

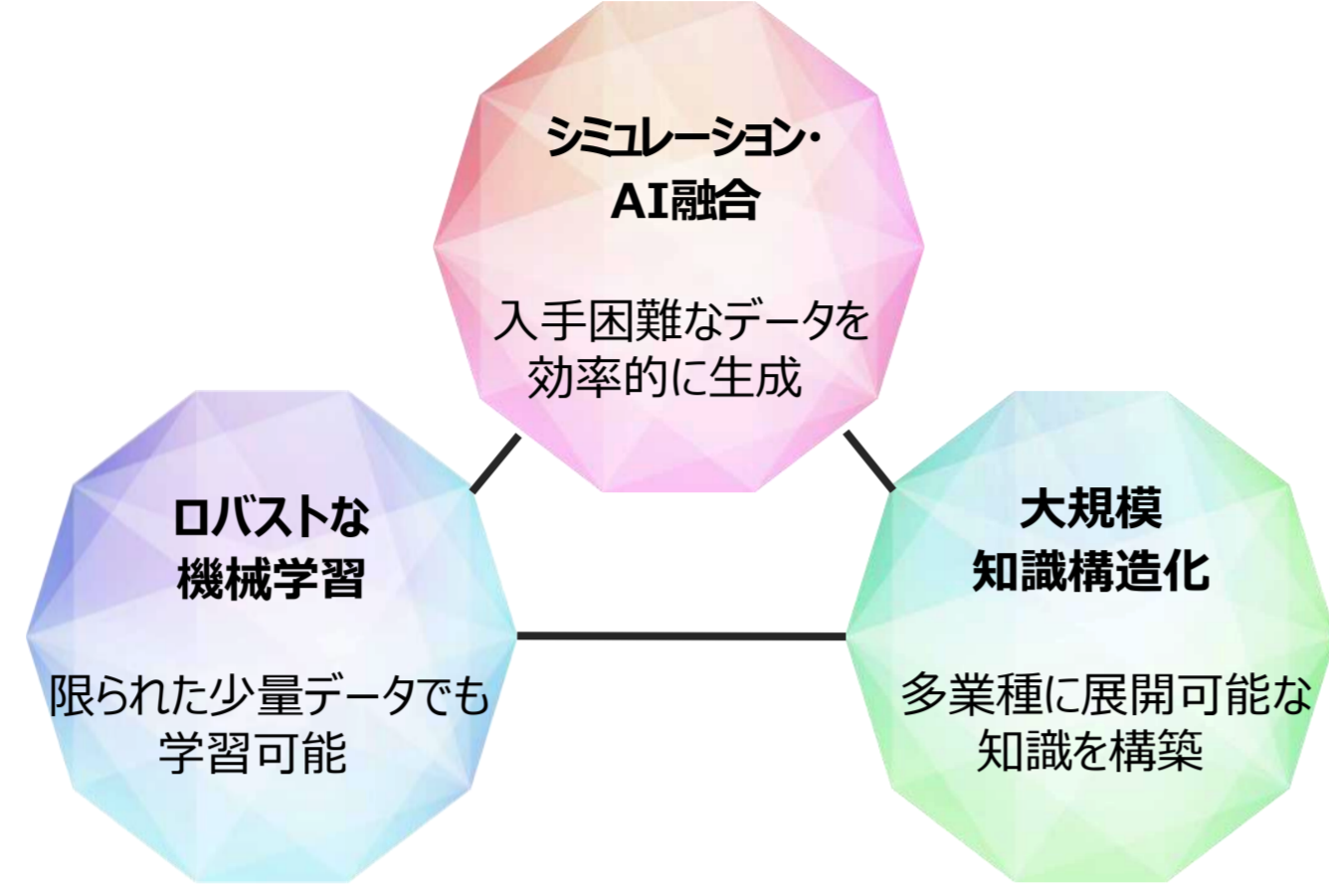
## 研究方針

人が幸せになるAI、自律成長と人との協調による課題解決

**共同研究テーマ**

**想定外を想定するAI**

環境の不確実な変化に対しても、的確な未来予測に基づき、より良い判断を支援



### ロバストな機械学習

- ディープラーニングを使って超音波画像から先天性心疾患を自動検知する技術を開発
- 因果探索技術にニューラルネットワークを適用し、非線形な因果関係を探索可能に拡張
- データの幾何学的構造を活用して僅かなデータで学習する技術を開発

### シミュレーション・AI融合

- データ科学を活用した材料開発プロセスの高効率化手法を開発、実証

### 大規模知識構造化

- 分散されているテキストから化学知識を構造化するための情報抽出技術を開発

## 因果推論

### 非線形因果探索

- より現実に即した仮定の因果探索手法の研究開発
- ものづくり、創薬などにおける最終性能への設計パラメータの影響の推定
  - 経済学、社会学における複雑なシステムダイナミクスの解明

### 提案手法 ICASSP2020採択

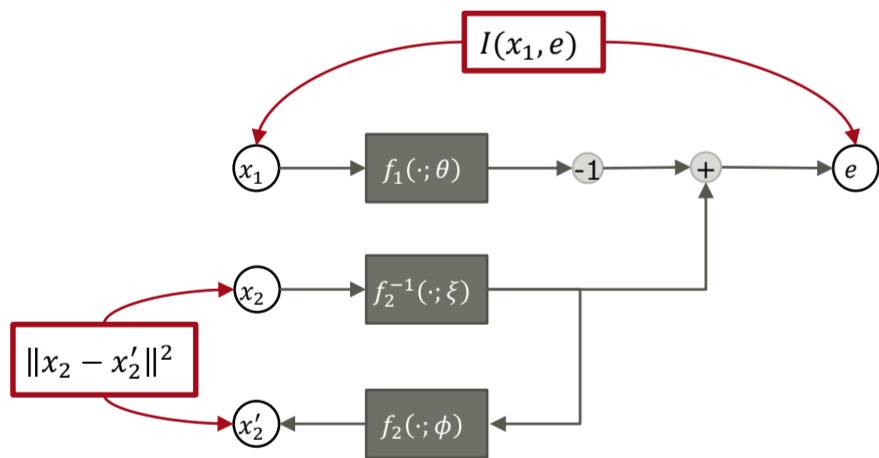
Post Nonlinearモデルを仮定し、NNでモデル推定

$$x_2 = f_2(f_1(x_1) + e)$$

$f_1, f_2$ は非線形関数、 $e$ はノイズ (= 未知の原因因子)

[条件1]  $f_2$ は可逆  
[条件2]  $e$ と $x_1$ は独立

- $f_1, f_2, f_2^{-1}$ をそれぞれNNでモデル化
- 誤差を以下で定式化してNN学習
  - [条件1] 再構成誤差  $\|x_2 - x_2'\|^2$ の最小化
  - [条件2] 変数の従属の度合い $V(x_1, e)$ の最小化

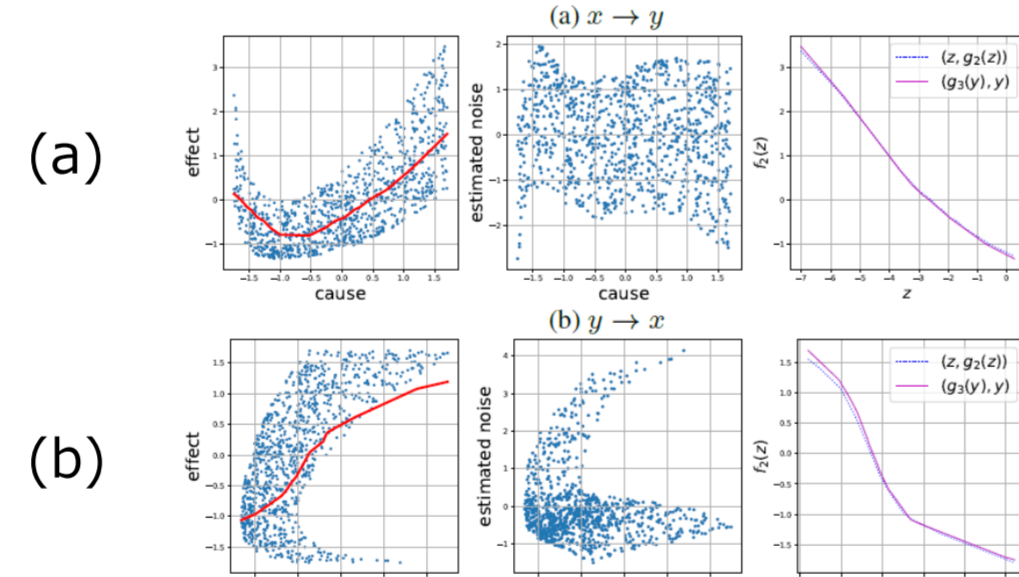


### 従来手法 (LINGAM)

- 変数間の因果関係が線形であることを仮定するが、現実問題の多くは非線形
- 仮定と現実のギャップにより精度やロバスト性、ユーザの技術受容性が低下

### モデル仮定を満たす人工データによる検証

- (a) 正しい方向( $x \rightarrow y$ )を仮定して推定した結果  
中央図: 推定した独立成分( $x, e$ )のプロット  
右図:  $f_2$ の可逆性検証プロット
- (b) 誤った方向( $y \rightarrow x$ )を仮定して推定した結果  
独立成分を抽出できないことを確認
- 誤差値より真の方向を推定可能



- 実データ(UCI ML Repo.)ベンチマーク
  - 人が付けた因果方向をデータから推定
  - LINGAMに比べ、多くの推定に成功 (31 vs 20 /41ケース)

## マテリアルズ・インフォマティクス

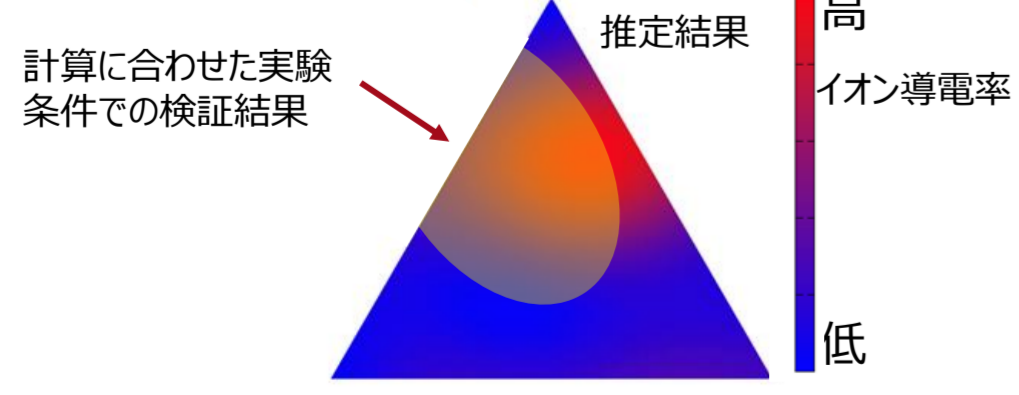
### 材料開発プロセス高効率化

- 所望の特性をもつ材料の組成だけでなく、その合成プロセスまで最適化
- 企業の利用を想定した高効率で高速なシミュレーション手法の開発: 精度とコストのバランスを重視

### 材料組成最適化 BCSJ採択

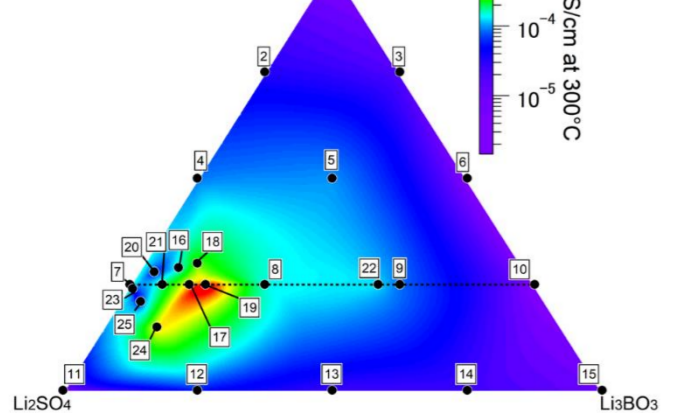
- 特定の実験条件下での高イオン導電率材料の最適組成推定 (18年度)

「第一原理計算 x ベイズ最適化」で定性的には予測可能



ジャーナル投稿

- 合成プロセスまで含めた最適化 (19年度)
  - 組成ごとに最適な合成条件を実験家が選択し、ベイズ最適化で組成探索
  - 合成プロセスまで含めた最適化を実証中



### 対象: Li2次電池用固体電解質

- 全固体Liイオン電池のキー材料、電池の小型化・大容量化、長寿命化、急速充電を可能に
- 世界中で高イオン導電率化に向け開発競争
- 富士通研究所には実験データの蓄積あり

### シミュレーション高速化

- 高イオン導電率材料探索に特化し、シミュレーションを高速化 (古典分子動力学法の採用)
  - 必要な情報(原子レベル)を小規模な第一原理計算から抽出
  - 単純なモデルでパラメータ化し、重要ではない機能(化学反応など)にかかわる部分の大幅カットを段階的に実行
- 第一原理計算で1年がかりの計算が30分程度で済むことを予備的計算で確認済
- ベイズ最適化、遺伝的アルゴリズムなどを利用したパラメータ作成作業を進行中

## 周産期医療

### 胎児心臓超音波スクリーニング

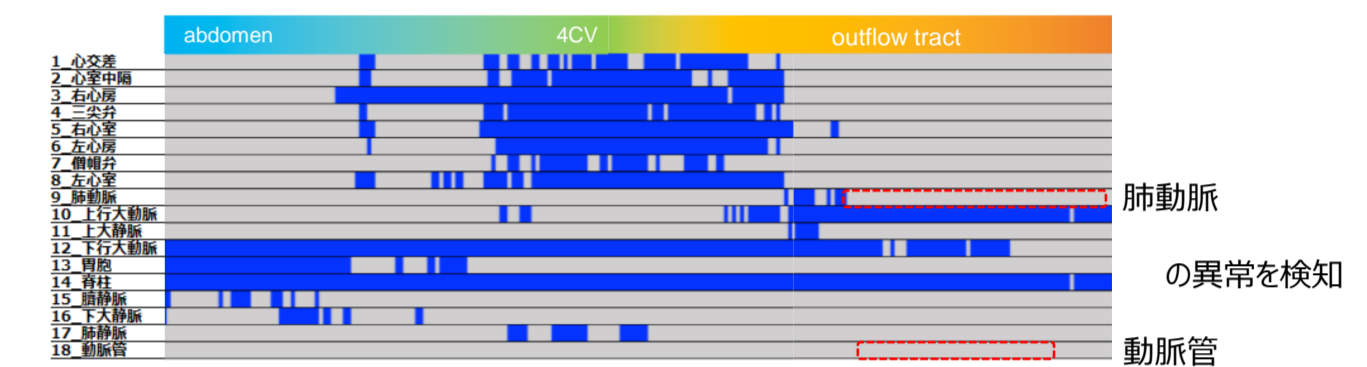
- 胎児超音波診断動画からの先天性心疾患検知
- 先天性心疾患は新生児死亡原因の20%
  - 先天性心疾患は出生前診断により治療が可能
  - 妊婦健診での超音波診断による早期発見が重要

### 技術的課題

- 異常例が極めて少ない(全出生児の1%)
- 母体・胎児の角度、高速な心臓の動き、超音波映像の特性(骨の影)、検査者の熟練度等により、正常画像のバリエーションが極めて多い
  - 従来の異常検知技術では困難
- 利用可能なデータの収集は始まったばかり

### 臨床的価値の検証 ISUOG2019受賞

- フォロー四徴症(TOF)での評価において、肺動脈および動脈管について正常胎児との明確な違いを捉えた
  - 疾患に特徴的な構造所見と一致

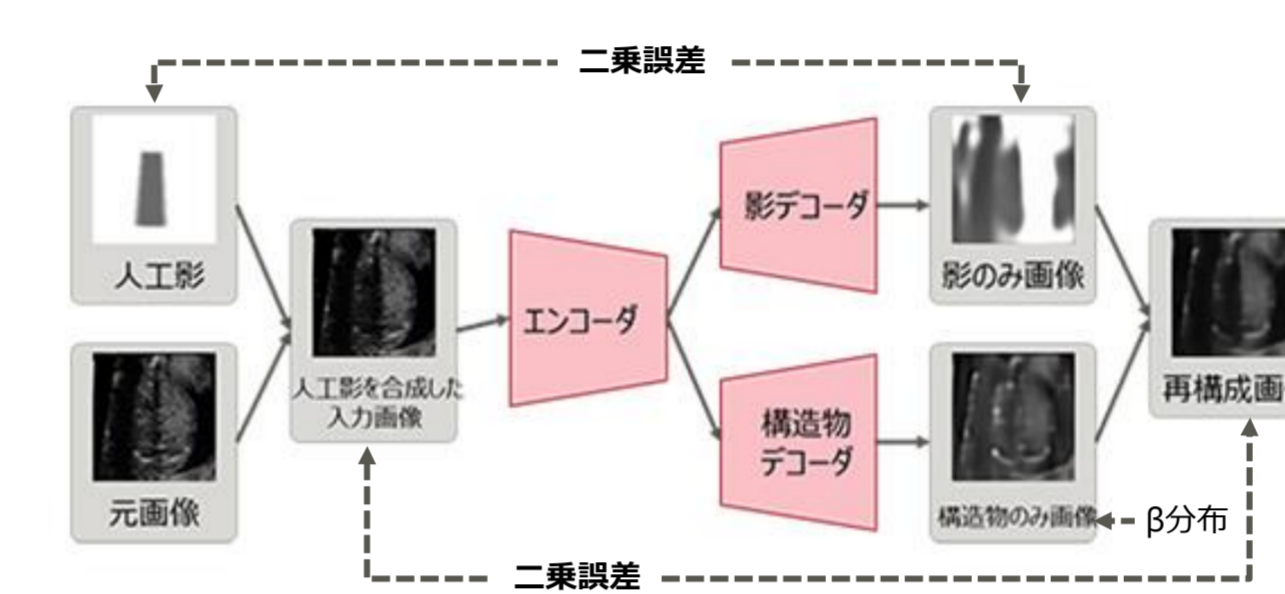


- 産婦人科超音波医学の最高峰国際会議ISUOG 2019で日本人唯一のAward受賞



### 影の自動検出 MIDL2019採択 プレス発表

- ラベルなしデータによる影の検出
- 影無し画像に付加した人工影を分離するように学習
- 従来の教師あり学習(セグメンテーション)と比較して、はるかに低コスト



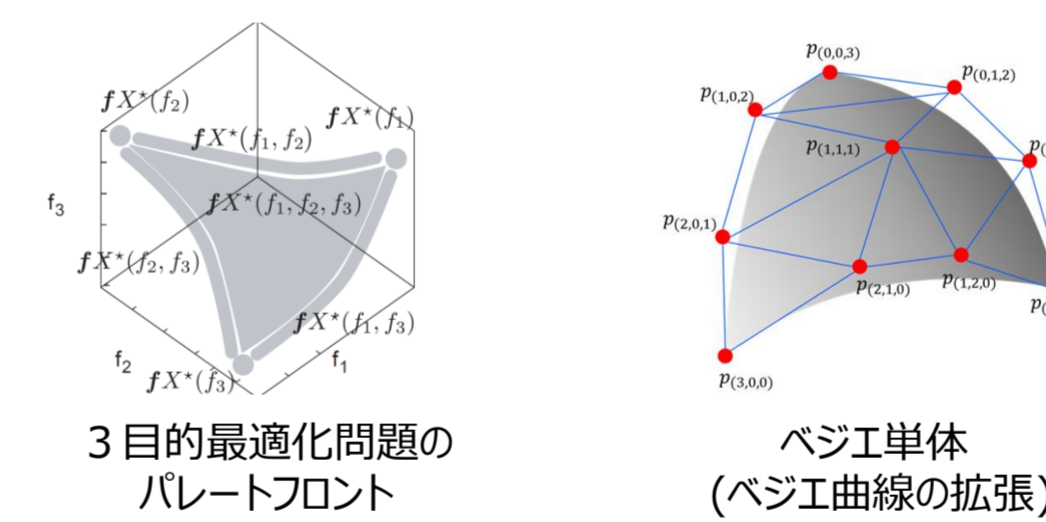
## Stratification Learning

### 滑層分割学習

- データの幾何学的構造(滑層)を活用し学習を効率化
- 工業製品設計: 様々な制約条件により、最適設計案の空間は滑層分割されている
  - SNS記事分析: 記事に付いたタグによってデータ空間が滑層分割されている

### 提案手法

- パレートフロントのベジエ単体フィッティング法 (18年度)
- 多目的最適化問題のパレートフロントをベジエ単体で近似する2つの手法を提案
    - 一斉推定法: 解集合全体から一様にサンプリングして近似
    - 骨格推定法: 各次元の部分単体から層別にサンプリングして低次元から近似



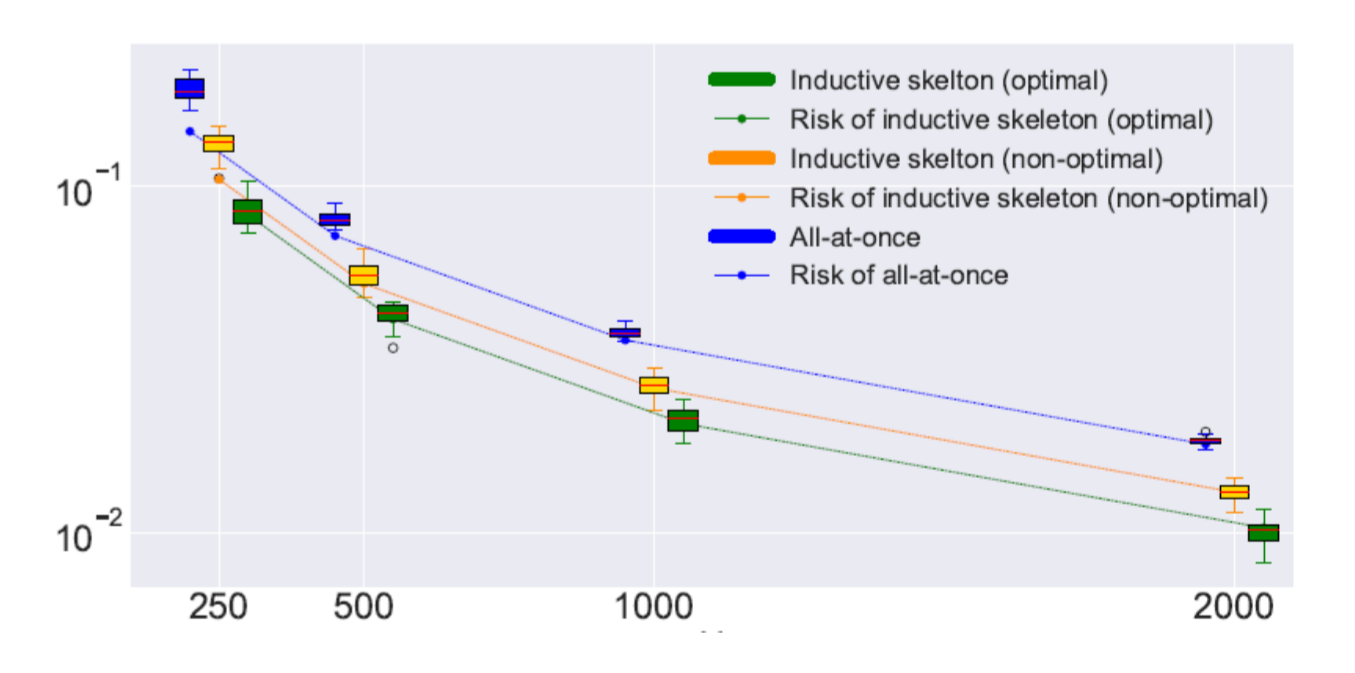
### 期待効果

- 環境変化による滑層分割構造の変化を把握できる
- 高次元データを低次元に分割して効率よく学習できる
- より少ない訓練データ(投資)で再学習できる

### AAAI 2年連続採択

ベジエ単体フィッティングの漸近リスク (19年度)

- 一斉推定と骨格推定の漸近リスクを理論的に評価
  - 一斉推定法/骨格推定法の漸近リスクを理論的に評価
  - 数値実験(人工データ、実データ)で評価結果の妥当性を検証

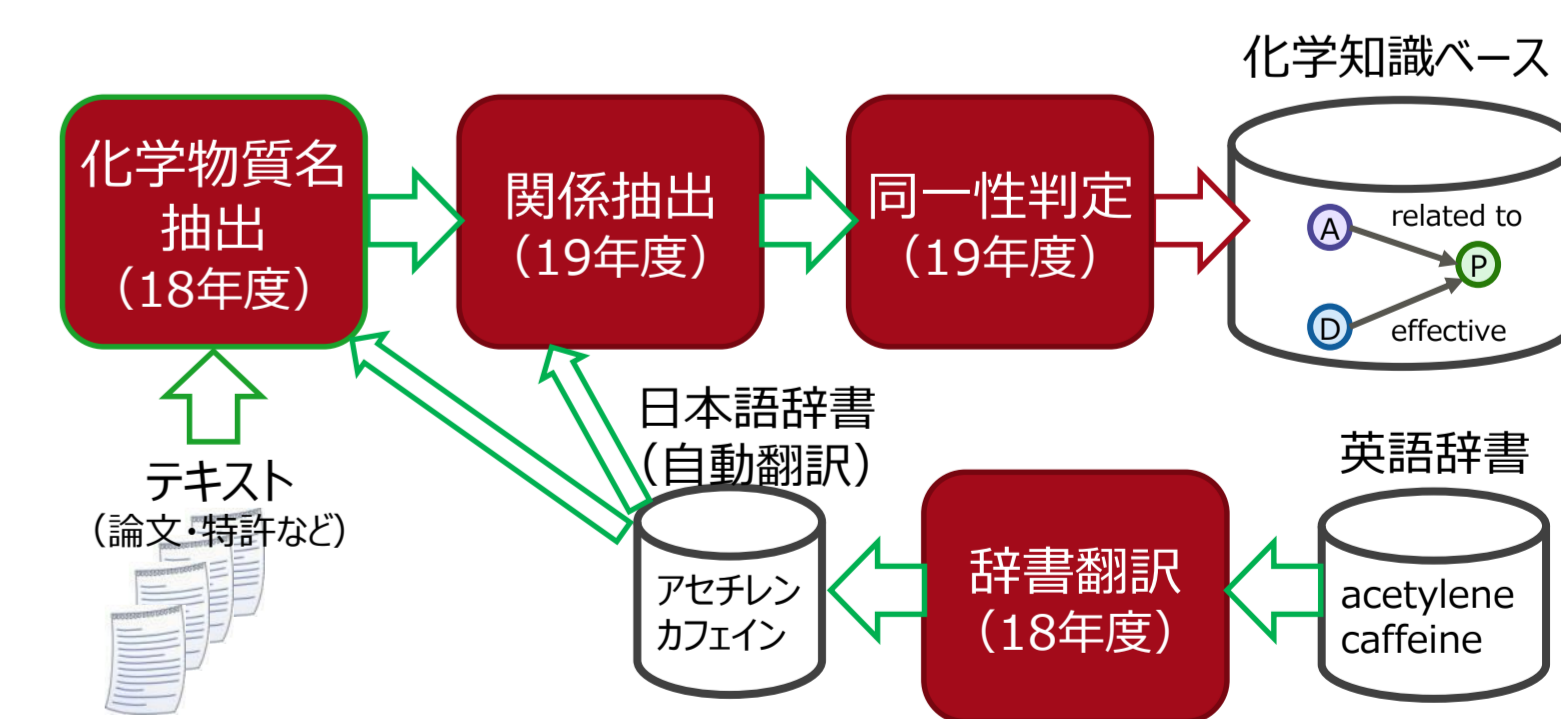


## マテリアルズ・インフォマティクス

## 大規模知識構造化

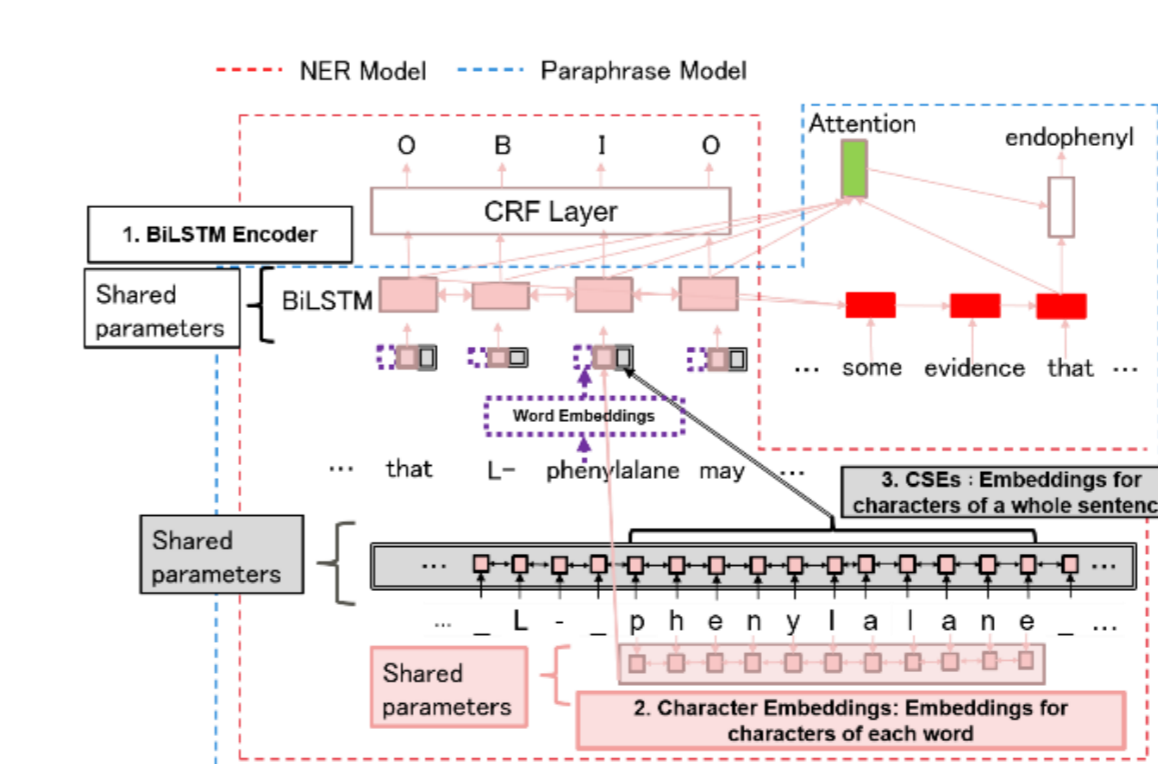
### 化学知識構造化

- テキストから化学知識を抽出する要素技術を開発
- 化学物質名抽出、辞書翻訳 (18年度)
  - 関係抽出、同一性判定 (19年度)



### 化学物質名抽出 EMNLP2019採択

- 固有表現抽出と言い換えのmulti-task学習手法
- CHEMDNERデータで最高精度達成



### 関係抽出

- 自動作成した補助タスクを用いたmulti-task学習手法
- CHEMPROTでBERTベースを超える最高精度達成
- 言語処理学会年次大会(20年3月)で発表予定

### 同一性判定

- Transformerに基づく原子数制約を考慮した構造推定学習で既存技術を大幅に超える精度改善を達成
- 言語処理学会年次大会(20年3月)で発表予定

