -- Data Intensive Scientific Discovery”  
という報告書が、2009 年に Microsoft Research Inc. の Tony Hey ら  
によって公表されたことです。 **科学の発達**が、「実験」・「理論」・「シミュ  
レーション」と共に、「**データ**」によってもたらされる時代になったとの認識  
です。この概念は、科学研究の世界に大きな一石を投じたと言えま  
しょう。(Science 6 Mar. 2009)

|      |                      |
|------|----------------------|
| 第1世代 | 「経験科学」 数千年           |
| 第2世代 | 「理論的科学」 数百年          |
| 第3世代 | 「コンピュータシミュレーション」 数十年 |
| 第4世代 | 「データ科学」              |



## 5-4 ビッグデータ時代における考え方

# データに対する考え方を変える必要

ビッグデータを加速する3要因と扱いの変革3点は下記の通り。

ビッグデータがこのように加速してきた大きな要因

- (1) インターネットの発展による人・物・情報のつながりの普及 (ICT)
- (2) 各種センサーの発達と発信元の多様化 (IoT)
- (3) コンピュータ能力(記憶容量・スピード)の増大と、解析技術(アナリティクス、特に人工知能)の進歩

その上に、データの扱いが大きく変わってきました。

- 第1 データを選別・抽出するのではなく、全データを扱う。
- 第2 仮説ドリブンでなくデータドリブンのアプローチをとる。
- 第3 因果関係でなく、相関関係を重視する。

「ビッグデータの正体」V・M ショーンベルガー& ケネス・クキエ著 2013

斎藤栄一郎訳 講談社 2013 年 5 月刊



## 5-5 健康医療分野でのビッグデータ マッキンゼー報告書

### 2013 年マッキンゼー社の調査報告書

The Big Data Revolution in Healthcare” A Study by Mckinsey & Co. by Drs. B.Kayyali, D.Knott and S.Kuiken April 2013

多様なソースからの患者データを疾病間にまたがって総合解析した結果の診断治療の成功例が数多く紹介されており、さらに今後の展望が述べられている。

最近の情報技術の進歩によって、従来分散していた患者データを多様なソース(病院・製薬企業・保険機関ほか)からまとめることによって、総合的な分析ができるようになった。特に疾病ごとに分かれていたデータを連携することによって新たな発見につながる例もみられ、大きな価値を見出している。

**Kaiser Permanent 社:** HealthConnect というコンピュータシステムに、多くの医療機関からのデータを集約している。その分析によって、心臓病関連で開発費および診療費の約 10 億ドル削減に成功した。

**Blue Shield 社:** NantHealth 社と共同で、統合システムによって医師・病院・健康プランの個人別対応ができるようになり、多くの疾病について効果を上げている。

**AstraZeneca 社:** HealthCore 社(WellPoint's data and analytics の子会社)と共同で、独自の臨床試験データを含む実世界データを用いて慢性疾患・生活習慣病に最も効果的な治療法を開発している。

統合ビッグデータ解析によって 2010 年以降でも 200 以上の新規ビジネスが生まれているとしている。

# ビッグデータ、データサイエンス 5-7日本の取り組み

## 1) 情報通信政策

2014 年版「情報通信白書」の第3章で「データが切り拓く未来社会」を論じています。ビッグデータ・オープンデータ・パーソナルデータの日本全体の総合的な現況と予測について社会の動きと情報通信政策が分かります。たいへん大事な資料で、その社会的環境の中で、科学技術・健康医療分野の展開を考えるべきでしょう。

## 2) 関係者－論文・総説

更に、下記のような論文・総説が、日本のバイオインフォマティクス指導者から、次々と出されるようになってきました。

「ビッグデータのゲノム・オミックス医療への応用」 田中 博 2014 年 12 月 5 日

「ライフサイエンスにおけるビッグデータの解析」 松田秀雄 2014 年 9 月 12 日

「ますます多様化・大規模化する生命情報に求められる情報技術とは？」

高木利久 2012 年 11 月 9 日

“Tackling Biomedical Big Data -Cancer System” 宮野 悟 2013 年 5 月 31 日

「ビッグデータ応用：ブームから挑戦へ」JST/CREST 田中 譲 2015 年 10 月 28 日

# 日本の取り組み： ビッグデータの利活用のための専門人材育成

## ビッグデータの利活用のための専門人材育成について

平成27年7月30日

大学共同利用機関法人 情報・システム研究機構

ビッグデータの利活用に係る専門人材育成に向けた産学官懇談会

大学共同利用機関法人 情報・システム研究機構 情報公開 会議報告書からダウンロード

[http://www.rois.ac.jp/open/pdf/bd\\_houkokusho.pdf](http://www.rois.ac.jp/open/pdf/bd_houkokusho.pdf)

<http://www.rois.ac.jp/open/index06.html>



# データサイエンティストに求められるスキルセット

## データサイエンティストに求められるスキルセット

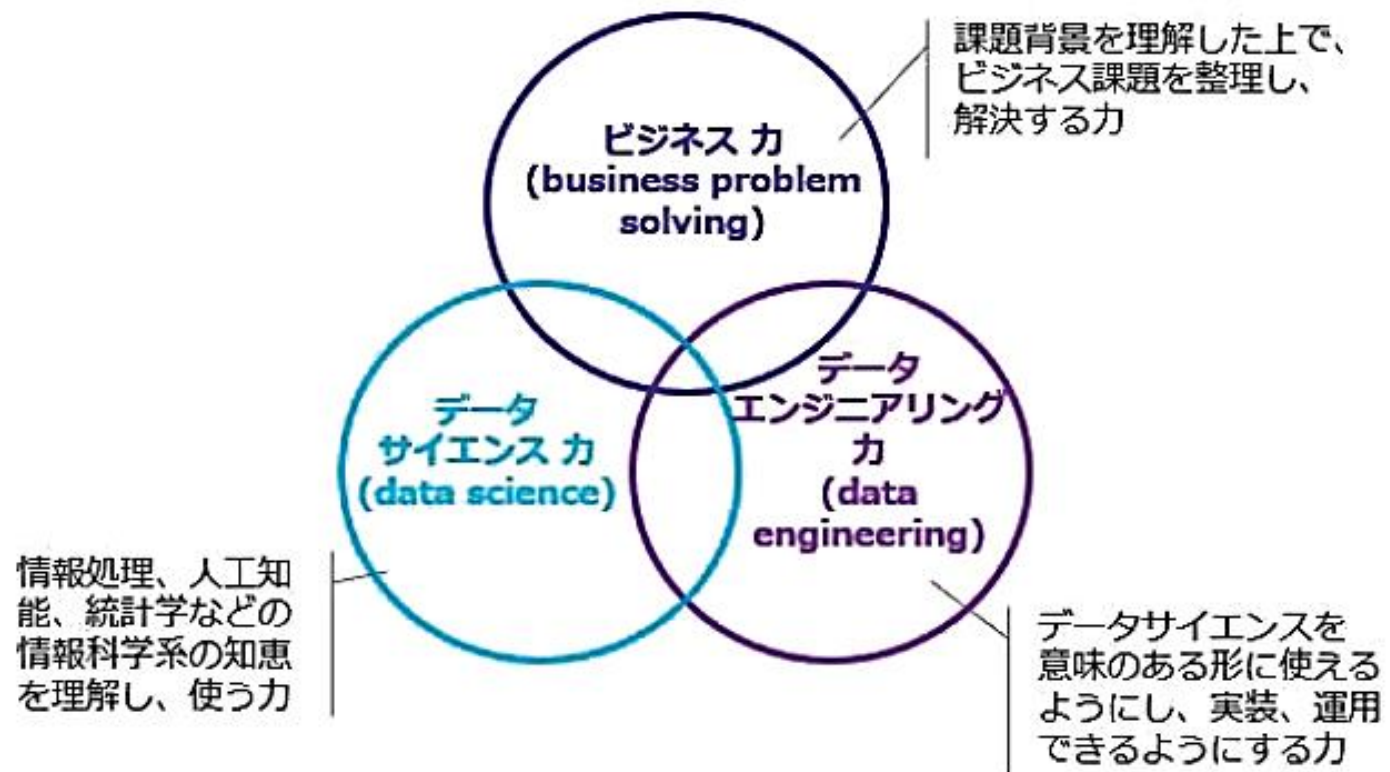


図 1. データサイエンティストに求められるスキルセット  
(データサイエンティスト協会スキル定義委員会の報告書より転載)



# 人材育成レベルと 毎年の育成目標人数

データサイエンティスト協会が定めたスキルレベル (2014年12月)

業界を代表するレベル  
Senior Data Scientist

棟梁レベル  
(full) Data Scientist

独り立ちレベル  
Associate Data Scientist

見習いレベル  
Assistant Data Scientist

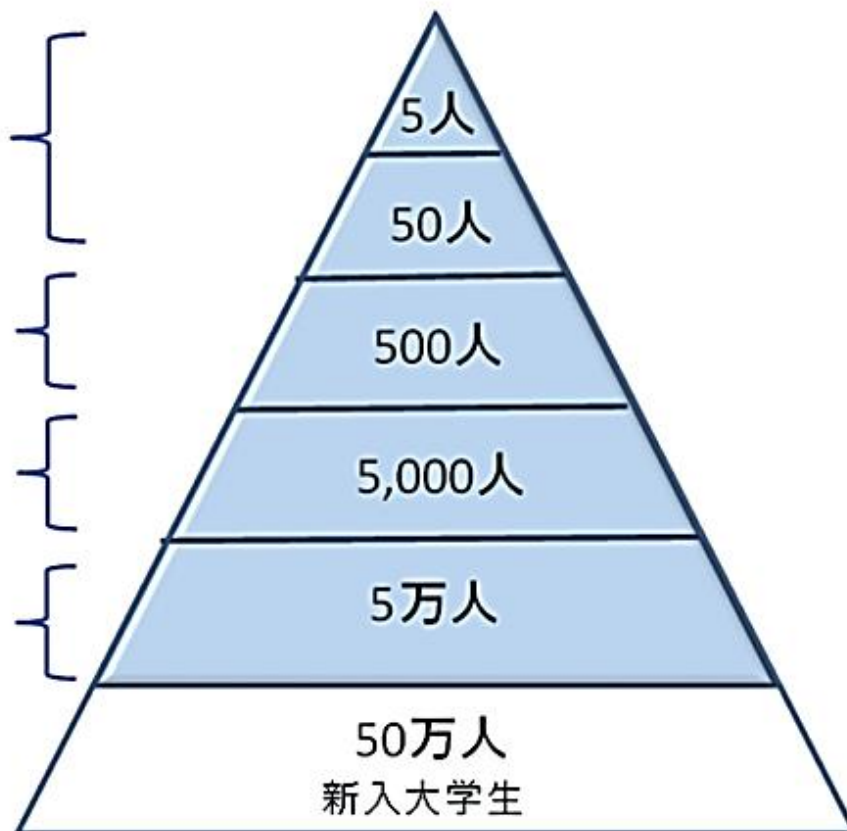


図 2. 育成レベルと、毎年の育成目標人数

# 棟梁レベルの人材不足

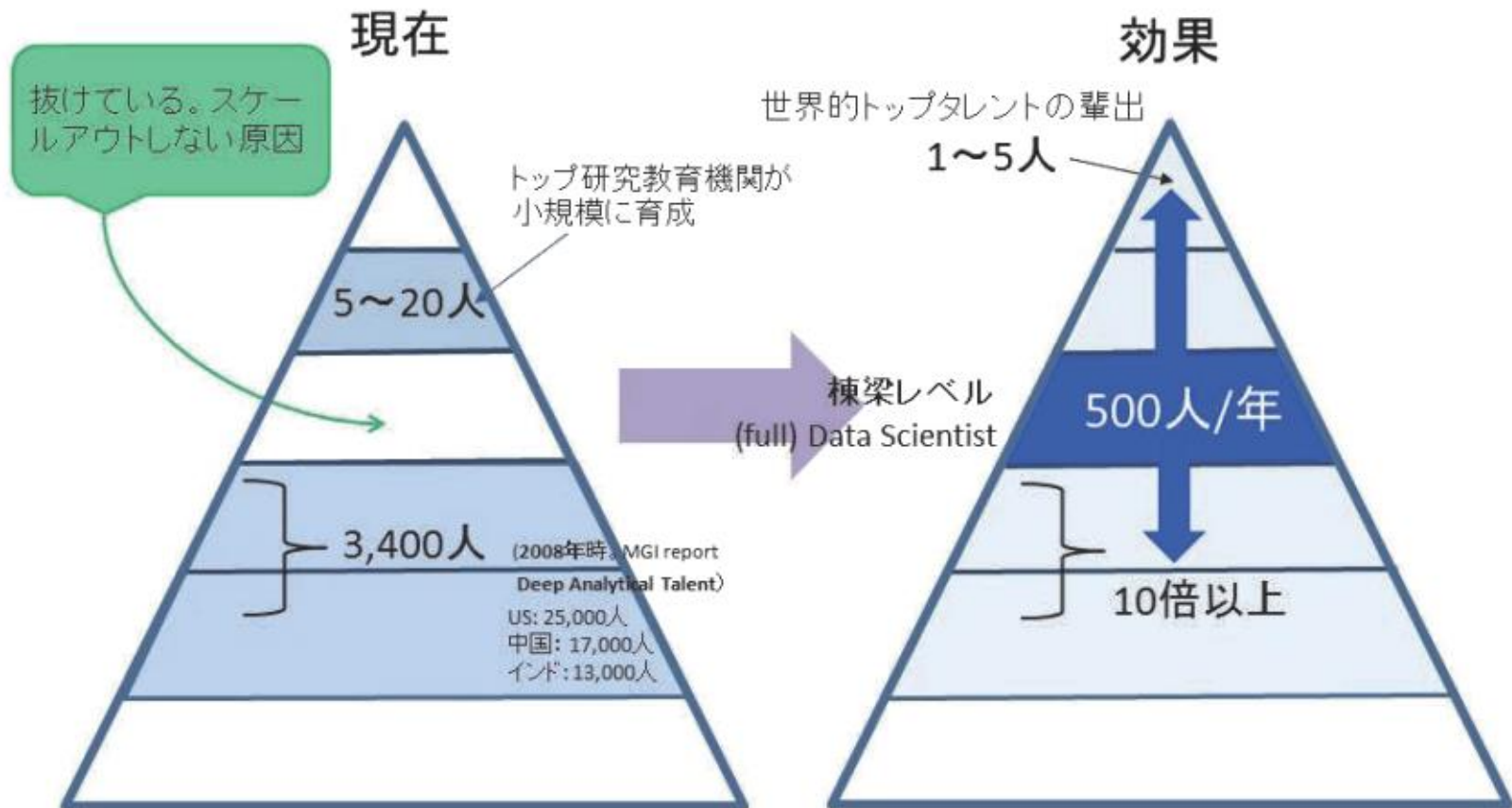


図 3. 育成が遅れ、最も欠けている棟梁レベル(左)と育成実現時の波及効果(右)

# 海外におけるデータサイエンス関連教育プログラム

表 1 海外におけるデータサイエンス関連教育プログラムの数

|              | US         | GB        | IE       | FR       | NL       | ES       | Others    | Total      |
|--------------|------------|-----------|----------|----------|----------|----------|-----------|------------|
| Bachelors    | 15         | 2         | 0        | 1        | 0        | 0        | 2         | 20         |
| Masters      | 126        | 37        | 7        | 5        | 6        | 5        | 28        | 214        |
| Doctorate    | 11         | 1         | 0        | 0        | 0        | 0        | 2         | 14         |
| Certificate  | 29         | 0         | 1        | 0        | 0        | 0        | 1         | 31         |
| <b>Total</b> | <b>181</b> | <b>40</b> | <b>8</b> | <b>6</b> | <b>6</b> | <b>5</b> | <b>33</b> | <b>279</b> |

Colleges with Data Science Degrees (<http://datascience.community/colleges>) より編集

# 文科省が大学にコース

## ビッグデータ分析のプロ、国挙げて育成 文科省が大学にコース

2016/6/28 12:30 | 日本経済新聞 電子版

文部科学省は2017年度から、企業が持つビッグデータを分析して事業戦略などに生かす専門家の育成に乗り出す。10前後の大学にコースを設け、博士課程の学生などから毎年250人を養成する。統計学に精通し、高度な知識を持つ専門家は毎年500人ほど育成が必要とみており、大手だけでなく地方企業への人材供給につなげる。



# 大手外資企業 データサイエンティスト育成の取り組み

## 1. チーム・データサイエンティスト

←スーパー・データサイエンティスト

- ① ビッグデータ収集(測定)
- ② ビッグデータ解析(手法)
- ③ ビジネス(ニーズ・事業展開)

## 2. 人材育成・確保

- ① 社内教育
- ② インターン受け入れ(新規雇用)
- ③ 外部からの中途雇用 → 企業文化変革への対応

→DS関連の人材育成のニーズ増大

1. ゲノム総合科学の展開と情報科学
2. タンパク質構造ゲノミクス・構造生物学の進展と情報科学
3. システムバイオロジー、合成バイオロジーの進展
4. 脳科学の進歩と情報科学、そして人工知能
5. 情報科学の新たな展開ーデータサイエンス、ビッグデータ

## 2. タンパク質構造ゲノミクス・構造生物学の進展と情報科学

### 2-1 タンパク質立体構造解析の進展と成果

- ① その重要性と解析技術
- ② 構造ゲノミクスプロジェクト(日本、米国)  
(ATP 合成酵素、リボゾーム、チトクローム C, タンパク質合成、アディポネクチン他)

### 2-2 構造バイオインフォマティクスの進歩 データベース、分子計算、医薬分子設計

- ① 立体構造データベースとその解析 PDB, SCOP
- ② 立体構造予測・モデリング CASP (構造予測コンテスト)  
インターネットコンテスト利用による技術開発の加速(事例)

- ③ 分子シミュレーション MD (Molecular Dynamics)

マルチスケールモデリング、スーパーコンピューティング

- ④ 構造ベース医薬分子設計 SBDD (Structure Based Drug Design)

### 2-3 構造解析・設計技術の最近の進歩

- ① X線自由電子レーザー法による動的状態測定
- ② 低温電子顕微鏡 Cryo-EM の新たな展開
- ③ 複合法による構造モデリング
- ④ タンパク質構造ダイナミクスの解析
- ⑤ 人工アミノ酸、人工塩基対による新規タンパク質

## 2-2④ 構造ベース医薬分子設計 (Structure Based Drug Design)

◆ 阻害あるいは活性化対象のタンパク質立体構造に対し医薬候補がどのように作用するかを解析し、効果の高い化合物を探し出す方法。

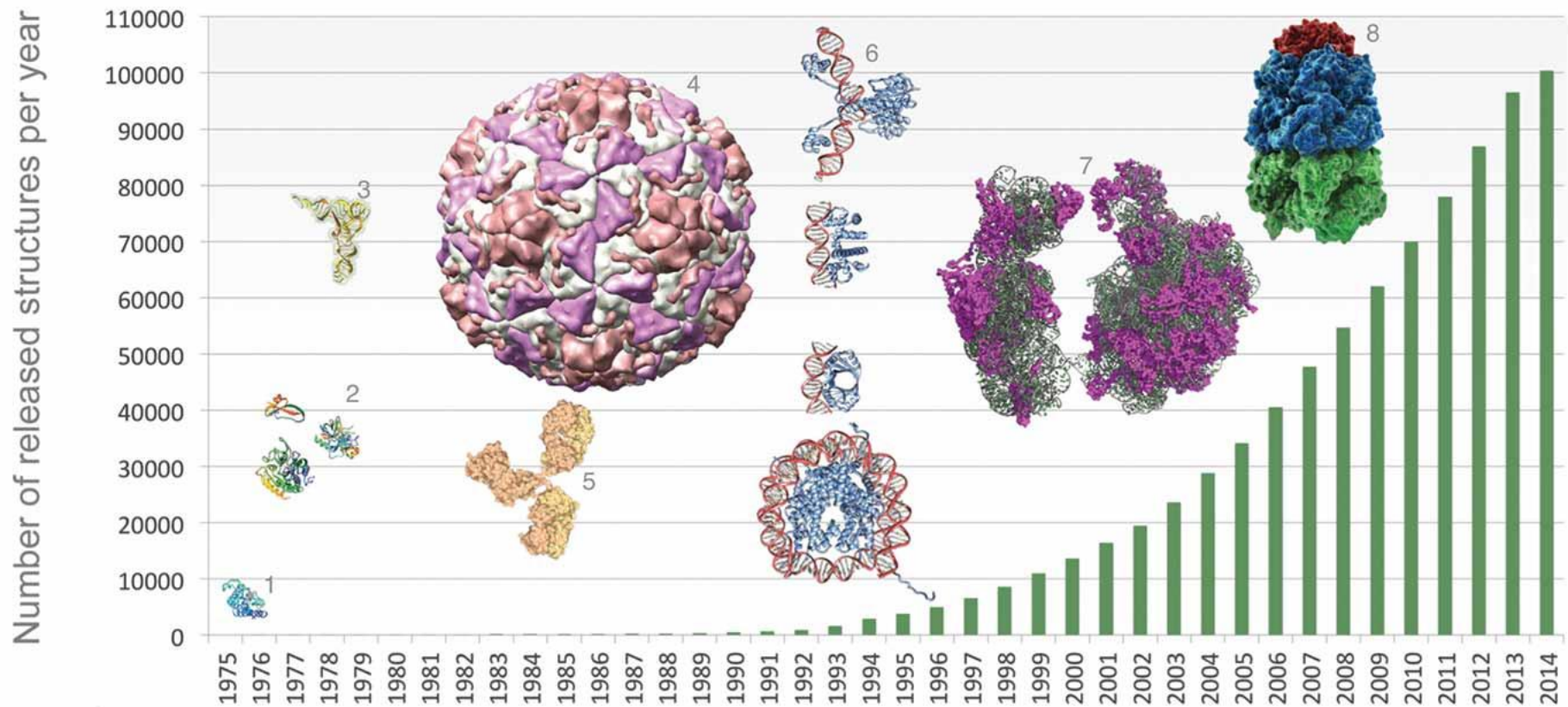
◆ SBDD (Structure Based Drug Design) という手法が1990年頃から始まる。FBDD (Fragment Based Drug Design) やLBDD (Ligand Based Drug Design) 法も含めて、すでに医薬開発の現場で日常的に使われるレベルまで一般化。

◆ 立体構造が解かれたタンパク質数が増えてきたのと、ホモロジーモデリング手法の発達に依ることです。2015年11月に理研横浜研究所で開かれたスウェーデンとの合同シンポジウムでもこの分野の先端的研究状況が紹介されました。(本間、広川 ほか)

→ PDB (Protein Data Bank) のデータについて



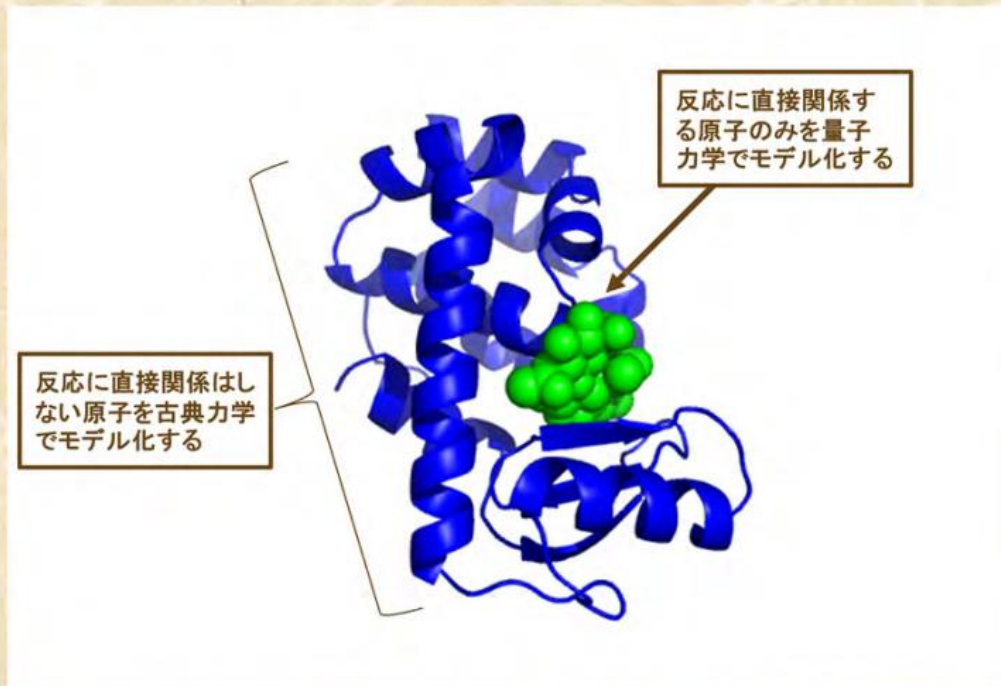
# Protein Data Bank (PDB)



PDB: 蛋白質(タンパク質)と核酸の3次元構造の構造座標(立体配座)を蓄積している国際的な公共のデータベース。 構造バイオインフォマティクス研究共同体 (RCSB; Research Collaboratory for Structural Bioinformatics) のプロジェクトの一つ。現在11万件以上のデータが登録。

# 2013年のノーベル化学賞は 計算生命科学に

古典力学と量子力学を統合した研究  
「複雑な化学システムのためのマルチスケール・モデル」



PDB:148Lを用いて作成した。

2014年江口至洋先生 講義スライドより抜粋

タンパク質分子動力学シミュレーションMD (Molecular Dynamics) が、タンパク質の詳細な動きを見るためにX線結晶構造だけでは分からない作用機序が理解できます。この方法は最近のコンピュータ能力の進歩により大きく複雑な系に対しても使われるようになってきました。Martin Karplus, Michael Levitt らはこの原理と方法の研究でノーベル化学賞 2013を受賞しています。

「生命科学研究への情報科学技術のインパクト」2016年 八尾徹



# タンパク質発現から創薬化学まで

**Molecular / Structural  
Biologist**

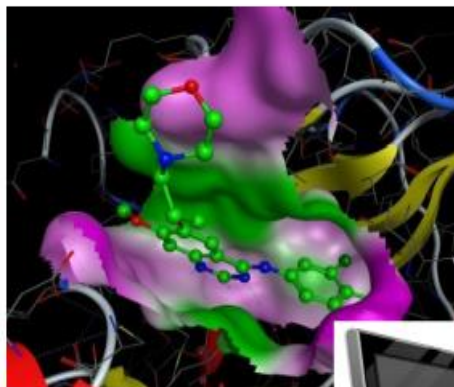


標的タンパク質発現

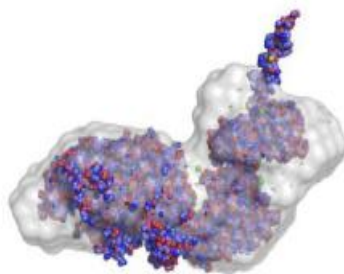


単結晶構造解析

**Computational  
Chemist**

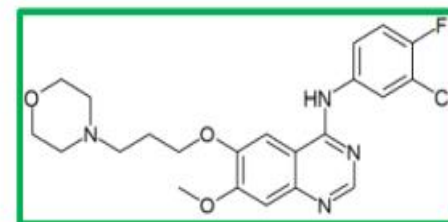


構造最適化



NMR / 小角散乱

**Medicinal  
Chemist**



候補化合物・医薬品



バーチャルスクリーニング  
ドッキングスタディ

構造・活性  
データ

|      |    |    |     |   |   |         |         |          |      |       |
|------|----|----|-----|---|---|---------|---------|----------|------|-------|
| ATOM | 1  | H  | OSY | A | 1 | 48.6942 | 51.0338 | 101.2884 | 0.01 | 27.20 |
| ATOM | 2  | CA | OSY | A | 1 | 45.6940 | 50.1330 | 100.3885 | 0.01 | 26.09 |
| ATOM | 3  | O  | OSY | A | 1 | 46.6932 | 49.6948 | 101.3088 | 0.01 | 26.80 |
| ATOM | 4  | O  | OSY | A | 1 | 46.6935 | 50.222  | 102.383  | 0.01 | 26.91 |
| ATOM | 5  | N  | SEB | A | 2 | 47.283  | 48.5316 | 100.951  | 1.00 | 26.26 |
| ATOM | 6  | CA | SEB | A | 2 | 48.277  | 47.777  | 101.762  | 1.00 | 26.17 |
| ATOM | 7  | C  | SEB | A | 2 | 49.214  | 47.777  | 100.945  | 1.00 | 26.21 |
| ATOM | 8  | O  | SEB | A | 2 | 49.214  | 47.777  | 102.850  | 1.00 | 26.31 |
| ATOM | 9  | CB | SEB | A | 2 | 46.277  | 46.777  | 102.404  | 1.00 | 27.09 |
| ATOM | 10 | OG | SEB | A | 2 | 46.147  | 46.147  | 101.270  | 1.00 | 27.05 |
| ATOM | 11 | H  | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 12 | CA | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 13 | C  | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 14 | O  | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 15 | CB | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 16 | OG | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 17 | MD | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 18 | CD | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |
| ATOM | 19 | CE | SEB | A | 3 | 46.147  | 46.147  | 100.609  | 1.00 | 27.44 |

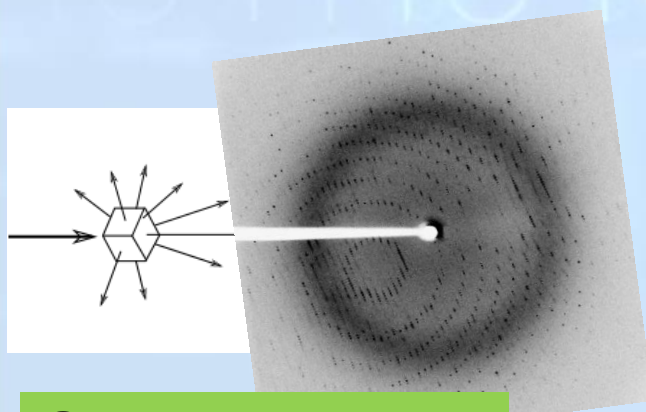
# X-線回折→電子密度マップ→立体構造解析 データ→モデル



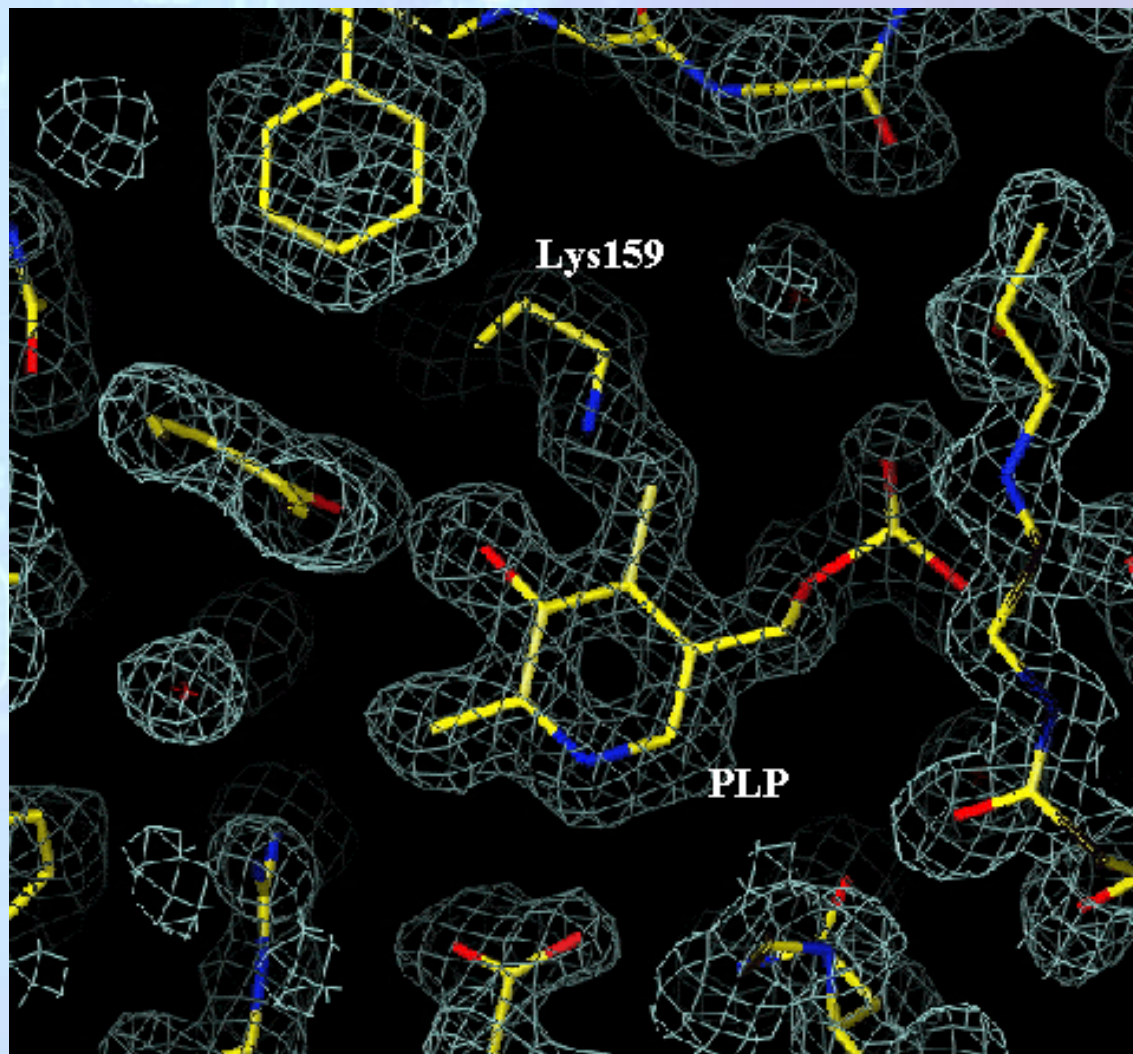
① タンパク質結晶化



② SPring-8



③ 単結晶X線回折像



研究目的に合ったデータの質 モデル？



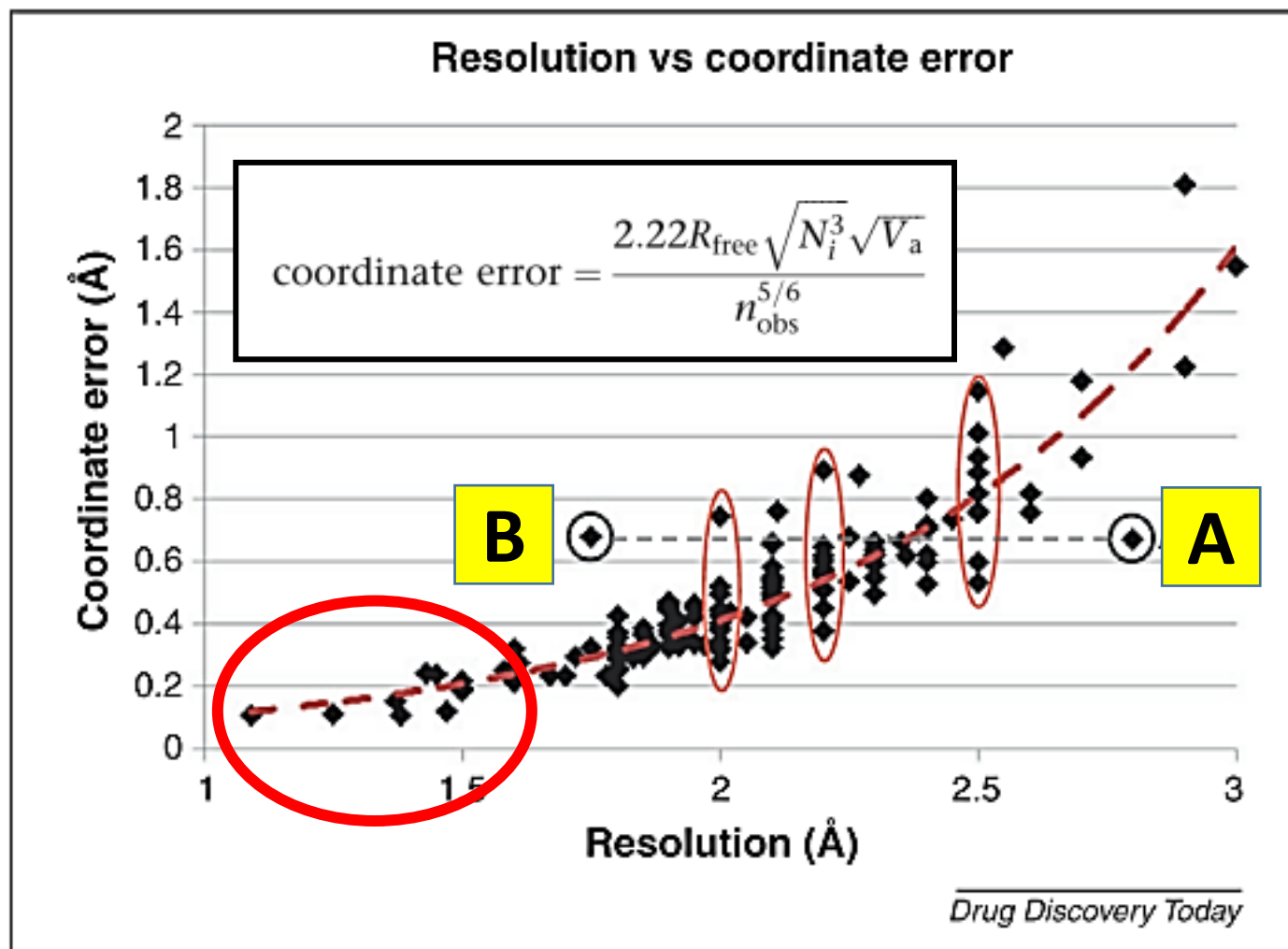
# PDB\_REDO: constructive validation, more than just looking for errors

Robbie P. Joosten,<sup>a,\*</sup> Krista Joosten,<sup>a</sup> Garib N. Murshudov,<sup>b</sup> and Anastassis Perrakis<sup>a</sup>

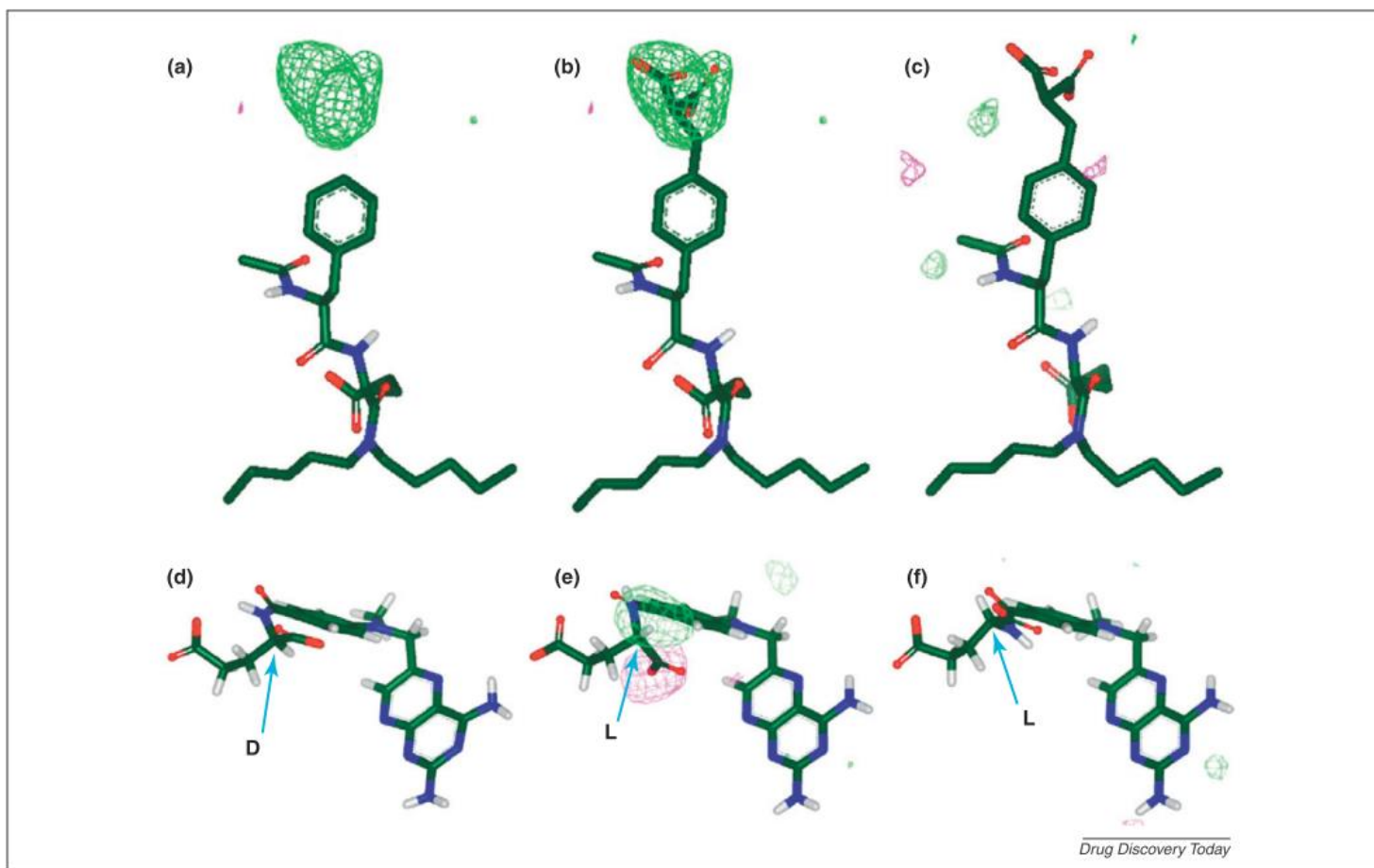
Acta Crystallogr D Biol Crystallogr. 2012 Apr 1; 68(Pt 4): 484-496.

*PDB\_REDO* works fully automatically without the need for intervention by a crystallographic expert. The pipeline was tested on 12 000 PDB entries and the great majority of the test cases improved both in terms of crystallographic criteria such as  $R_{\text{free}}$  and in terms of widely accepted geometric validation criteria. It is concluded that *PDB\_REDO* is useful to update the otherwise 'static' structures in the PDB to modern crystallographic standards. The publically available *PDB\_REDO* database provides better model statistics and contributes to better refinement and validation targets.

# Essential considerations for using protein-ligand structures in drug discovery.



# Essential considerations for using protein-ligand structures in drug discovery.

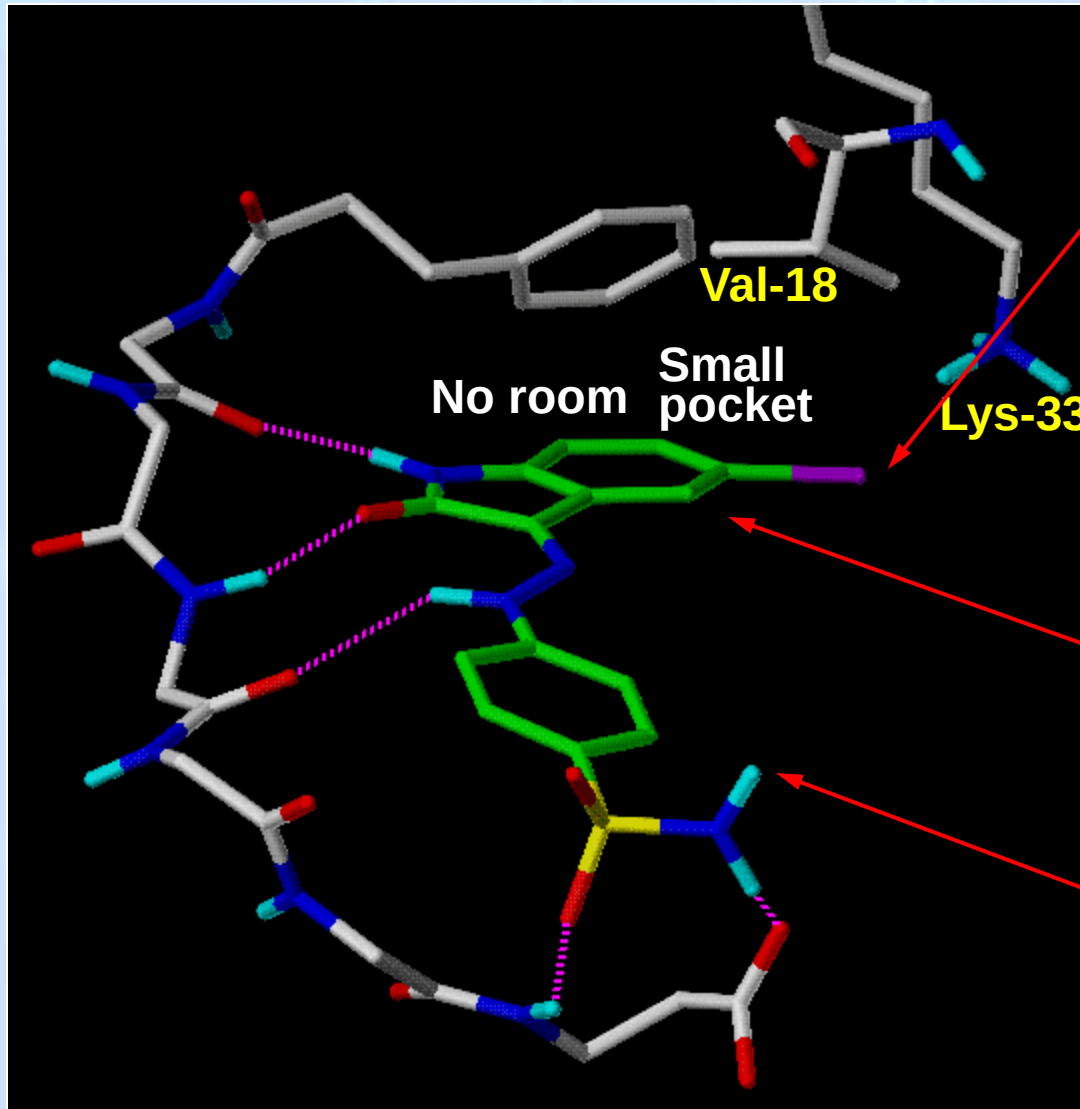


Warren GL1, Do TD, Kelley BP, Nicholls A, Warren SD.  
*Drug Disc. Today*, **12**, 1270 (2012)

**A Highly Trustworthy Protein-Ligand Structure**

**DatabaseOpenEyeIridium:** <http://www.eyesopen.com/database-iridium>

# 酵素阻害剤のドッキングモデル



Lys-33との  
水素結合形成の可能性

疎水相互作用の可能性

溶媒相互作用の領域。化合物の活性に影響を与えないで物性改良の手がかり

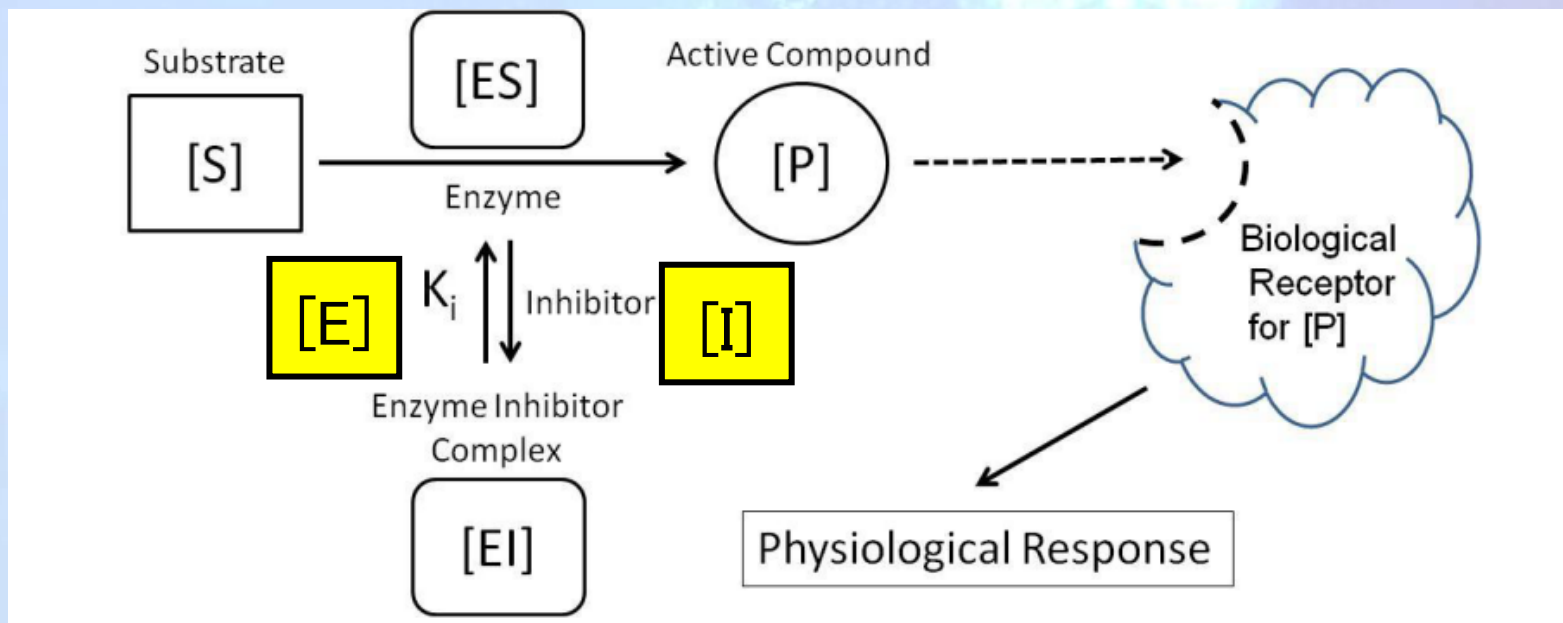


# 競合阻害剤の $K_i$ 値 と $IC_{50}$ 値について

## ”The $IC_{50}$ Concept Revisited”

Caldwell GW, Yan Z, Lang W, Masucci JA.

Curr Top Med Chem. 2012 Jun 1;12(11):1282-90.



$$K_i = \frac{[E]_f [I]_f}{[EI]}$$

$$IC_{50} = K_i \left( 1 + \frac{[S]_b}{K_s} \right)$$

## ～製薬企業での個人的経験～

■Biologist

バケツ一杯



■Medicinal Chemist 耳かき一杯



■Computational Chemist ??

1. ゲノム総合科学の展開と情報科学
2. タンパク質構造ゲノミクス・構造生物学の進展と情報科学
3. システムバイオロジー、合成バイオロジーの進展
4. 脳科学の進歩と情報科学、そして人工知能
5. 情報科学の新たな展開ーデータサイエンス、ビッグデータ

## 4. 脳科学の進歩と情報科学、そして人工知能

脳科学は生命科学の中でも際立って広範な研究分野。

分子・細胞レベルから精神・意識などに及びます。

通常の科学における要素還元主義のアプローチだけでは到達できそうもない対象を含む。

### 4-2 脳科学研究 EU, 米国, 日本

① EU の Brain Project ② 米国 BRAIN Initiative ③ 日本の脳科学総合研究

### 4-3 意識の統合情報理論

意識はいつ生まれるのか、トノーニの著書紹介、意識の統合情報理論

覚醒・意識・記憶 オプトジェネティクス(井ノ口・利根川)

脳とグリア細胞(工藤)

### 4-4 人工知能研究

その歴史・現況・展望

日本の近況 人工知能研究センター設立、産業界の動き、AI 利用成功事例 など

辻井、山川・一杉・高橋・松尾、トヨタ・IBM・楽天・リクルート、奥野・三宅

計算能力・記憶能力代替から思考能力・発見へー北野提唱



## 4-4 人工知能研究

元々、コンピュータ開発の目的は、人間の計算能力・記憶能力・思考能力の模倣と代替にあり、**人間の脳のメカニズムの解明はコンピュータ開発と不可分**のことでした。

しかし、人間の脳はあまりにも高度で、まだ解明できていないことが山積しています。その中であって、**計算能力および記憶能力はコンピュータがほとんど代替可能なレベル**に達し、世の中の様々な局面に充分利用されるようになってきました。

最近の人工知能と呼ばれるものは、この範囲のものが多いです。計算能力や記憶能力を組み合わせれば、かなりのレベルの能力を持つ「人工知能」は作れますし、これからは多種多様な人工知能製品が出回るようになるでしょう。

# 最近の日本の人工知能関連の動き

- 1) 2015 年 5 月に、産業技術総合研究所にAIセンターが設立
- 2) 2016 年度に、理化学研究所に AIIP (Advanced Integrated Intelligence Platform Project) が設立される予定。
- 3) 2015 年 8 月に、NPO 法人全脳アーキテクチャーイニシャティブ (WBAI) が設立されました(世話人: **山川宏**・一杉裕志・高橋恒一・松尾豊ほか)。 脳全体のアーキテクチャーに学んだ汎用人工知能の研究を目指す。
- 4) 2016 年 1 月、トヨタ自動車(株)が米国シリコンバレーに人工知能研究所 (Toyota Research Inst.)を設立。更に、楽天はボストンとシンガポールに人工知能の研究拠点を設立。リクルートは、研究所を新たにして総合人材サービス領域での人工知能研究を加速。
- 5) 2015 年 12 月 14 日に、日本オミックス医療学会が「AI 創薬シンポジウム」を開催され、(中略)「ケミカルゲノミクス情報に基づくインシリコ創薬システム (CGBVS 法) を駆使して活性化化合物構造を生成する AI の成功例が示される (**奥野恭史**)。

# IBM Watson コグニティブ・ビジネス

## BM Watsonプロジェクトは「医療」から始まった 日本IBMの武田浩一氏に聞く

2015/03/17 00:00 日経デジタルヘルス の抜粋

2011年2月に米国のクイズ番組「Jeopardy!」でチャンピオンを破ったことで一躍有名になった、米IBM社のコグニティブコンピューティング技術「IBM Watson（以下、Watson）」の開発には、日本IBMも携わった。同社は2015年2月、Watsonの日本市場での展開に向けてソフトバンクテレコムと提携。2社が共同で、日本での市場開拓やエコシステムの構築を目指す。

<http://techon.nikkeibp.co.jp/article/FEATURE/20150316/409350/?ST=health&P=1>

### 医療への事業展開

- 新薬・効果的な治療方法の発見
- がん診断支援
- 次世代ヘルスケア



<http://cs.utdallas.edu/ibmwatson/>





# IBM Watsonを「10年/1000億円」の創薬に 第一三共が導入

ロボット新産業2016——デジタルヘルスへの本格活用、始まる

**IBM Watsonを「10年/1000億円」の創薬に 第一三共が導入**

2016/02/24 04:00 大下 淳一＝日経デジタルヘルスから抜粋

**「創薬の“匠（たくみ）”と呼ばれてきた研究者の英知や勘を、Watsonに持たせる。**これにより、開発した薬は必ず製品化に成功するようにしたい」（第一三共 執行役員 研究開発本部長 兼 研究開発企画部長の赤羽浩一氏）。

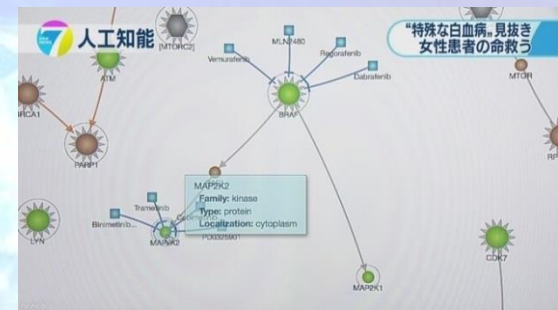
2016年2月18日に日本語版の提供が始まった米IBM社のコグニティブ・コンピューティング・システム「IBM Watson」）。日本の製薬企業として、その利用にいち早く手を挙げたのが第一三共だ。2月18日に日本IBMとソフトバンクが開催したWatson日本語版の発表会では、顧客企業の1社として登壇。新薬開発を統括する赤羽氏が、導入の狙いを語った。



# 人工知能 病名突き止め患者の命救う 国内初か 「IBM Watson」

東京大学医科学研究所が導入した2000万件もの医学論文を学習した人工知能が、専門の医師でも**診断が難しい特殊な白血病を僅か10分ほどで見抜き、治療法を変えるよう提案した結果、60代の女性患者の命が救われたことが分かりました。**

人工知能は、このほかにも医師では診断が難しかった2人のがん患者の病名を突き止めるなど合わせて41人の患者の治療に役立つ情報を提供していて、専門家は「**人工知能が人の命を救った国内初のケースだと思う**」と話しています。



# Google DeepMind

DeepMindはロンドンに拠点を置くベンチャー企業で、2010年にDemis Hassabisらが創業した。DeepMindは**ビデオゲームを見るだけで、驚異的な速度でプレーの仕方を学習し、世界を驚かせた**。Googleが2014年1月に買収し、今では「Google DeepMind」として研究開発を進めている。同時に、GoogleはDeepMindをYouTubeなどの製品に組み込む作業に着手した。

## 汎用的に学習できるシステム

Google CEOのLarry Page(上の写真右側)は、トークショーホストCharlie Rose(同左側)との対談で、DeepMindの機能とGoogleの人工知能戦略について語った。Pageは、**DeepMindは驚異的な速さで、人の手助け無く、自ら学習するシステム**であると説明した。その結果、DeepMindは、**人間のエキスパートを凌駕するスコアをマークする**。



# DeepMind 医療への応用

1. 人工知能を使って過去の700件の症例を分析
2. 機械学習を使って放射線治療計画の作成を効率的に行えるようにサポート
3. 治療計画作成にかかる時間を  
**4時間から約1時間にまで短縮**
4. どの部分にどれくらい照射するか  
など治療計画の**決定権は**、人工  
知能ではなく**医師**にある



By liz west



Anonymised CT and MRI scans from 700 former University College London Hospital radiotherapy patients will be analysed by Google's artificial intelligence division, DeepMind.

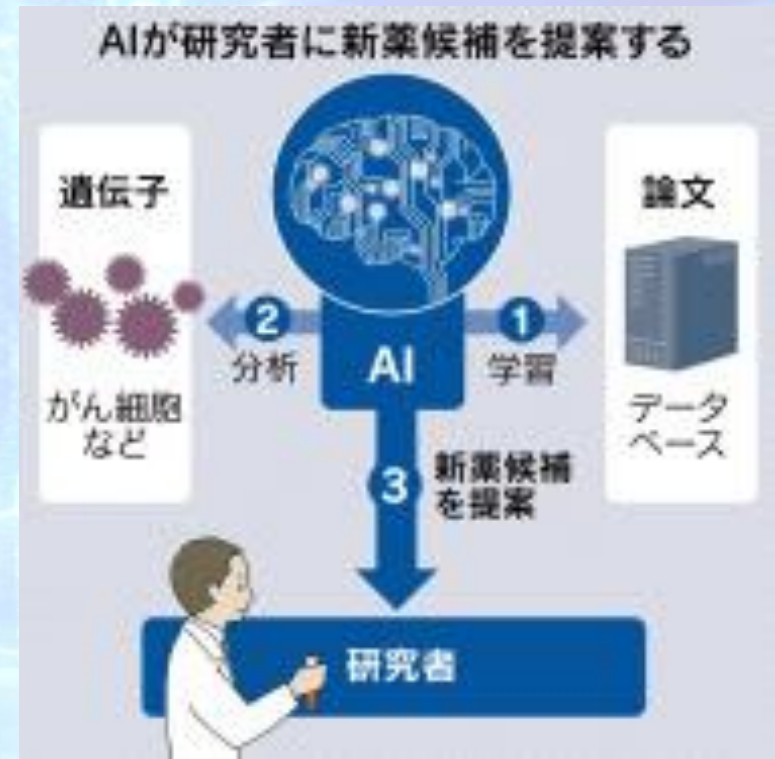
# 新薬候補 AIが提案 論文学習、新物質探る 抗がん剤などの開発費を抑制 厚労省後押し

2016/9/23付

日本経済新聞 朝刊より抜粋

**厚生労働省は人工知能(AI)を使い、高い効果の見込める画期的新薬の開発を後押しする。抗がん剤といった新薬のもとになるシーズ(種)と呼ぶ新規物質を見つけ、数年内に研究者らに提案することを目指す。**

**グローバルに新薬開発競争が激しさを増す中、巨額の費用が必要で成功率も低い新薬の開発に向けて国の支援を強化する。AI活用で開発を効率化し、医療費全体の抑制につなげる狙いもある。**





## AI(ArtificialIntelligence)創薬への動き～ディープラーニングとは何か？～『健康医療分野のビッグデータ活用・研究会』レポート

医薬産業政策研究所 統括研究員 森田 正実

医薬産業政策研究所 主任研究員 鈴木 雅……………17

## 欧州のBigDataforBetterOutcomes

医薬産業政策研究所 主任研究員 鈴木 雅……………23

政策研ニュース No.47 2016年3月

<http://www.jpma.or.jp/opir/news/>よりダウンロード可

# 参考資料 「人工知能」 WIKKI

「人工知能」という名前は1956年にダートマス会議でジョン・マッカーシーにより命名された。現在では、記号処理を用いた知能の記述を主体とする情報処理や研究でのアプローチという意味あいでも使われている。**日常語としての「人工知能」という呼び名は非常に曖昧なものになっており、多少気の利いた家庭用電気機械器具の制御システムやゲームソフトの思考ルーチンなどがこう呼ばれることもある。**

プログラミング言語 LISP による「ELIZA」というカウンセラーを模倣したプログラムがしばしば引き合いに出されるが(人工無脳、会話ボット)、計算機に人間の専門家の役割をさせようという**「エキスパートシステム」**と呼ばれる研究・情報処理システムの実現は、**人間が暗黙に持つ常識の記述が問題となり、実用への利用が困難視**されている現状がある。

人工的な知能の実現へのアプローチとしては、「ファジィ理論」や「ニューラルネットワーク」などのようなアプローチも知られているが、従来の人工知能[との差は記述の記号的明示性にあると言えよう。近年では「サポートベクターマシン」が注目を集めた。また、自らの経験を元に学習を行う強化学習という手法もある。

日本には人工知能学会があり、オンラインで機関誌も読める

Northwestern大学の Grzybowski教授らの手法

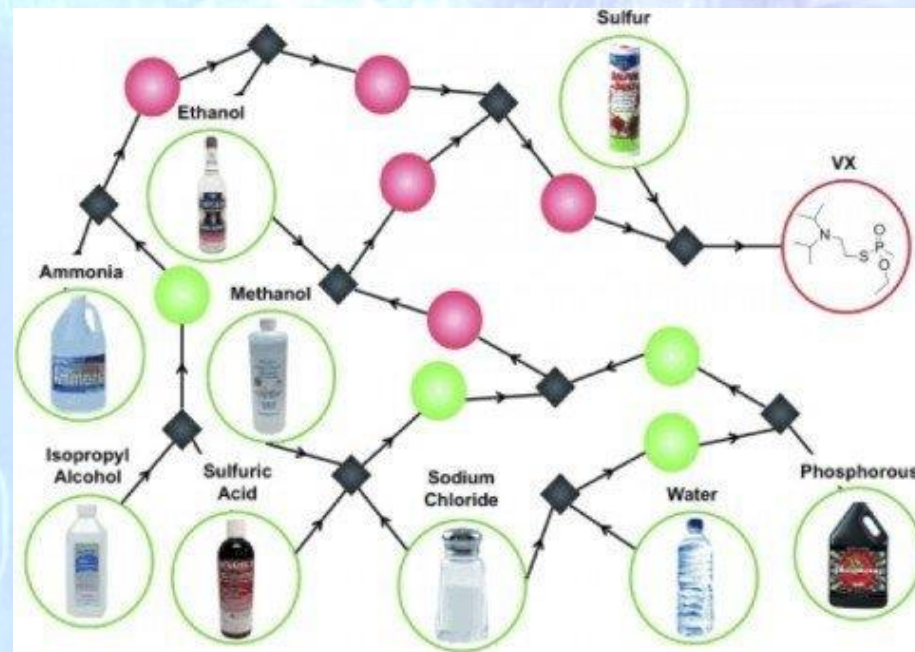
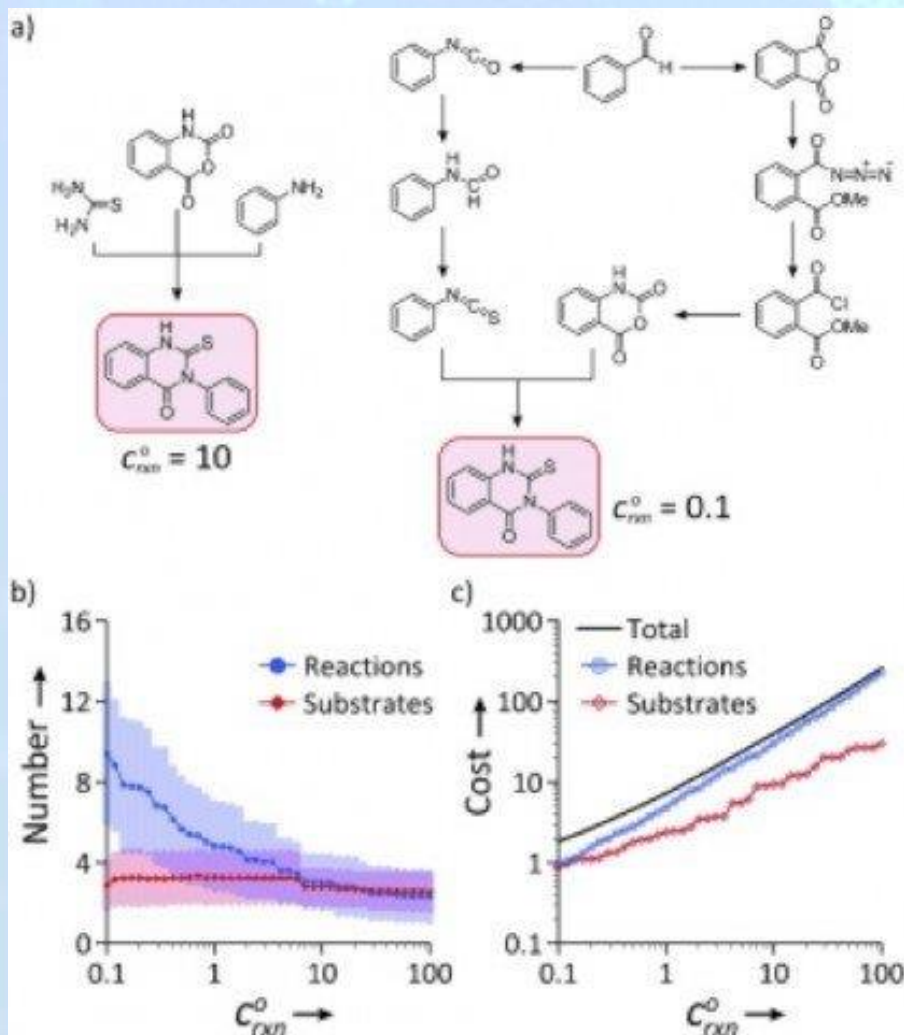
“Rewiring Chemistry: Algorithmic Discovery and Experimental Validation of One-Pot Reactions in the Network of Organic Chemistry”

Angew. Chem. Int. Ed. 51, 7922-7927 (2012). doi:10.1002/anie.201202155

- 現在までに報告されている数百万にもものぼる有機化合物と、反応に関するデータベースを構築
- それらの関連性すなわちネットワークを解析するという手法を採用



# The Network of Organic Chemistry

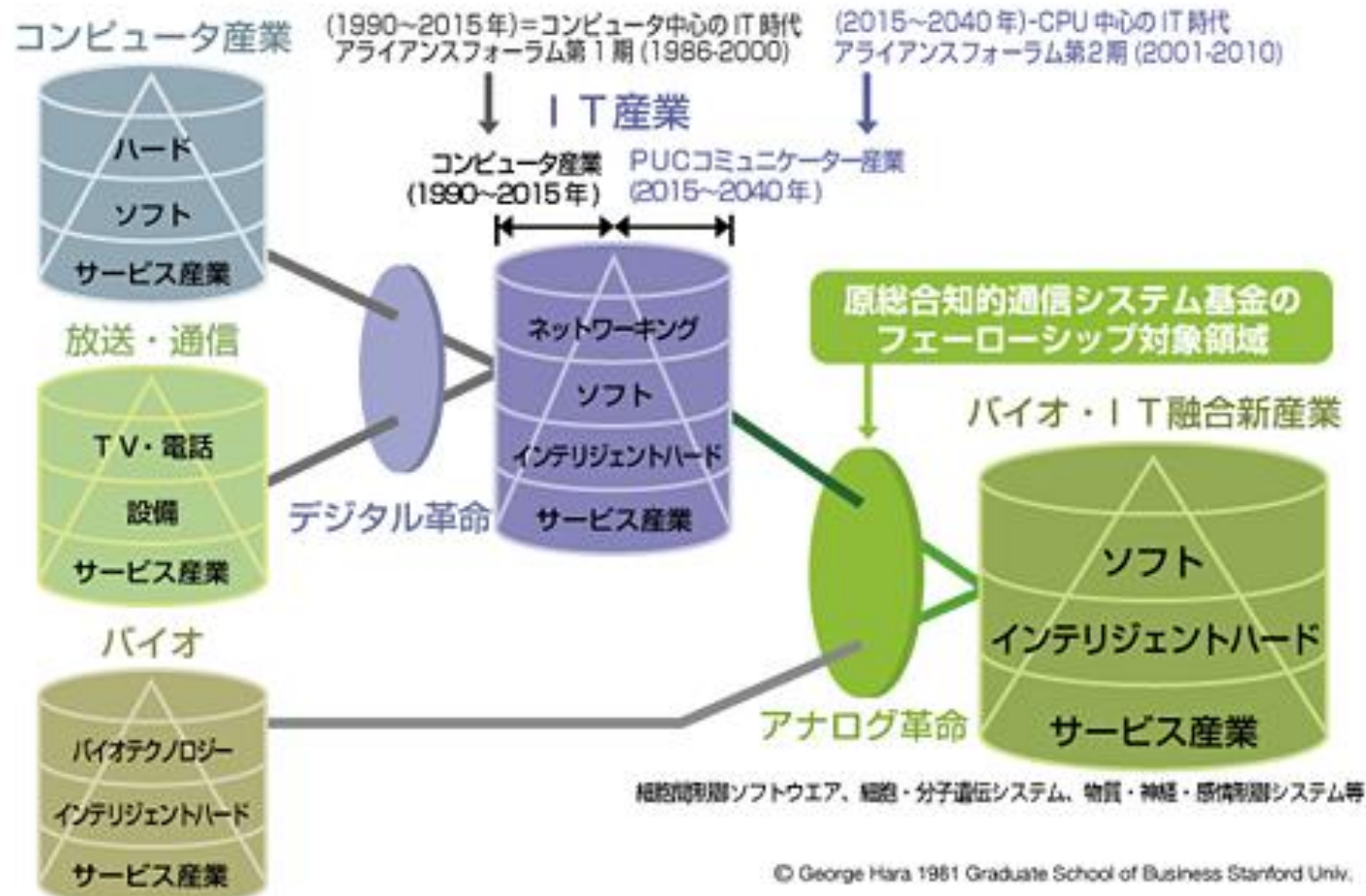


人口知能と、自動合成ロボットが組み合わさった時、有機化学者の仕事は？



# 参考資料: アナログ革命

## 1990年創立時に構想した 公益財団法人 原総合知的通信システム基金 ～先進技術の融合と産業化への波及効果～



# 本日の講義内容

1. 遠隔講義「計算生命科学の基礎」について
2. 計算生命科学とは
3. ノーベル賞と計算生命科学
4. 「生命科学研究への情報科学技術のインパクト」から
  - ① データサイエンス
  - ② 構造ベース創薬 (タンパク質構造データベース PDB)
  - ③ 人工知能 (AI)
5. 2016年度シラバス紹介
6. まとめ Q&A

# 2016 計算生命科学の基礎III

## 講義内容

### 第1編 「バイオインフォマティクス」

(担当:木下 賢吾、江口 至洋)

### 第2編 「構造生命科学のための分子シミュレーション」

(担当:田中成典)

### 第3編 「計算生命科学の最前線」

(担当:森一郎)

# 第1編 「バイオインフォマティクス」

## ① ゲノムに記された遺伝ビッグデータを読む

ーヒトゲノム計画から大規模個人ゲノム解読時代の到来までー

岩手医科大学 いわて東北メディカル・メガバンク機構: 清水 厚志

## ② ゲノム情報からの生命現象・病理現象の統計解析

京都大学大学院 医学研究科附属ゲノム医学センター 統計遺伝学: 山田 亮

## ③ ゲノム・タンパク質のバイオインフォマティクス入門

東北大学大学院 情報科学研究科 生命システム情報科学: 木下 賢吾

## ④ 人工知能研究と生命科学

ーディープラーニングのバイオテクノロジーへの応用可能性ー

東京工業大学 情報理工学院 情報工学系 : 石田 貴士



## 第2編 「構造生命科学のための分子シミュレーション」

### ① 計算生命科学のための量子化学基礎

東京大学 生産技術研究所:佐藤 文俊

### ② フラグメント分子軌道法の基礎と応用

京都大学 福井謙一記念研究センター :北浦 和夫

### ③ QM/MM法を用いたタンパク質機能解析

広島市立大学大学院 情報科学研究科 医用情報科学専攻:鷹野優

### ④ 分子シミュレーションを活用した創薬支援技術

産総研 創薬分子プロファイリング研究センター, 筑波大学: 広川 貴次

### ⑤ ドッキングソフトの原理と実際

産総研 創薬分子プロファイリング研究センター :福西 快文

## 第3編 「計算生命科学の最前線」

- ① 創薬における計算生命科学：インフォマティクスとシミュレーションを  
融合したインシリコスクリーニングと設計

理化学研究所 ライフサイエンス技術基盤研究センター：本間 光貴

- ② 製薬企業におけるデータサイエンス

塩野義製薬株式会社 解析センター：都地 昭夫、北西 由武

- ③ 計算生物学によるシステムの理解からの創薬への展開

医薬基盤・健康・栄養研究所 バイオインフォマティクス：水口 賢司

- ④ 全脳アーキテクチャ・アプローチ：脳全体のアーキテクチャに学び  
人間のような汎用人工知能の構築を目指す

ドワンゴ人工知能研究所：山川 宏

- ⑤ 計算生命科学がもたらすものへの期待

日経BP社 医療局特命編集委員：宮田 満

# 受講者の皆さんへ

この講義は実験生命科学 及び 広く理工系の方々を対象

**生命科学を志す皆さんには、**

意識するとしなにかかわらず、研究室にある実験・計測装置から生み出される大量のデータや、世界的に公開されている膨大な生命科学データベースを基盤に研究を進めることが求められてくると思います。

**理学、工学を志す皆さんには、**

その基礎的な理論を生命科学に生かしていく場がますます広がってきているとともに、生命科学の方々から連携を強く求められてくると思います。

**この講義により理学、工学、生物学などの枠を超えて、「理論、計算、観察・実験の協力関係がより一層深まって行くことを願っています。」**





皆さんのご聴講を  
心よりお待ちしております。