

# 材料およびプロセス開発のための インフォマティクスの基礎と研究開発最前線

Chemoinformatics, Materials-informatics and Process-informatics; Fundamentals and Frontiers

監修：船津 公人（東京大学大学院 工学系研究科 教授）

- 材料開発・プロセス開発にいまや必須のインフォマティクス！
- 各分野共通の基礎となる研究から、分野別最新動向を1冊で！
- インフォマティクスの理解を深める事例となる研究を豊富に掲載！
- 参考資料として「データ解析の進め方」を掲載、利便性を高めた！

### ＜発行要項＞

- 発行：2020年8月1日
- 定価：本体（白黒） 90,000円＋税  
本体＋CD（カラー）100,000円＋税
- 体裁：A4判・並製・本文270頁
- ISBN：978-4-904482-84-1

### 刊行にあたって

手順を通して目標とする分子構造や材料の設計を行い、最後にその妥当性の検証のために実験により確認することになる化学は帰納的な学問としてとらえられている。こうした中で特に明確な数式モデルを構築して、求める物性を満たす分子構造や材料を設計する流れが強くなってきた。実験、測定、そして計算を行うと必ず数字（データ）が得られる。このデータが、例えば有機低分子化合物の沸点だとすると、化合物と沸点とを関連付けたものが、データの上位の情報と呼ばれるものである。この情報の件数が100件、1000件と増えていくと、構造と沸点との関係性を明らかにしたいという願望が出てくる。この構造と沸点（つまり物性）との間の関係を数式として表現したものが構造物性相関モデル（ $y=f(X)$ ）と呼ばれ、情報の上位に位置する知識となる。この知識を利用することで沸点（物性）未知の構造情報をモデルのXに入力することでyとして沸点（物性）を予測できるようになる。また、逆に特定の物性をyに指定することで、その沸点（物性）を満たすX、つまり構造を求めることにつながる。予測と設計が可能となるのである。このようなことを目的の一つとしてケモインフォマティクスが発展してきたし、その土壌としてのデータの取納法、化学構造の表現法など、化学情報の取り扱いに関する様々な研究がケモインフォマティクスとして幅広く研究されてきた。

所望の機能を想定した場合、それを満たす分子、材料、デバイスを手に入れようとするが、そのために1) 何をやるか、2) それをどう作るか、3) それをはきたか、4) それを製品としてどう作るか、そして5) どう物質循環させるか、という項目が必然的に関わってくる。ケモインフォマティクスは化学に関わる広範な領域を対象とする知識情報処理と言える。それぞれの項目は、様々なデータ・情報から誘導される知識によってサポートされる。

マテリアルズインフォマティクスはその中で「何を作るか」に特にフォーカスしたケモインフォマティクスの一領域とみなせる。しかしながら、材料特性が作り方（プロセス）に依存して変わることから、「どう作るか」、「製品としてどう作るか」という項目が必然的に関わってくる。私はこの意味で、「プロセスインフォマティクス」を提言してきたが、もはやこれは当たり前のごとく定着しつつある。様々なフェーズで観測される化学データ・情報は相互に関係しあって高度利用可能な知識へと昇華していく。データや情報を通して材料設計の奥深さを感じる。材料設計の奥深さに合わせてマテリアルズインフォマティクスのスペクトルは大きく広がった。マテリアルズインフォマティクス領域の研究者がこの数年の間に多く活躍するようになったが、その主だった方々が今回のこの書の発刊の企画に賛同し、マテリアルズインフォマティクス最前線の研究動向を御執筆下さった。監修者として改めてここに深く感謝申し上げる次第である。この書がマテリアルズインフォマティクスを推進する多くの研究者にとって羅針盤となることを切に願うものである。

船津 公人

### 執筆者一覧（執筆順）

船津 公人	東京大学大学院 工学系研究科 化学システム工学専攻 教授	塩見 淳一郎	東京大学 大学院工学系研究科 機械工学専攻 教授
山野 仁詩	奈良先端科学技術大学院大学 先端科学技術研究科 研究システムズ生物学研究室、東京応化工業株式会社 AI 推進部	岩崎 悠真	日本電気株式会社 NEC システムプラットフォーム研究所 主任
清水 宏明	東京応化工業株式会社 AI 推進部	高際 良樹	国立研究開発法人 物質・材料研究機構 エネルギー・環境材料研究拠点 独立研究者
小野 直亮	奈良先端科学技術大学院大学 データサイエンス創造センター/先端科学技術研究科 研究システムズ生物学研究室 准教授	宮里 一旗	北海道大学大学院 理学研究院化学部門 特任助教
Altaf-Ul-Amin	奈良先端科学技術大学院大学 先端科学技術研究科 研究システムズ生物学研究室 准教授	高橋 啓介	北海道大学大学院 理学研究院化学部門 准教授
黄 銘	奈良先端科学技術大学院大学 先端科学技術研究科 研究システムズ生物学研究室 助教	高橋 崇宏	静岡大学 工学部 准教授
森田(平井) 晶	奈良先端科学技術大学院大学 先端科学技術研究科 研究システムズ生物学研究室	松本 凌	国立研究開発法人 物質・材料研究機構 若手国際研究センター ICYS 研究員
金谷 重彦	奈良先端科学技術大学院大学 先端科学技術研究科 研究システムズ生物学研究室/データサイエンス創造センター 教授、公益社団法人新化学技術推進協会	高野 義彦	国立研究開発法人 物質・材料研究機構 国際ナノアーキテクトニクス研究拠点 (MANA) ナノフロンティア超伝導材料グループ グループリーダー
知京 豊裕	国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門デバイス材料設計グループ	藤元 伸悦	日鉄ケミカル&マテリアル株式会社 総合研究所 基盤技術センター 主任研究員
木野 日織	国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門デバイス材料設計グループ 主任研究員	南 拓也	昭和電工株式会社 融合製品開発研究所 計算科学・情報センター リサーチャー
小山 幸典	国立研究開発法人 物質・材料研究機構 統合型材料開発・情報基盤部門デバイス材料設計グループ 主幹研究員	日沼 洋陽	東京工業大学 科学技術創生研究院 研究員
五十嵐 康彦	筑波大学 システム情報系コンピュータサイエンス専攻 准教授	蒲池 高志	福岡工業大学 工学部生命環境科学科 教授
田中 るみ子	筑波大学 図書館情報メディア研究科 博士後期課程	鳥屋 尾隆	北海道大学 触媒科学研究所 助教
中山 伸一	筑波大学 図書館情報メディア系 教授	清水 研一	北海道大学 触媒科学研究所 教授
白井 泰博	シネックスジャパン株式会社	吉田 将隆	九州大学 先導物質化学研究所
浅原 彰規	株式会社日立製作所 研究開発グループ 人工知能イノベーションセンタ知能情報研究部 主任研究員	斎藤 雅史	九州大学 先導物質化学研究所
西浦 廉政	北海道大学 電子科学研究所 名誉教授	辻 雄太	九州大学 先導物質化学研究所 助教
赤城 和人	東北大学 材料科学高等研究所 准教授	蒲池 高志	福岡工業大学 工学部 教授
齋藤 健	先端素材高速開発技術研究組合 (ADMAT) 技術部 研究員	吉澤 一成	九州大学 先導物質化学研究所 教授
芹沢 貴之	旭化成ファーマ株式会社 (旧所属)	右田 啓哉	株式会社日本触媒 事業創出本部 データサイエンス&インフォマティクス推進室 室長
櫻井 篤	新潟大学 工学部工学科 機械システム工学プログラム 准教授	高橋 ローレンニコール	北海道大学大学院 理学研究院化学部門 学術研究員
津田 宏治	東京大学大学院 新領域創成科学研究科 メディカル情報生命専攻	杉本 学	熊本大学大学院 先端科学研究部 准教授
		田中 大佑生	熊本大学大学院 先端科学研究部
		鈴木 耕太	東京工業大学 物質理工学院 助教 JST さきがけ 研究員

注文書		メルマガ登録	登録済み	登録希望	お申込み・お問合せ
品名	材料およびプロセス開発のためのインフォマティクスの基礎と研究開発最前線	価格	定価（書籍）	90,000円＋税	
会社名			書籍＋CDセット	100,000円＋税	
部課名	TEL	メルマガ会員 書籍	81,000円＋税	編集発行： <b>(株)シーエムシー・リサーチ</b> 101-0054 東京都千代田区神田錦町2-7 東和錦町ビル3F	
お名前	FAX	書籍＋CDセット	90,000円＋税		
住所	E-mail				TEL: 03 (3293) 7053 FAX: 03 (3291) 5789
					URL: <a href="http://www.cmcre.com">http://www.cmcre.com</a> E-mail: <a href="mailto:re@cmcre.com">re@cmcre.com</a>

\*書籍はご注文を受けた翌営業日に納品書・請求書とともに送付します。\*お支払いには請求書指定口座に納品日の翌月末日までに振り込みをお願いします。

# 構成および内容

## 第I編 基礎(理論)

### 第1章 マテリアル・インフォマティクスに向けたデータベースの活用法 山野仁詩, 清水宏明, 小野直亮, Altaf-Ul-Amin, 黄 銘, 森田(平井) 晶, 金谷重彦

- はじめに
- 公開データベースを活用したマイニング法
- ホモポリマー物性相関
  - 物性量の相関
  - ホモポリマーの構造物性相関
- ポリマーブレンドデータの活用
- マテリアルズ・データベース

### 第2章 Rではじめるケモ・マテリアル・インフォマティクス:社内データの活用法! 金谷重彦

- はじめに
- 多変量データ解析
- Rパッケージ reshape
- Rパッケージ reshape2
- 多変量解析法を進めるためのRパッケージ

### 第3章 マテリアルズインフォマティクスの動向 知京豊裕, 木野日織, 小山幸典

- はじめに
- マテリアルインフォマティクス黎明期
- 新しいマテリアルズインフォマティクスの登場
  - アメリカにおけるデータ駆動型材料開発とハイスループット実験
  - 欧州におけるデータ駆動型材料開発研究
  - アジアにおけるデータ駆動型材料開発
  - 日本におけるデータ駆動型材料開発
- マテリアルズインフォマティクスのビジネス展開
- 今後のマテリアルインフォマテックスマテリアルインフォマティクスの将来に関するキー
- まとめ

## 第II編 最新研究・開発事例(MI向けAI/DB技術)

### 第1章 スパースモデリングによる物質・材料設計のための基盤技術の構築 五十嵐康彦

- マテリアルズインフォマティクスとスパースモデリング
- スパースモデリングの基礎
- スパースモデリングによる物質・材料設計への展開
- スパースモデリングを用いた計測インフォマティクスによる新規材料・物質の探索

### 第2章 化学物質名のNER研究の現状と課題 田中るみ子, 中山伸一

- はじめに
- 化学物質名のNER研究の概要
- 日本語文章における化学物質名のNER研究
  - 日本語の特許公開情報に対する化学物質名のNER研究 (1) 化学文章コーパスの作成 (2) 形態素解析による化学物質名を含む単語の切り出し (3) 化学物質名の識別
  - 日本語文章における化学物質名のNER研究の今後の展望
- おわりに

### 第3章 インフォマティクス計算に適した計算環境 白井泰博

- はじめに
- スケールアップ型サーバ
- スケールアウト型システム
- まとめ

### 第4章 AIの技術を活用するITソリューションとしてのマテリアルズインフォマティクス 浅原彰規

- はじめに
- 機械学習の技術とMI
- ITソリューションとしてのMI
- 材料特性を改善するためのパーソナルスクリーニング
- 材料データを分析するためのIT環境が要する機能

- システムの構成
- 分析と可視化
- おわりに

### 第5章 データ解析と数理モデル〜ミクロとマクロをつなぐ〜 西浦康政, 赤木和人

- はじめに
- 静的記述子
- 動的記述子
- まとめ

### 第6章 拡張OCTA:高分子複合材料のマテリアルズ・インフォマティクスを志向したOCTAの機能拡張 齋藤 健

- はじめに:OCTAと拡張OCTA
- 拡張OCTAの構成と機能
  - 拡張GOURMET
  - Image Loader
  - AI Tool
- 拡張OCTAの入手方法
- まとめ

## 第III編 最新研究・開発事例(ケモインフォマティクス,有機低分子/医薬)

### 第1章 旭化成ファーマにおけるAI/機械学習の創薬研究への活用 芹沢貴之

- はじめに
- AIを活用する創薬プロセス
- 創薬プロジェクトへの活用事例
  - ジェネレーターの性能検証
  - Ligand Based Drug Design (LDBB)への応用事例
  - Structure Based Drug Design (SBDD)への応用事例
  - 逆合成解析ツールの開発
- 現状のAI創薬における課題
  - 予測モデルの精度
- おわりに

### 第2章 深層学習による溶解度予測と深層学習モデルの化学的解釈 船津公人

- 序論
- 手法
  - ニューラルネットワークと深層学習
  - 高次元データの可視化
  - 提案手法
- 結果と考察
  - データセット
  - モデルのパラメータ
  - 予測精度の比較
  - Isomapを用いた可視化
- 結論

## 第IV編 最新研究・開発事例(無機材料開発)

### 第1章 機械学習を用いた超狭帯域熱放射多層膜の開発 櫻井 篤, 津田宏治, 塩見淳一郎

- はじめに
- COMBOによる超狭帯域熱放射デバイスの設計
  - 最適化条件
  - COMBOで予想された最適構造とその実証実験について
- 現象の考察
- まとめ

### 第2章 説明可能AIを用いた熱電材料開発 岩崎悠真

- 説明可能AIとマテリアルズ・インフォマティクス
- スピニ熱電材料
- 説明可能AIを用いたスピニ熱電材料開発
- おわりに

### 第3章 無機材料の組成式を元にした物性予測のための記述子開発 船津公人

- 緒言
- 手法
  - Random Forest
  - モデル評価指標
- 記述子
  - 元素の個数および割合に関する記述子
  - 各元素に付随する要素に関する記述子
  - 組成比に関する記述子
- ケーススタディ
  - 生成エネルギー予測 (1) データセットと前処理 (2) モデル構築に寄与する記述子 (3) 予測結果

- 密度の予測 (1) データセットと前処理 (2) モデル構築に寄与する記述子 (3) 予測結果
- 屈折率の予測 (1) データセットと前処理 (2) モデル構築に寄与する記述子 (3) 予測結果 (4) 各サンプルの予測

### 第4章 機械学習を用いた低コストかつ無害な新規熱電材料の出力向上 高際良樹

- 背景
- FAST材料
- 機械学習を用いたFAST材料の出力特性向上
- 今後の展望

### 第5章 計測インフォマティクスの登場と期待 宮里一旗, 高橋啓介

- 緒言
- 提案する自動実験計画法の概要
- 自動実験計画法の例証
  - 仮想的なCVD装置と実験結果
  - 自動実験計画策定過程の例証

### 第7章 マテリアルズ・インフォマティクスによる新超伝導物質の発見 松本 凌, 高野義彦

- はじめに
- 高圧力印加による物質の電子状態エンジニアリング
- 新規超伝導候補物質のスクリーニング
- 高効率な高圧力下電気抵抗測定
- 第一候補物質SnBi<sub>2</sub>Se<sub>4</sub>における圧力誘起超伝導の発見
- 第二候補物質PbBi<sub>2</sub>Te<sub>4</sub>における圧力誘起超伝導の発見
- おわりに

## 第V編 最新研究・開発事例(有機高分子)

### 第1章 構造用高分子材料の実用型最適設計・総合評価支援ツールの開発 藤元伸悦

- 高分子材料を対象としたマテリアルズエンジニアリング
- 研究開発の概要と主な成果
- 研究開発成果の活用と展望

### 第2章 分子設計・材料設計・プロセス設計のためのデータ駆動型化学材料設計・プロセス設計・品質管理と制御の運動 船津公人

- 材料設計の現状と課題
- プロセスも含めたポリマー材料設計戦略
- プロセス・インフォマティクスの展開

### 第3章 ベイズ最適化を活用した耐熱性ポリマーの効率的設計 南 拓也

- 緒言
- 手法
  - 機械学習モデルの構築
  - ポリマー設計効率の評価方法
- 結果
  - ポリマー物性予測モデルの性能評価
  - ポリマー設計効率の評価
- 結言

## 第VI編 最新研究・開発事例(触媒開発)

### 第1章 触媒インフォマティクスの創成に向けた実験・理論・データ科学研究 日沼洋陽, 蒲池高志, 鳥屋尾 隆, 清水研一

- TiO<sub>2</sub>表面における分子吸着のフロンティア軌道理論
- 酸化物表面における酸素欠陥生成エネルギーの第一原理計算
- 機械学習による吸着エネルギーの予測

### 第2章 合金表面上におけるメタン活性化の触媒インフォマティクス 吉田特隆, 齋藤雅史, 辻 雄太, 蒲池高志, 吉澤一成

- 序論

- 計算法
  - 結果と考察 (1) 学習データセットの作成
  - 結果と考察 (2) 回帰モデルを利用した触媒探索
- 総括

### 第3章 日本触媒におけるキャタリストインフォマティクスの適用事例 右田啓哉

- はじめに
- 選択性向上のための調製条件の最適化
  - 背景
  - データ
  - 解析手法
  - 結果と考察
  - 結論
- 触媒組成と性能への影響の推定
  - 背景
  - データ
  - 解析手法
  - 結果と考察
  - 結論
- おわりに

### 第4章 マテリアルズ・キャタリストインフォマティクスの概要と展望 高橋啓介, 高橋ローレンニコール

- データ
- データから知識
- プラットフォーム
- 今後の展望

## 第VII編 最新研究・開発事例(電池・電解質・電解液関連)

### 第1章 ペロブスカイト太陽電池用ホール輸送材料のキャリア移動度予測のための計算科学インフォマティクス研究 杉本 学, 田中大佑生

- はじめに
- キャリア輸送特性のための予測モデルの構築
  - 記述子の導入に関する理論的考察
  - エネルギー記述子の数値的評価
- 帰帰モデルの構築と予測性能
- まとめと今後の展望

### 第2章 リチウム導電体探索手法の発展と現状 鈴木耕太

- はじめに
- リチウム導電性固体電解質の探索
  - 古典的な探索手法 (1) 元素置換による探索 (2) 構造に基づく探索 (3) 組成に基づく探索
  - 理論計算やインフォマティクスを使ったアプローチ (1) 計算化学による構造を基にした探索 (2) 機械学習を活用した組成に基づく探索
- おわりに

## 第VIII編 今後の展望 ケモインフォマティクス・マテリアルズインフォマティクスの将来展望 船津公人

### 第IX編 参考資料 データ解析の進め方 船津公人

- データセットの表現
- 前処理
- シグマ法
- Hampel identifier
- Savitzky-Golay (SG) 法
- 主成分分析 (Principal Component Analysis, PCA)
- 独立成分分析 (Independent Component Analysis, ICA)
- 最小二乗法による線形重回帰分析
- Partial Least Squares (PLS)
- Support Vector Machine (SVM)
- Support Vector Regression (SVR)
- Online SVR (OSVR)
- Least Absolute Shrinkage and Selection Operator (LASSO) 法
- ステップワイズ法による変数選択
- Genetic Algorithm-based PLS (GAPLS)
- Genetic Algorithm-based WaveLength Selection (GAWLS)
- k-nearest neighbor (k-NN) 法
- One-Class SVM (OCSVM)
- 各種統計量