

メディアテクノロジーシリーズ 13

# Webテクノロジー

— 自然言語処理・検索・推薦・大規模言語モデル —

牛尾剛聡・佃 洸撰

【編著】

河合由起子・若宮翔子・中島伸介

【共著】

コロナ社

## メディアテクノロジーシリーズ 編集委員会

|                |                                  |
|----------------|----------------------------------|
| 編集委員長          | 近藤 邦雄 (元東京工科大学, 工学博士)            |
| 編集幹事           | 伊藤 貴之 (お茶の水女子大学, 博士 (工学))        |
| 編集委員<br>(五十音順) | 五十嵐悠紀 (お茶の水女子大学, 博士 (工学))        |
|                | 稲見 昌彦 (東京大学, 博士 (工学))            |
|                | 牛尼 剛聡 (九州大学, 博士 (工学))            |
|                | 大淵 康成 (東京工科大学, 博士 (情報理工学))       |
|                | 竹島由里子 (東京工科大学, 博士 (理学))          |
|                | 鳴海 拓志 (東京大学, 博士 (工学))            |
|                | 馬場 哲晃 (東京都立大学, 博士 (芸術工学))        |
|                | 日浦 慎作 (兵庫県立大学, 博士 (工学))          |
|                | 松村誠一郎 (東京工科大学, 博士 (学際情報学))       |
|                | 三谷 純 (筑波大学, 博士 (工学))             |
|                | 三宅陽一郎 (株式会社スクウェア・エニックス, 博士 (工学)) |
|                | 宮下 芳明 (明治大学, 博士 (知識科学))          |

(2023 年 5 月現在)

## 編著者・執筆者一覧

|     |                  |                      |
|-----|------------------|----------------------|
| 編著者 | 牛尼 剛聡 (2 章, 6 章) | 佃 洸撰 (2 章, 4 章, 5 章) |
| 執筆者 | 河合由起子 (1 章)      | 若宮 翔子 (3 章)          |
|     | 中島 伸介 (6 章)      |                      |

## 刊行のことば

“Media Technology as an Extension of the Human Body and the Intelligence”

「メディアはメッセージである (The medium is the message)」というマクルーハン (Marshall McLuhan) の言葉は、多くの人々によって引用される大変有名な言葉である。情報科学や情報工学が発展し、メディア学が提唱されたことでメディアの重要性が認識されてきた。このような中で、マクルーハンのこの言葉は、つねに議論され、メディア学のあるべき姿を求めてきたといえる。

人間の知的コミュニケーションを助けることができるメディアは生きていくうえで欠かせない。このようなメディアは人と人との関係をより良くし、視野を広げ、新しい考え方に目を向けるきっかけを与えてくれる。

また、マクルーハンは「メディアはマッサージである (The medium is the massage)」ともいっている。マッサージは疲れた体をもみほぐし、心もリラックスさせるが、メディアは凝り固まった頭にさまざまな情報を与え、考え方を広げる可能性があるため、マッサージという言葉はメディアの特徴を表しているともいえるだろう。

さらにマクルーハンは“人間の身体を拡張するテクノロジー”としてメディアをとらえ、人間の感覚や身体的な能力を変化させ、社会との関わりについて述べている。現在、メディアは社会、生活のあらゆる場面に存在し、五感を通してさまざまな刺激を与え、多くの技術が社会生活を豊かにしている。つまり、この身体拡張に加え、人工知能技術の発展によって“知能拡張”がメディアテクノロジーの重要な役割を持つと考えられる。このために物理的な身体と情報や知識を扱う知能を融合した“人間の身体と知能を拡張するメディアテクノロジー”を提案・開発し、これらの技術を活用して社会の構造や仕組みを変革し、

どのような人にとっても住みやすく、生活しやすい社会を目指すことが望まれている。

一方、大学におけるメディア学の教育は、東京工科大学が1999年にメディア学部を設置して以来、全国の大学でメディア関連の学部や学科が設置され文理芸分野を融合した多様な教育内容が提供されている。その体系化が期待されメディア学に関する教科書としてコロナ社から「メディア学大系」が発刊された。この第一巻の『改訂メディア学入門』には、メディアの基本モデルの構成として「情報の送り手、伝達対象となる情報の内容（コンテンツ）、伝達媒体となる情報の形式（コンテナ）、伝達形式としての情報の提示手段（コンベア）、情報の受け手」と書かれている。これからわかるようにメディアの基本モデルには文理芸に関連する多様な内容が含まれている。

メディア教育が本格的に開始され20年を過ぎるいま、多くの分野でメディア学のより高度で急速な展開が見られる。文理芸の融合による総合知によって人間生活や社会を理解し、より良い社会を築くことが必要である。

そこで、このメディア分野の研究に関わる大学生、大学院生、さらには社会人の学修のため「メディアテクノロジーシリーズ」を計画した。本シリーズは“人間の身体と知能を拡張するメディアテクノロジー”を基礎として、コンテンツ、コンテナ、コンベアに関する技術を扱う。そして各分野における基本的なメディア技術、最近の研究内容の位置づけや今後の展開、この分野の研究をするために必要な手法や技術を概観できるようにまとめた。本シリーズがメディア学で扱う対象や領域を発展させ、将来の社会や生活において必要なメディアテクノロジーの活用方法を見出す手助けとなることを期待する。

本シリーズの多様で広範囲なメディア学分野をカバーするために、電子情報通信学会、情報処理学会、人工知能学会、日本ソフトウェア科学会、日本バーチャルリアリティ学会、ヒューマンインタフェース学会、日本データベース学会、映像情報メディア学会、可視化情報学会、画像電子学会、日本音響学会、芸術科学会、日本図学会、日本デジタルゲーム学会、ADADA Japan などにおいて第一線で活躍している研究者の方々に編集委員をお願いし、各巻の執筆者選

定、目次構成、執筆内容など検討を重ねてきた。

本シリーズの読者が、新たなメディア分野を開拓する技術者、クリエイター、研究者となり、新たなメディア社会の構築のために活躍されることを期待するとともにメディアテクノロジーの発展によって世界の人達との交流が進み相互理解が促進され、平和な世界へ貢献できることを願っている。

2023 年 5 月

編集委員長 近藤邦雄

編集幹事 伊藤貴之

### 表紙・カバーデザインについて

私たちは五感というメディアを介して世界を知覚し、自己の存在を認知することができます。メディア技術の進歩によって五感が拡張され続ける中、「人」はなにをもって「人」と呼べるのか、そんな根源的な問いに対する議論が絶えません。

本書の表紙・カバーデザインでは、二値化された五感が新しい機能や価値を再構築する様子をシンプルなストライプ柄によって表現しました。それぞれのストライプは 5 本のゆらぎを持った線によって描かれており、手描きのような印象を残しました。

しかし、この細かなゆらぎもプログラム制御によって生成されており、十分に細かく量子化された表現によって「デジタル」と「アナログ」それぞれの存在がゆらぐ様子を表しています。乱雑に描かれたストライプをよく観察してみてください。本書を手にとった皆さんであれば、きっともう一つ面白いことに気づくでしょう。

デザインを検討するにあたって、同じコンセプトに基づき、いくつかのグラフィックパターンを生成可能なウェブアプリケーションを準備しました。下記 URL にて公開していますので、あなただけのカバーを作ってみてください。読者の数だけカバーデザインが存在するのです。世界はあなたの五感を通じて存在しているのですから。

馬場哲晃

〈Cover Generator〉 ぜひお試しください

<https://tetsuakibaba.github.io/mtcg/>

(2023 年 5 月現在)



# ま え が き

インターネットと Web テクノロジーは、われわれの日常生活やビジネス活動に深く浸透し、その影響力は年々増大している。本書でいう「Web テクノロジー」とは、インターネット上でサービスや情報を提供・流通させるための基盤技術の総称であり、通信プロトコルやデータ処理だけでなく、自然言語処理・情報検索・推薦・大規模言語モデル（LLM）といった高度な情報処理技術をも包含する。情報通信、エンターテインメント、教育、医療、そしてビジネスの多くの側面が、これらの技術によって支えられ、革新され続けている。本書『Web テクノロジー』は、これらの急速に進化する技術の基礎から応用までを包括的に解説し、読者が現代の Web テクノロジーを深く理解するための道しるべとなることを目指すものである。

近年、大規模言語モデルを中核とする生成 AI の進展により、Web サービスの設計と実装を取り巻く前提が大きく変わりつつある。従来の検索や推薦に加え、対話型の情報アクセス、コンテンツ生成、パーソナライズといった機能が実用化し、Web サービスに求められる技術要素も多様化している。本書は、こうした変化を踏まえて、基盤となる原理と実装上の要点を体系的に学べる構成とした。その上で、Web テクノロジーに興味を持つすべての読者、特にこれからこの分野でのキャリアを目指す学生や技術者にとって貴重な知識と洞察を提供することを目的とする。

本書は、全 6 章から構成される。まず 1 章（河合）では、インターネットと Web の基礎を概観する。インターネット通信の仕組み、URL と HTTP、クライアント技術、Web サービス設計などを通じて、Web を構成する基盤技術の全体像を把握する。2 章（牛尼、佃）では、機械学習の基礎を扱う。機械学習は、現代の多くの Web テクノロジーを支える中核技術であり、その基本的な概念と手法の理解は、Web テクノロジー全般の理解を深める上で不可欠である。

3章（若宮）では自然言語処理を扱い、言語解析の基礎から実応用タスクまでを学ぶことで、人間の言語を計算機的に扱う原理と方法を習得する。さらに、4章（佃）の情報検索と5章（佃）の情報推薦では、膨大なデータからユーザにとって有用な情報を精度高く見つけ出すために必要な検索・推薦アルゴリズムや評価指標を詳述する。最後に6章（牛尼，中島）では、生成AIの基盤となる大規模言語モデルとその応用を取り上げ、大規模言語モデルがどのように開発され、どのように実世界の問題解決に活用できるのかを学ぶ。

本書で扱うおもな数式や記号の表記ルールは、つぎのとおりである。ベクトルや行列などの具体的な記号が表す内容は章によって異なり、各章の本文中で適宜説明しているため、ここでは例示のみとする。

- スカラー（通常の変数）：小文字の斜体で表記する（例： $x, y, a$ ）。
- ベクトル：小文字の太字で表記する（例： $\mathbf{x}, \mathbf{y}$ ）。
- 行列：大文字の太字で表記する（例： $\mathbf{A}, \mathbf{B}$ ）。
- 集合：大文字の斜体で表記する。実数や整数の集合には白抜き文字を用いる（例： $A, B, \mathbb{R}, \mathbb{Z}$ ）。
- 関数：1文字の関数名は斜体，2文字以上の関数名や三角関数・対数関数などは立体（ローマン体）で表記する（例： $f(x), \exp(x), \sin x, \log y$ ）。
- その他：ノルムは  $\|\mathbf{x}\|$ ，内積は  $\langle \mathbf{x}, \mathbf{y} \rangle$ ，和と積はそれぞれ  $\sum_{i=1}^n x_i, \prod_{i=1}^n a_i$  のように表記する。

最後に、本書の執筆機会を与えてくださった「メディアテクノロジーシリーズ」の編集委員会および、長期間にわたり献身的なご支援をいただいたコロナ社の皆さんに、心より謝意を表する。本書が読者の学びと発展に寄与し、Webテクノロジーの未来を切り開く一助となることを願う。

2026年6月

編著者 牛尼剛聡・佃 洸撰

<sup>†1</sup> 本書の書籍詳細ページ (<https://www.coronasha.co.jp/np/isbn/9784339013832/>) から URL リンク集などの補足情報がダウンロードできます。

<sup>†2</sup> 本書で使用している会社名，製品名は一般に各社の登録商標です。本書では® や ™ は省略しています。

<sup>†3</sup> 本書で紹介している URL は 2026 年 4 月現在のものです。



# 目 次

## 第 1 章

### Web 技術の基礎と仕組み

|            |                        |           |
|------------|------------------------|-----------|
| <b>1.1</b> | <b>インターネットと Web</b>    | <b>1</b>  |
| <b>1.2</b> | <b>クライアント・サーバモデル</b>   | <b>6</b>  |
| 1.2.1      | URL                    | 8         |
| 1.2.2      | HTTP リクエストと HTTP レスポンス | 13        |
| 1.2.3      | HTTP バージョンの進化          | 15        |
| <b>1.3</b> | <b>Web クライアント</b>      | <b>16</b> |
| 1.3.1      | HTML と Web ページ         | 17        |
| 1.3.2      | CSS                    | 18        |
| 1.3.3      | JavaScript             | 19        |
| <b>1.4</b> | <b>Web のデータ通信と構造</b>   | <b>21</b> |
| 1.4.1      | XML と JSON             | 21        |
| 1.4.2      | Ajax                   | 23        |
| 1.4.3      | SOAP                   | 25        |
| 1.4.4      | REST                   | 27        |
| 1.4.5      | OAuth                  | 29        |
| <b>1.5</b> | <b>モダン Web 開発の潮流</b>   | <b>30</b> |
| 1.5.1      | クロスプラットフォーム開発フレームワーク   | 33        |
| 1.5.2      | クラウドサービス               | 34        |
| <b>1.6</b> | <b>ま と め</b>           | <b>35</b> |



## 第 2 章

## 機 械 学 習

|                          |    |
|--------------------------|----|
| 2.1 機械学習とは.....          | 37 |
| 2.2 教師あり学習.....          | 39 |
| 2.2.1 学 習 の 流 れ          | 39 |
| 2.2.2 分 類 の 手 法          | 40 |
| 2.2.3 分類の評価指標            | 51 |
| 2.2.4 回 帰 の 手 法          | 57 |
| 2.2.5 回帰の評価指標            | 61 |
| 2.3 教師なし学習.....          | 62 |
| 2.3.1 階層的クラスタリング         | 63 |
| 2.3.2 K-means            | 64 |
| 2.3.3 DBSCAN             | 66 |
| 2.4 深 層 学 習.....         | 68 |
| 2.4.1 ニューラルネットワークの基本構造   | 68 |
| 2.4.2 パラメータの学習           | 71 |
| 2.5 ま      と      め..... | 72 |

## 第 3 章

## 自然言語処理

|                           |    |
|---------------------------|----|
| 3.1 テキストデータの扱い方と基礎概念..... | 73 |
| 3.1.1 テキストの前処理            | 73 |
| 3.1.2 言語データの統計的性質         | 76 |
| 3.2 形 態 素 解 析.....        | 77 |
| 3.2.1 形態素解析とは             | 77 |

|            |                      |            |
|------------|----------------------|------------|
| 3.2.2      | 形態素解析器               | 78         |
| <b>3.3</b> | <b>固有表現抽出</b> .....  | <b>80</b>  |
| 3.3.1      | 固有表現抽出とは             | 80         |
| 3.3.2      | 固有表現抽出の方法            | 81         |
| 3.3.3      | 応用と発展                | 82         |
| <b>3.4</b> | <b>構文解析</b> .....    | <b>83</b>  |
| 3.4.1      | 構文解析とは               | 83         |
| 3.4.2      | 構文解析の種類              | 83         |
| 3.4.3      | 構文解析器と出力例            | 85         |
| <b>3.5</b> | <b>意味解析</b> .....    | <b>87</b>  |
| 3.5.1      | 意味解析とは               | 87         |
| 3.5.2      | 単語レベルの課題：語義の曖昧性解消    | 87         |
| 3.5.3      | 文レベルの課題：意味役割付与       | 87         |
| 3.5.4      | 単語埋め込み               | 88         |
| 3.5.5      | 伝統的な意味表現：統計的手法とベクトル化 | 91         |
| 3.5.6      | 意味解析のための知識資源の活用      | 93         |
| <b>3.6</b> | <b>応用タスク</b> .....   | <b>95</b>  |
| 3.6.1      | 文書分類                 | 95         |
| 3.6.2      | 機械翻訳                 | 97         |
| 3.6.3      | 要約                   | 99         |
| 3.6.4      | 質問応答                 | 101        |
| <b>3.7</b> | <b>言語資源</b> .....    | <b>103</b> |
| 3.7.1      | コーパス                 | 104        |
| 3.7.2      | 辞書                   | 104        |
| 3.7.3      | 言語資源の整備と共有           | 106        |
| <b>3.8</b> | <b>まとめ</b> .....     | <b>108</b> |

## 第 4 章

## 情 報 検 索

|                                   |     |
|-----------------------------------|-----|
| 4.1 情報検索の流れ.....                  | 109 |
| 4.2 転置インデックス .....                | 111 |
| 4.3 単語の出現傾向を用いたランキング手法 .....      | 114 |
| 4.3.1 ベクトル空間モデル .....             | 114 |
| 4.3.2 TF-IDF .....                | 116 |
| 4.3.3 BM25 .....                  | 117 |
| 4.4 リンク構造を用いたランキング手法.....         | 118 |
| 4.4.1 PageRank .....              | 119 |
| 4.4.2 HITS .....                  | 122 |
| 4.5 機械学習を用いたランキング手法 .....         | 124 |
| 4.5.1 ラ ン ク 学 習 .....             | 124 |
| 4.5.2 深層学習に基づく Web ページランキング ..... | 127 |
| 4.6 情報検索の評価.....                  | 131 |
| 4.6.1 Precision .....             | 132 |
| 4.6.2 Recall .....                | 133 |
| 4.6.3 F-measure .....             | 133 |
| 4.6.4 MAP .....                   | 133 |
| 4.6.5 MRR .....                   | 135 |
| 4.6.6 nDCG .....                  | 135 |
| 4.6.7 その他の評価指標・評価方法 .....         | 137 |
| 4.7 クエリ自動補完・クエリ推薦 .....           | 138 |
| 4.7.1 クエリ自動補完 .....               | 139 |
| 4.7.2 ク エ リ 推 薦 .....             | 141 |
| 4.8 ま と め.....                    | 144 |

## 第 5 章

## 情 報 推 薦

|                                    |     |
|------------------------------------|-----|
| <b>5.1 Web における情報推薦</b> .....      | 146 |
| <b>5.2 ユーザの嗜好表現</b> .....          | 148 |
| 5.2.1 明示的フィードバック                   | 148 |
| 5.2.2 暗黙的フィードバック                   | 150 |
| <b>5.3 アイテムの内容に基づく情報推薦</b> .....   | 151 |
| 5.3.1 メタ情報に基づく情報推薦                 | 151 |
| 5.3.2 特徴量に基づく情報推薦                  | 153 |
| <b>5.4 協調フィルタリングに基づく情報推薦</b> ..... | 155 |
| 5.4.1 ユーザ間型メモリベース法                 | 155 |
| 5.4.2 アイテム間型メモリベース法                | 157 |
| 5.4.3 モデルベース法：MF                   | 159 |
| 5.4.4 モデルベース法：BPR                  | 161 |
| 5.4.5 モデルベース法：FM                   | 162 |
| <b>5.5 深層学習を用いた情報推薦</b> .....      | 165 |
| 5.5.1 ニューラルネットワークに基づく推薦            | 165 |
| 5.5.2 グラフニューラルネットワークに基づく推薦         | 166 |
| 5.5.3 時系列に基づく推薦                    | 168 |
| <b>5.6 ハイブリッド型の情報推薦</b> .....      | 169 |
| 5.6.1 混 合                          | 170 |
| 5.6.2 切 り 替 え                      | 171 |
| 5.6.3 重 み 付 け                      | 172 |
| 5.6.4 カ ス ケ ード                     | 173 |
| <b>5.7 推薦手法の評価</b> .....           | 174 |
| 5.7.1 オフライン評価                      | 174 |

|       |            |     |
|-------|------------|-----|
| 5.7.2 | オンライン評価    | 180 |
| 5.7.3 | ユーザスタディ    | 184 |
| 5.7.4 | 三つの評価方法の比較 | 186 |
| 5.8   | ま と め      | 188 |

## 第 6 章

### 大規模言語モデルとその応用

|       |                            |     |
|-------|----------------------------|-----|
| 6.1   | 言語モデルの概念と歴史的背景             | 190 |
| 6.1.1 | 初期の言語モデル                   | 191 |
| 6.1.2 | ニューラル言語モデル                 | 192 |
| 6.1.3 | リカレントニューラルネットワーク (RNN)     | 194 |
| 6.1.4 | Seq2Seq                    | 198 |
| 6.1.5 | 注 意 機 構                    | 200 |
| 6.2   | Transformer                | 202 |
| 6.2.1 | 自己注意機構                     | 202 |
| 6.2.2 | 位置エンコーディング                 | 208 |
| 6.2.3 | エンコーダ-デコーダアーキテクチャ          | 209 |
| 6.3   | Transformer を基盤とした言語モデル    | 212 |
| 6.3.1 | GPT                        | 212 |
| 6.3.2 | BERT                       | 214 |
| 6.3.3 | 構造と応用の比較                   | 216 |
| 6.4   | Transformer を基盤とした言語モデルの学習 | 216 |
| 6.4.1 | 事 前 学 習                    | 217 |
| 6.4.2 | ファインチューニング                 | 219 |
| 6.4.3 | 人間の意図との整合：アライメントと RLHF     | 222 |
| 6.5   | LLM を活用したアプリケーション          | 223 |
| 6.5.1 | プロンプトエンジニアリングと出力制御         | 223 |

|         |                     |     |
|---------|---------------------|-----|
| 6.5.2   | 自然言語インタフェースとしての LLM | 228 |
| 6.5.3   | 検索拡張生成 (RAG)        | 231 |
| 6.5.4   | Web アプリケーションにおける活用例 | 234 |
| 6.6     | LLM の 課 題           | 235 |
| 6.7     | ま と め               | 237 |
| 引用・参考文献 |                     | 240 |
| 索 引     |                     | 250 |

# 第 1 章

## Web 技術の基礎と仕組み

1990 年 12 月、ジュネーブにある欧州原子核研究機構（CERN）において、世界初の Web サイト <http://info.cern.ch/> が公開された。その後、Yahoo!, Google, YouTube, Twitter（現 X）, Instagram などのサービスが普及するにつれて、World Wide Web（WWW、以下 Web と呼ぶ）は急速に拡大し、2026 年頃には、世界の Web サイト数は 14 億を超えたと報告されている。これほどの規模となった Web は、サイバー空間とフィジカル空間の融合技術として、学術分野だけでなく社会構造を革新した。本章では、この Web の仕組みを支えるインターネットの構造とクライアント・サーバモデルを中心に、Web 技術の基本的な構成要素を体系的に学ぶ。まず、1.1 節でインターネットにおけるデータ通信の流れを、1.2 節で Web の基本モデルとなる URL の構造や HTTP 通信の仕組みを解説する。つぎに 1.3 節では、クライアント側の技術として、HTML, CSS, JavaScript を取り上げ、ユーザインタフェースの構築方法を学ぶ。さらに 1.4 節で、Web におけるデータ交換の基礎となるデータ形式や、Web サービスの設計手法などを紹介し、最後に 1.5 節では、モダンな Web 開発の基礎を学び、サイバーフィジカルシステムやデジタルツインに必要な Web サービス構築技術の全体像を把握することを目指す。

### 1.1 インターネットと Web

**World Wide Web（WWW）** は、インターネット上で動作するハイパーメディアである。文書や画像などの情報資源をハイパーリンクで相互に結び、利用者がリンクをたどることで情報を探索できる仕組みを提供する。本書では、WWW を略して **Web** と呼ぶ。

Web は「世界規模の網の目」という語感から、インターネットと混同されやすい。インターネットは情報を運ぶための通信インフラであり、Web はその上で情報を提供・参照するための仕組みである。つまり、インターネットが「道路網」だとすれば、Web はその道路網の上で提供される「サービス」に相当する。

Web が成立するためには、世界規模で端末同士が通信できる基盤が必要であり、その基盤がインターネットである。インターネットは、インターネットプロトコル (Internet Protocol, **IP**) を中心としたプロトコル群により、複数のネットワークが相互接続されて形成される広域の情報通信網である。各端末やサーバは、IP アドレスによって自身および通信相手を識別し、データをパケット (packet) として送受信する。図 1.1 にインターネットでのデータ送信の流れを示す。

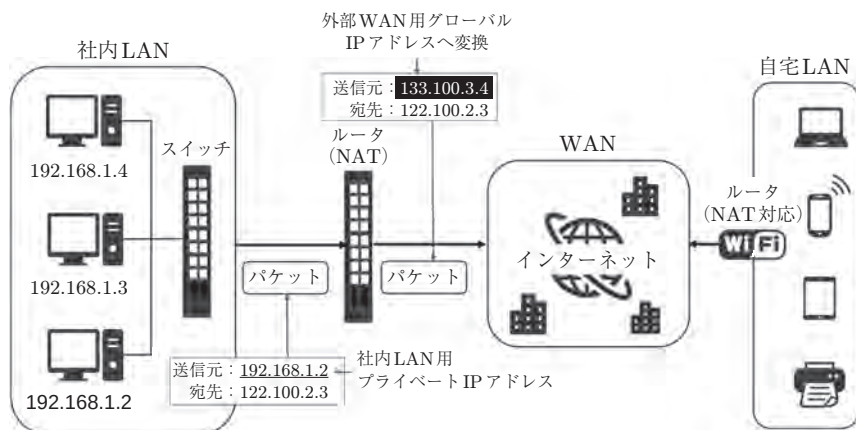


図 1.1 情報通信網となるインターネット (WAN) への IP によるデータ送信の流れ

一般に、建物のフロアや部署内などに限定されたネットワークをローカルエリアネットワーク (local area network, **LAN**) と呼び、それらをルータなどにより相互接続して広域化したネットワークを広域ネットワーク (wide area network, **WAN**) と呼ぶ。インターネットでは、このような複数のネットワークが相互接続され、組織や拠点をまたいだ通信が行われる。



IP アドレスとは、IP において通信相手を識別するための番号である。図 1.1 では、社内 LAN で割り振られたプライベート IP アドレスの端末から送出されたパケットが、境界装置（ルータ）で **NAT**（network address translation）によりグローバル（パブリック）IP アドレスに対応付けられ、外部ネットワークへ転送される流れを示している。

インターネットの通信では、IP のほかにも TCP や UDP、さらに SMTP, SSH, HTTP など多数のプロトコルが利用される。これらはそれぞれ役割が異なり、実際の通信は複数のプロトコルを重ね合わせて実現される。このような役割分担を整理する考え方が**階層（レイヤ）**であり、代表的な整理方法として**TCP/IP**がある。TCP/IP では、通信機能を四つの階層に分けて役割を分担している（表 1.1）。

**表 1.1** TCP/IP の 4 階層モデル

| 階層名                  | おもな役割                     | 代表的なプロトコル、技術                                   |
|----------------------|---------------------------|------------------------------------------------|
| アプリケーション層            | ユーザにサービスを提供               | HTTP, HTTPS, SMTP, FTP, SSH, DNS, Telnet, SNMP |
| トランスポート層             | 通信の信頼性とデータの分割、再構成を管理      | TCP (コネクション型), UDP (コネクションレス)                  |
| インターネット層             | データのルーティングとアドレッシングを担当     | IP (IPv4, IPv6), ICMP, IGMP, ARP               |
| ネットワークインタフェース層（リンク層） | 実際の物理的な通信処理（LAN やデバイス間通信） | Ethernet, Wi-Fi, PPP                           |

以下では、インターネットにおけるデータ送受信の流れを、各通信プロトコルとともに上位階層（アプリケーション層）から順に見ていく。

まず、最上位の**アプリケーション層**では、Web やメールなどのサービスでやりとりするメッセージ形式や手順を定めた、表 1.1 に示されているアプリケーションプロトコルを利用する。例えば、メール送信の **SMTP** やファイル転送の **FTP**、Web の **HTTP** といったプロトコルがある。これらのメールや Web などの通信では、**ポート番号**により、受信側ホスト上のどのアプリケーション（プロセス）にデータを渡すかを識別する。ポート番号は 0 から 65535

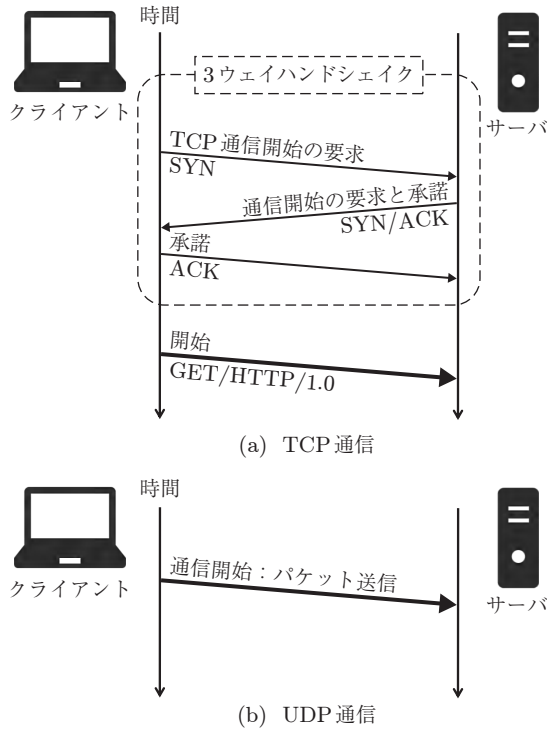
までの範囲があり、このうち 0~1023 は IANA (Internet Assigned Numbers Authority) により管理されるウェルノウンポート (well-known port) で、代表的なサービスに割り当てられている。例えば、SMTP (メール送信) にはポート番号 **25**, HTTP (Web 通信) には **80**, HTTPS (暗号化通信) には **443** が使用される。

つぎに、トランスポート層では、アプリケーション間の通信の信頼性を確保する。**TCP** プロトコルまたは **UDP** プロトコルを使用し、上位層のアプリケーション要求を受け取り、TCP ではデータの送信元と送信先間でコネクションを確立し、制御情報とともにデータを下位層へ渡す。なお、受信側では TCP/UDP ヘッダなどの制御情報を解釈し、宛先のアプリケーション (プロセス) へデータを引き渡す。図 1.2 に、TCP と UDP によるデータ送信の流れを示す。TCP では、信頼あるコネクションを確立するために、**3 ウェイハンドシェイク** (3-way handshake) と呼ばれる、以下の 3 段階の手順を踏む。

1. クライアントがサーバに、SYN フラグを設定した TCP セグメントを送信し、接続を要求する。
2. サーバはそれを受信すると、承諾 (ACK) に加えて自身の接続要求 (SYN) も示すため、SYN/ACK フラグを設定した TCP セグメントを返信する。
3. クライアントはサーバの SYN に対する承諾として、ACK フラグを設定した TCP セグメントを返信し、コネクションが確立する。

UDP はコネクションを確立せずにデータ (データグラム) を送信するため、到達保証や順序保証は行わない。TCP は通信の正確性と順序性を保証する一方、UDP は高速性を重視し、リアルタイム性が求められる音声や映像などの通信に適している。

インターネット層では、ネットワーク間のデータ転送を担い、IP アドレスによるルーティングを行う。ルータや L3 スイッチなどを通して、IP アドレスに基づいてパケットを転送し、ルーティングプロトコルなどにより得られた経路情報に従って宛先へ到達させる (図 1.1)。ICMP はネットワーク診断 (例: ping) に用いられ、ARP は IP アドレスから MAC アドレス (各通信機器に付与され



(a) TCP 通信は 3 ウェイハンドシェイクによりコネクションを確立し、再送や順序制御などにより信頼性を確保する。(b) UDP 通信はコネクション確立を行わず、必要最小限のヘッダでパケット（データグラム）を送るため、オーバーヘッドが小さい。

図 1.2 TCP と UDP による通信

た世界で一意な番号）を取得するために使われる。

最下位のネットワークインタフェース層は、物理的に接続された機器間でビット列を信号（光、電気、無線）に変換し、光ケーブルや LAN ケーブル、無線などでデータを送り、受信側で信号をビット列に戻す役割で、**Ethernet** や **PPP** などの技術が用いられる。

# 索引

|               |                |          |
|---------------|----------------|----------|
| 【あ】           | エンコーダーデコーダ注意機構 | 切り替え     |
| アイテム間型メモリベース法 | 211            | キーワード検索  |
| 157           | エンコーディング       | 74       |
| アクセストークン      | エンティティリンキング    | 82       |
| 29            | エントロピー         | 42       |
| アノテーション       | 【お】            |          |
| 104           |                |          |
| アノテータ         | オーソリティ         | 122      |
| 104           | オフセット          | 111      |
| アプリケーション層     | オフライン評価        | 174      |
| 3             | 重み付け           | 172      |
| アライメント        | オンライン評価        | 180      |
| 222           |                |          |
| アンサンブル学習      |                |          |
| 44            |                |          |
| 暗黙のフィードバック    |                |          |
| 150           |                |          |
| 【い】           |                |          |
|               | 【か】            |          |
| 言い換え          |                |          |
| 97            | 回帰             | 38, 57   |
| 依存構造解析        | 階層             | 3        |
| 84            | 階層のクラスタリング     | 63       |
| 位置エンコーディング    | 過学習            | 40       |
| 202, 208      | 確率の勾配降下法       | 159      |
| 一方向自己注意機構     | 隠れ層            | 69       |
| 212           | カスケード          | 173      |
| イベントハンドラ      | カスタムスキーム       | 10       |
| 20            | 活性化関数          | 70       |
| 意味解析          | カーネル関数         | 49       |
| 87            | カーネルトリック       | 48       |
| 意味役割付与        | 【き】            |          |
| 88            |                |          |
| 因果マスク         | 機械翻訳           | 97       |
| 217           | 逆文書頻度          | 92       |
| インターネット層      | 教師あり学習         | 38, 124  |
| 4             | 教師なし学習         | 38, 62   |
| インターネットプロトコル  | 協調フィルタリング      | 147, 155 |
| 2             | 局所表現           | 91       |
| インターリーピング     |                |          |
| 182           |                |          |
| インデックス        |                |          |
| 111           |                |          |
| 【え】           |                |          |
|               |                |          |
| 永続的識別子        |                |          |
| 9             |                |          |
| エッジ           |                |          |
| 41            |                |          |
| エンコーダ         |                |          |
| 198           |                |          |
| エンコーダーデコーダアーキ |                |          |
| 202           |                |          |
| テクチャ          |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |
|               |                |          |

|              |     |
|--------------|-----|
| 広域ネットワーク     | 2   |
| 交差エントロピー損失   | 60  |
| 構文解析         | 83  |
| 語義の曖昧性解消     | 87  |
| コーパス         | 104 |
| 固有表現抽出       | 81  |
| コールドスタート問題   | 169 |
| 混 合          | 170 |
| コンスティテューショナル |     |
| AI           | 222 |
| コンテキストベクトル   | 199 |
| 混同行列         | 51  |

【さ】

|             |         |
|-------------|---------|
| 再現率         | 53, 133 |
| 最尤推定法       | 60      |
| サブワード分割     | 76      |
| サポートベクトルマシン | 46      |
| 残 差         | 58      |
| 残差接続        | 211     |
| サンプリング      | 214     |

【し】

|          |              |
|----------|--------------|
| シェアードタスク | 106          |
| 自己注意機構   | 202          |
| 事前学習     | 216, 217     |
| 自然言語処理   | 73           |
| シソーラス    | 93           |
| ジップの法則   | 76           |
| 質問応答     | 101          |
| 次トークン予測  | 217          |
| ジニ不純度    | 42           |
| 重回帰      | 58           |
| 出現頻度     | 92           |
| 述語項構造解析  | 88           |
| 出力層      | 69           |
| 情報検索     | 109          |
| 情報推薦     | 146          |
| 情報要求     | 109          |
| 新規性      | 178          |
| 深層学習     | 68, 127, 165 |

【す】

|        |     |
|--------|-----|
| 推移確率行列 | 121 |
| 推薦システム | 146 |

【せ】

|           |     |
|-----------|-----|
| 正解率       | 52  |
| 正規化       | 82  |
| 生成型・抽象型要約 | 99  |
| 正則化       | 59  |
| 正則化項      | 159 |
| セッション     | 141 |
| セレンディピティ  | 179 |
| 線形回帰      | 57  |
| 全結合層      | 69  |
| 全文検索      | 111 |

【そ】

|             |     |
|-------------|-----|
| 層正規化        | 211 |
| 双方向自己注意機構   | 214 |
| ソフトマージン SVM | 48  |
| 損失関数        | 58  |

【た】

|           |        |
|-----------|--------|
| 大規模言語モデル  | 216    |
| 多クラス分類    | 40, 96 |
| 多項式回帰     | 57     |
| 多層パーセプトロン | 68     |
| ターム辞書     | 113    |
| 多様性       | 178    |
| 単回帰       | 57     |
| 単語埋め込み    | 88     |
| ダンピングファクタ | 120    |

【ち】

|         |     |
|---------|-----|
| チェインルール | 191 |
| 逐次検索    | 111 |
| 知識グラフ   | 94  |
| チャンク    | 232 |
| 注意機構    | 200 |
| 中間層     | 69  |
| 抽出型要約   | 99  |

【て】

|          |         |
|----------|---------|
| ディープリンク  | 10      |
| 適合性      | 110     |
| 適合率      | 53, 132 |
| デコーダ     | 198     |
| デコーディング  | 212     |
| テストデータ   | 39, 174 |
| デフォルトポート | 10      |
| 転移学習     | 219     |
| 転置インデックス | 111     |

【と】

|          |     |
|----------|-----|
| トークン     | 75  |
| トークン化    | 75  |
| トランスポート層 | 4   |
| 貪欲法      | 213 |

【な】

|            |     |
|------------|-----|
| ナイーブベイズ分類器 | 49  |
| 生コーパス      | 104 |

【に】

|             |         |
|-------------|---------|
| 二値分類        | 95      |
| 入力層         | 69      |
| ニューラル言語モデル  | 193     |
| ニューラルネットワーク | 68, 165 |
| ニューロン       | 69      |

【ね】

|           |    |
|-----------|----|
| ネットワークインタ |    |
| フェース層     | 5  |
| 根ノード      | 41 |

【の】

|     |        |
|-----|--------|
| ノード | 41, 69 |
|-----|--------|

【は】

|           |         |
|-----------|---------|
| ハイパーテキスト  | 6       |
| ハイパーパラメータ | 40, 175 |
| ハイパーメディア  | 1       |
| ハイパーリンク   | 7       |

## ハイブリッド型の情報推薦

|             |               |
|-------------|---------------|
|             | 169           |
| パケット        | 2             |
| バックエンド      | 21            |
| ハードマージン SVM | 48            |
| 葉ノード        | 41            |
| ハ ブ         | 122           |
| ハルシネーション    |               |
|             | 100, 227, 236 |

## 【ひ】

|             |     |
|-------------|-----|
| 非階層的クラスタリング | 64  |
| 被覆率         | 179 |
| ビームサーチ      | 213 |
| 評価値         | 148 |
| 評価値行列       | 148 |

## 【ふ】

|               |          |
|---------------|----------|
| ファインチューニング    |          |
|               | 216, 219 |
| フィードフォワード層    | 211      |
| 深 さ           | 41       |
| フラグメント識別子     | 11       |
| プロンプト         | 223      |
| プロンプトエンジニアリング |          |
|               | 223      |
| 分散表現          | 88       |
| 文書分類          | 95       |
| 分 類           | 38, 40   |

## 【へ】

|            |        |
|------------|--------|
| 平易化        | 97     |
| 平均絶対誤差     | 61     |
| 平均二乗誤差     | 58, 61 |
| ベクトル空間モデル  | 114    |
| ベクトル検索エンジン | 232    |

## 【ほ】

|           |     |
|-----------|-----|
| 補助情報      | 162 |
| ポスティングリスト | 113 |
| ポート番号     | 3   |

## 【ま】

|            |          |
|------------|----------|
| マイクロ平均     | 55       |
| マクロ平均      | 55       |
| マスク化       | 218      |
| マスク化自己注意機構 | 211      |
| マルチヘッド注意機構 |          |
|            | 206, 211 |
| マルチラベル分類   | 96       |

## 【め】

|            |     |
|------------|-----|
| 明示的フィードバック | 148 |
| メソッド       | 13  |

## 【も】

|         |    |
|---------|----|
| 文字符号化方式 | 74 |
|---------|----|

## 【ゆ】

|              |     |
|--------------|-----|
| 有向グラフ        | 118 |
| 尤 度          | 50  |
| 尤度関数         | 60  |
| ユーザ間型メモリベース法 |     |
|              | 155 |
| ユーザスタディ      | 184 |

## 【よ】

|     |    |
|-----|----|
| 要 約 | 99 |
|-----|----|

## 【ら】

|                |     |
|----------------|-----|
| ランク学習          | 124 |
| ランダムサーファーマodel |     |
|                | 120 |
| ランダムフォレスト      | 44  |

## 【り】

|            |     |
|------------|-----|
| リカレントニューラル |     |
| ネットワーク     | 194 |
| 利 得        | 136 |
| 隣接行列       | 121 |

## 【ろ】

|               |    |
|---------------|----|
| ローカルエリアネットワーク |    |
|               | 2  |
| ロジスティック回帰     | 59 |

## 【A】

|           |     |
|-----------|-----|
| A/B テスト   | 181 |
| AI エージェント | 238 |
| Ajax      | 23  |
| AUC       | 54  |

## 【B】

|          |              |
|----------|--------------|
| BaaS     | 31           |
| BERT     | 91, 130, 214 |
| BERT4Rec | 168          |

|      |        |
|------|--------|
| BLEU | 97     |
| BM25 | 117    |
| BoW  | 91, 92 |
| BPR  | 161    |

## 【C】

|      |    |
|------|----|
| CBOW | 90 |
| CSS  | 18 |

## 【D】

|         |    |
|---------|----|
| DBpedia | 94 |
|---------|----|

|             |     |
|-------------|-----|
| DBSCAN      | 66  |
| DCG         | 137 |
| DELETE メソッド | 15  |

## 【F】

|           |         |
|-----------|---------|
| F 値       | 53, 133 |
| fastText  | 90      |
| fetch API | 23      |
| FM        | 163     |
| FTP       | 3       |

|               |     |              |               |                |               |
|---------------|-----|--------------|---------------|----------------|---------------|
| <b>【G】</b>    |     | <b>【M】</b>   |               | SGD            | 159           |
| GET メソッド      | 13  | MAE          | 61, 176       | Skip-gram      | 90            |
| GloVe         | 90  | MAP          | 133           | SMTP           | 3             |
| GPT           | 212 | MeCab        | 78            | SOAP           | 26            |
| GRU           | 197 | MF           | 159           | softmax 関数     | 71            |
| GSK           | 107 | MLP          | 69            | SRL            | 88            |
|               |     | MPC          | 140           | Sudachi        | 80            |
|               |     | MRR          | 135           | SVM            | 46            |
| <b>【H】</b>    |     | <b>【N】</b>   |               | <b>【T】</b>     |               |
| HEAD メソッド     | 14  |              |               | TCP            | 4             |
| HITS          | 122 | n-gram モデル   | 76            | TCP/IP         | 3             |
| HTML          | 17  | NAT          | 3             | TF             | 92            |
| HTTP          | 3   | nDCG         | 136           | TF-IDF         | 91, 92, 116   |
|               |     |              |               | Transformer    | 202, 212, 216 |
| <b>【I】</b>    |     | <b>【O】</b>   |               | <b>【U】</b>     |               |
| IaaS          | 31  | OAuth        | 29            | UD             | 84            |
| IDF           | 92  | One-hot ベクトル | 91            | UDP            | 4             |
| IDR           | 107 |              |               | Unicode        | 74            |
| IOB2 形式       | 81  |              |               | Unicode 正規化    | 75            |
| IP            | 2   |              |               | UniDic         | 105           |
| IPA 辞書        | 104 |              |               | URI            | 8             |
| IR            | 109 |              |               | URL            | 6, 8          |
| <b>【J】</b>    |     | <b>【P】</b>   |               | <b>【w】</b>     |               |
| JavaScript    | 19  | PaaS         | 31            | WAN            | 2             |
| JMED-DICT     | 106 | PageRank     | 119           | Wikidata       | 94            |
| JSON          | 22  | Pairwise 法   | 126           | WordNet        | 93            |
| JUMAN 辞書      | 105 | Pointwise 法  | 126           | Word2Vec       | 90            |
| JUMAN/JUMAN++ | 80  | POST メソッド    | 14            | World Wide Web | 1             |
|               |     | PUT メソッド     | 15            | WSD            | 87            |
| <b>【K】</b>    |     | <b>【Q】</b>   |               | <b>【x】</b>     |               |
| K-means       | 64  | QUIC         | 16            | XML            | 21            |
| KPI           | 181 |              |               | XMLHttpRequest | 22            |
| <b>【L】</b>    |     | <b>【R】</b>   |               | ~~~~~          |               |
| LAN           | 2   | RAG          | 231, 236, 238 | 2 クラス分類        | 40            |
| Listwise 法    | 127 | REST         | 27            | 3 ウェイハンドシェイク   | 4             |
| LoRA          | 220 | RESTful API  | 29            | 3 層アーキテクチャ     | 7             |
| LSTM          | 195 | RMSE         | 61, 177       |                |               |
|               |     | RNN          | 194           |                |               |
|               |     | ROC 曲線       | 55            |                |               |
|               |     | ROUGE        | 100           |                |               |
|               |     | <b>【S】</b>   |               |                |               |
|               |     | Seq2Seq      | 198           |                |               |

—— 編著者・著者略歴 ——

**牛尼 剛聡**（うしあま たけとし）

1994年 名古屋大学工学部情報工学科卒業  
1996年 名古屋大学大学院工学研究科博士課程  
前期課程修了（情報工学専攻）  
1999年 名古屋大学大学院工学研究科博士課程  
後期課程満了（情報工学専攻）  
1999年 九州芸術工科大学助手  
2002年 博士（工学）名古屋大学  
2003年 九州大学助手  
2007年 九州大学助教  
2011年 九州大学准教授  
2024年 九州大学教授  
現在に至る

**佃 洸摂**（つくだ こうせつ）

2010年 京都大学工学部情報学科卒業  
2011年 京都大学大学院情報学研究科修士課程  
修了（社会情報学専攻）  
2014年 京都大学大学院情報学研究科博士後期  
課程修了（社会情報学専攻）  
博士（情報学）  
2014年 日本学術振興会特別研究員（PD）  
2015年 産業技術総合研究所研究員  
2019年 産業技術総合研究所主任研究員  
現在に至る

**河合由起子**（かわい ゆきこ）

1997年 九州工業大学情報学部電子情報工学科  
卒業  
1998年 奈良先端科学技術大学院大学情報科学  
研究科博士前期課程修了（情報システ  
ム学専攻）  
2001年 奈良先端科学技術大学院大学情報科学  
研究科博士後期課程修了（情報システ  
ム学専攻）  
博士（工学）  
2001年 通信総合研究所（現 情報通信研究機構）  
2006年 京都産業大学講師  
2018年 京都産業大学教授  
2018年 大阪大学 CMC 特任教授（常勤）  
2025年 関西大学教授  
現在に至る

**若宮 翔子**（わかみや しょうこ）

2009年 兵庫県立大学環境人間学部環境人間学  
科卒業  
2011年 兵庫県立大学大学院環境人間学研究科  
博士前期課程修了（環境人間学専攻）  
2013年 兵庫県立大学大学院環境人間学研究科  
博士後期課程修了（環境人間学専攻）  
博士（環境人間学）  
2013年 日本学術振興会特別研究員（PD）  
2014年 京都産業大学研究員  
2015年 奈良先端科学技術大学院大学博士研究  
員  
2018年 奈良先端科学技術大学院大学特任助教  
2020年 奈良先端科学技術大学院大学准教授  
現在に至る

**中島 伸介**（なかじま しんすけ）

1995年 神戸大学工学部生産機械工学科卒業  
1997年 神戸大学大学院自然科学研究科博士前  
期課程修了（機械工学専攻）  
2004年 京都大学大学院情報学研究科博士後期  
課程修了（社会情報学専攻）  
博士（情報学）  
2004年 情報通信研究機構専攻研究員  
2005年 奈良先端科学技術大学院大学助手  
2007年 奈良先端科学技術大学院大学助教  
2008年 京都産業大学准教授  
2015年 京都産業大学教授  
現在に至る



# Web テクノロジー —自然言語処理・検索・推薦・大規模言語モデル—

Web Technologies

—Natural Language Processing, Information Retrieval, Recommender Systems,  
and Large Language Models—

© Ushiamo, Tsukuda, Kawai, Wakamiya, Nakajima 2026

2026 年 8 月 3 日 初版第 1 刷発行

★

検印省略

|       |                 |
|-------|-----------------|
| 編 著 者 | 牛 尼 剛 聡         |
|       | 佃 洸 撰           |
| 著 者   | 河 合 由 起 子       |
|       | 若 宮 翔 子         |
|       | 中 島 伸 介         |
| 発 行 者 | 株式会社 コ ロ ナ 社    |
|       | 代 表 者 牛 来 真 也   |
| 印 刷 所 | 三 美 印 刷 株 式 会 社 |
| 製 本 所 | 株式会社 グ リ ー ン    |

112-0011 東京都文京区千石 4-46-10

発 行 所 株式会社 コ ロ ナ 社  
CORONA PUBLISHING CO., LTD.

Tokyo Japan

振替 00140-8-14844・電話 (03) 3941-3131(代)

ホームページ <https://www.coronasha.co.jp>

ISBN 978-4-339-01383-2 C3355 Printed in Japan

(松岡) G



**JCOPY** <出版者著作権管理機構 委託出版物>

本書の無断複製は著作権法上での例外を除き禁じられています。複製される場合は、そのつど事前に、出版者著作権管理機構（電話 03-5244-5088, FAX 03-5244-5089, e-mail: info@jcopy.or.jp）の許諾を得てください。

本書のコピー、スキャン、デジタル化等の無断複製・転載は著作権法上での例外を除き禁じられています。購入者以外の第三者による本書の電子データ化及び電子書籍化は、いかなる場合も認めていません。落丁・乱丁はお取替えいたします。