

ElixにおけるAI創薬 構造生成を中心に最新事例も交えて

CEO 結城伸哉



創薬の課題 / AI創薬導入の課題

創薬の課題

膨大な時間と
コストと
低い成功率

典型的には10-15年もの期間と20億ドル以上にも及ぶ費用がかかる。また多くの創薬プロジェクトは、失敗に終わる。

AI創薬の課題

AI人材や
経験の不足

多くの製薬企業はAI人材やAI創薬プロジェクトの経験が不足。基本的にAI企業と組んで始めることになるが、どのAI企業と組むのがよいのかも分からない。

AI創薬の課題

最適なソリューション
が見つからない

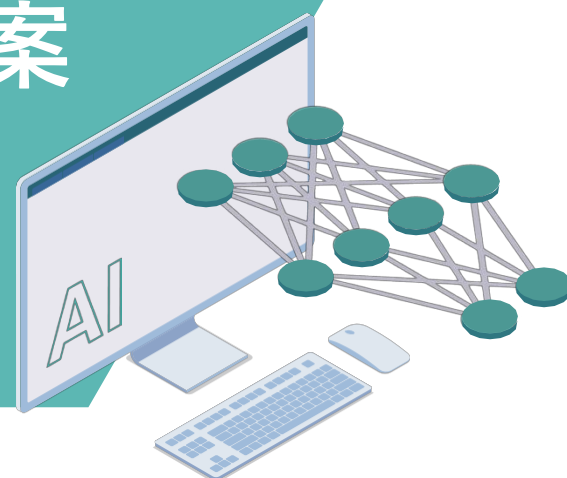
既存のAI創薬ソフトウェアは実際にメドケムが使いこなせるようにはなっていない。また、特に有名なAI企業との共同研究は高額になり過ぎて中堅企業には特に難しい状況。

Elix Discovery™と特徴

ELIX Discovery™

メドケムが本当に使える、AI創薬プラットフォーム

人間が思いつかない
構造を提案



構造最適化の加速による生産性向上だけでなく、メドケムが思いつかない新規構造の提案による課題解決も重視。**成功事例も複数。トライアルも実施中。**

コンサルティング
サポート



単純にAI製品だけを導入しても、なかなかケミストが使いこなすに至らないケースが多い。Elixでは製品操作方法のレクチャーはもちろんのこと、機械学習の基礎からAI創薬のコツまで**どこよりも手厚く支援。**

手法やノウハウを
オープンに開示



Elixは自社の手法を再現可能なレベルでオープンに共有している世界的に見てもユニークなスタンス。製薬会社様がAI創薬の**ノウハウを獲得し、より深い議論が可能**になることで共同研究の成功確率が向上するというメリットも。

取引実績等

製薬企業



合計15社以上（世界トップ30に入る製薬企業2社を含む）との取引実績

パートナー

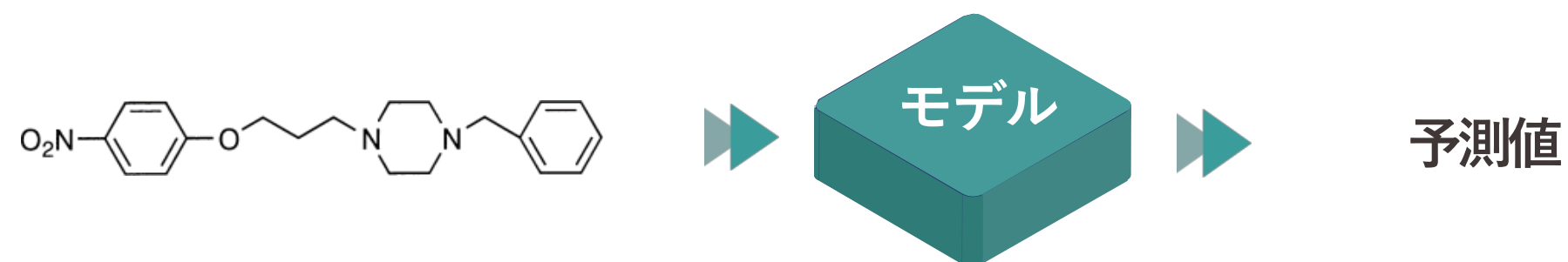


京都大学



Elix Discovery™ メイン機能

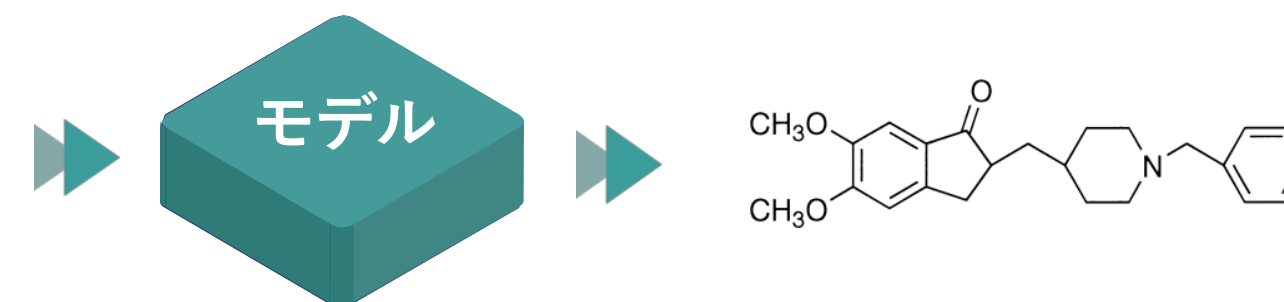
予測モデル



- 活性・物性・ADMET等を予測
- 古典的な機械学習モデルから最新の深層学習モデルまで。分類と回帰の両方に対応
- 自動で最適な予測モデルを作成

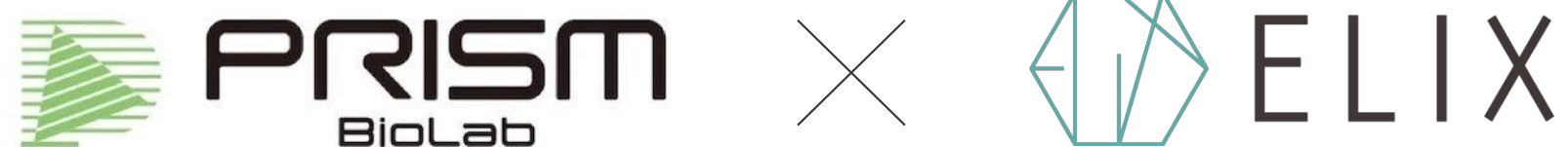
構造生成モデル

所望のプロファイル
/ 起点構造



- 所望のプロファイルを持つ化合物を生成
- 活性・物性・ADMETに加え、合成容易性等を考慮して構造を最適化
- De novo生成、リード最適化、リンカー生成など様々なタスクに対応

PRISM BioLab様との共同研究:概要



共同研究の目的

- PRISM様が持つ化合物構造を起点に、活性を維持しつつ新たな構造を見出す。もしくは活性を向上させる。

役割分担

- PRISM様：データ提供及び合成とアッセイ
- Elix：AIを活用した化合物デザイン

データ

- 利用したのは活性データのみ
- タンパクの構造情報は利用できない。かなりチャレンジングな問題設定

PRISM BioLab様との共同研究: 化合物デザインに活用した技術と流れ

構造生成モデル

- SmilesFormer: Elixの自社開発モデル
- Lib-INVENT: アストラゼネカが開発したオープンソースのモデル
- 自社開発、オープンソース問わず、良さそうなものを使うスタンス

予測モデル

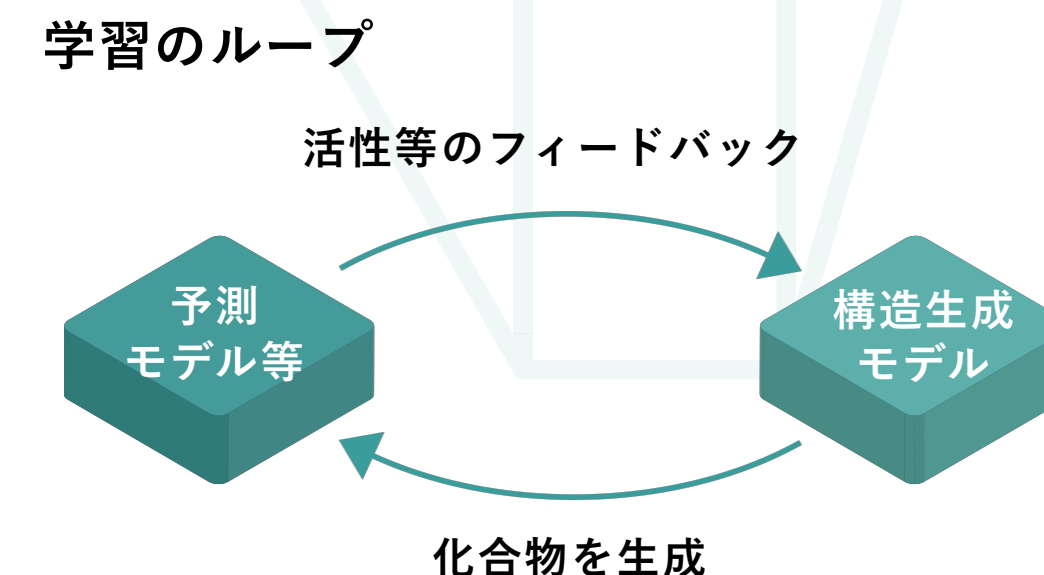
- 古典的なものからグラフベースなど複数作成したが、一番良かったのはMLP。さらにそのアンサンブルモデルも作成。精度も良好
- 関連の低い方のデータで事前学習し、関連の高いデータでfine-tune

最適化項目 (強化学習のreward)

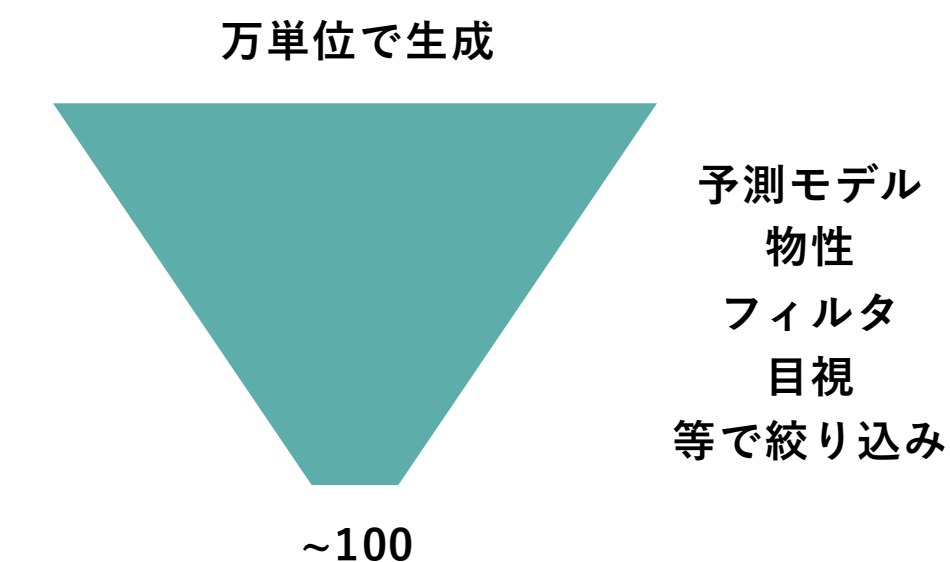
- 活性だけでなく、物性や合成容易性など複数考慮
- モデル自体にはばかり注目が行きがちだが、この設定も結果に大きく影響

生成化合物の絞り込み

- 大量に生成してから絞り込み。Rewardと同様にノウハウが影響する部分



生成化合物の絞り込み



PRISM BioLab様との共同研究：共同研究の進め方と合成

共同研究の進め方

- 化合物リストを共有してフィードバックをもらうというループを数ヶ月のうちに数回繰り返した
- 一般的にElixとの共同研究では、いきなり合成に進むのではなく、まずは共同研究先のメドケムからフィードバックをもらって構造生成モデルの最適化項目を調整するという作業を行う。ここはあまりイメージを持っていない人が多いと思われる。実際に今回のプロジェクトでもかなり影響した。
- AIというと、一般的には合成して増えた化合物を学習データにしてモデルが良くなっていくのがメインと考えがち。しかし、実際にデータは急に増えるわけではない。**実際に多く行われるのは最適化項目の調整の方。**

合成

- 結局合成は**1サイクル**しか行っていない
- 合成した化合物は**たったの6個**

PRISM BioLab様との共同研究: 結果と考察

6個中2個がヒット

- 2つのヒットのうち、1つは別のアッセイ方法でも活性を確認
- 学習データとしては片方のアッセイ方法しか利用していない

自社開発モデルの優位性

- 合成の候補となった6個中5個がElixの自社開発モデルから生成
- 一番良いヒット（2種類のアッセイで確認）も自社開発モデルから

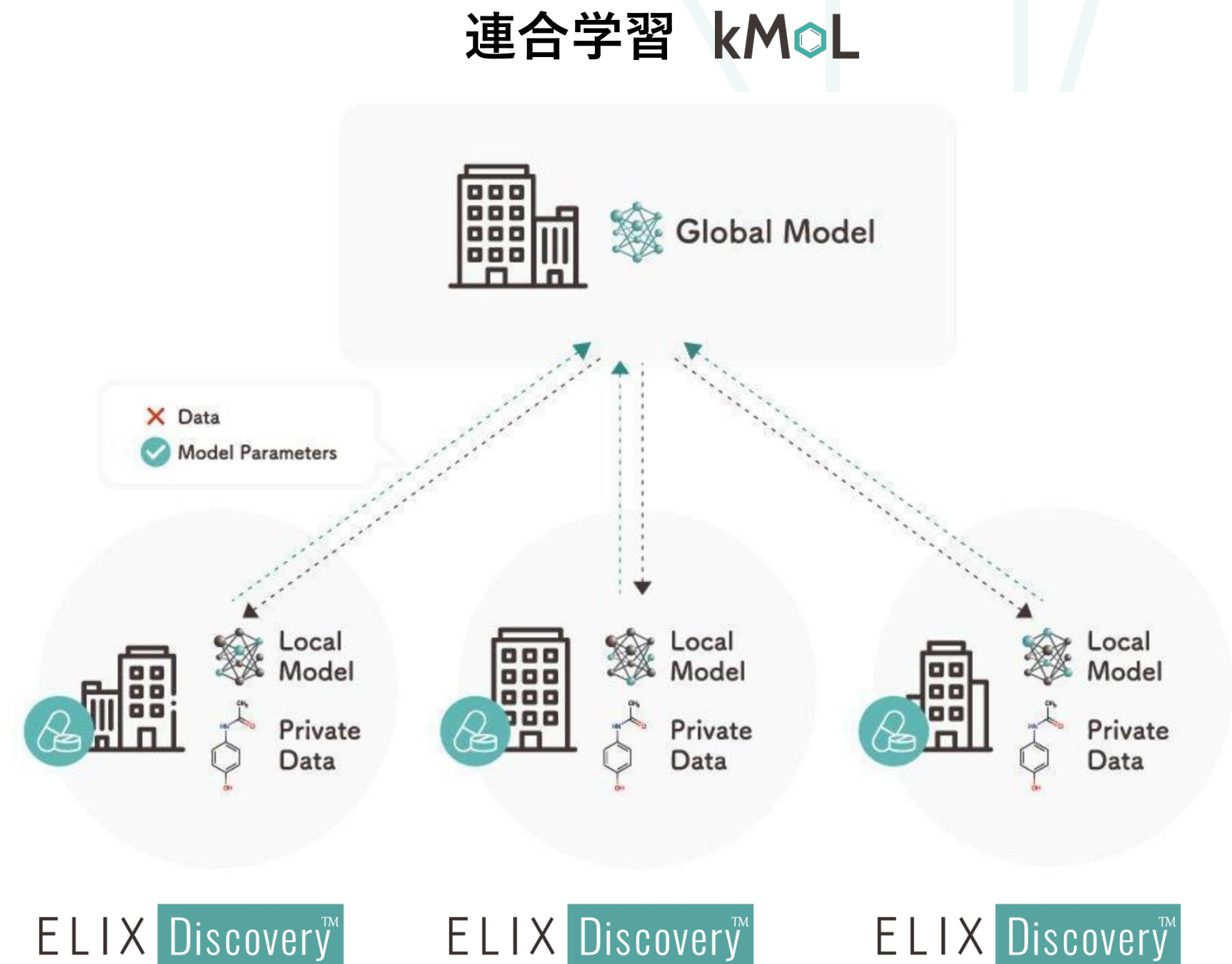
新しいケミカルクラスのヒットを創出

- AI創薬は期間短縮に焦点があたりがちだが、**人間だけではたどり着くことができなかつたであろう構造の化合物を創出にも役立つ**。Elixでは後者の新規性の方をより重視
- 今回のケースでも人間だけでは合成に進めるという判断にならなかつたであろう化合物を合成し、ヒットの創出に繋がった
- 弊社の他の事例なども踏まえても、そもそも人間には思いつかない構造を生成したり、人間には活性があるようには思えない構造でも活性が出たりという例が複数出てきている

ID	活性	備考
化合物1	ヒット	2種類のアッセイ方法で活性を確認
化合物2	ヒット	
化合物3	弱いヒット	
化合物4	弱いヒット	
化合物5	活性なし	
化合物6	活性なし	

連合学習 (Federated Learning)

- 最高の予測モデルを作成するためには、あらゆる製薬企業からデータを集めて学習させれば良い。
→ 現実的にはデータを共有することは不可能
- 連合学習：**データの秘匿性を保ったまま**複数のデータソースから学習可能。データではなく、モデルのパラメータをやり取りする
- **kMoL™**：Elixが京都大学と共同開発。世界でも唯一のAI創薬特化型の連合学習用ライブラリ。予測モデルも併せてオープンソース化。論文も近々投稿予定
- DAIIA：アカデミアが中心となり、**製薬企業17社が集まって連合学習**を行う**世界でも唯一**のアクティブなプロジェクト。ここでもkMoL™が使われている。
- 事業化：ここで学習されたモデルが**Elix Discovery™**上でも利用可能になる予定



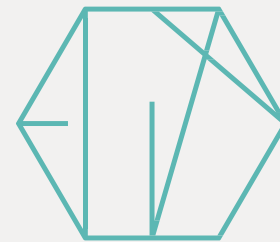
グローバル人材

- 約10カ国から集まり、全員が日本で勤務
- 研究者のうち78%がPhD取得者
- 社内公用語は英語
(社外のやり取りは日本語・英語両方可)



学際的なチーム

- AI研究者、エンジニア：AIはElixのコア技術。既存技術を活用するだけでなく、オリジナルのAIモデルを開発する能力も持つ
- 創薬化学、生物、物理など多様なバックグラウンドを持つメンバーの割合が増えてきている。創薬はやはりAIだけではダメで、異なる領域の専門家が協調することが重要



ELIX

Contact us: bizdev@elix-inc.com