

目 次

刊行のことば	…2
第2巻について	…3

第1章 ソフトウェア概説

荻野綱男 7

1. はじめに	7
2. ソフトウェアのタイプと入手法	8
ソフトウェアの入手と使用	8／ローカル、ネットワーク、クラウドコンピューティング
3. カナ漢字変換システム	11
ATOK	11／Google日本語入力
複数のカナ漢字変換システムを使う場合の注意点	14／より効果的にカナ漢字変換システムを使う方法
单語登録	15／单語登録の際は正しい品詞の指定を
4. 文書作成・編集	20
ワープロソフト	20／エディタ
書いた論文の長期保存	25／表作成
グラフ作成	27／プレゼンテーション
5. 日本語処理	29
テキストデータの処理	29／数値データの処理
汎用の処理	31／汎用の処理
6. ユーティリティソフト	34
ファイル	34／ファイルの圧縮・解凍
バッファアップ	36／ファイルのバッファアップ
文字コード変換	37／ソート
不要ファイルの削除	39／不要ファイルの削除
ファイルサイズ	41／ファイルサイズ
キーの入れ替え	42／キーの入れ替え
ユーティリティソフトの選び方・使い方	43／ユーティリティソフトの選び方・使い方
7. ウェブ関連	44
アンチウィルスソフト	44／ブラウザ
メールソフト	45／メールソフト
ホームページ作成ソフト	46／ホームページ作成ソフト
8. ソフトのアップデートとバージョンアップ	49
9. ソフトのユーザ登録	51
10. ソフトのカスタマイズ	52
11. ソフトの相性	54
12. おわりに	54

第2章 検索

広瀬英史 57

1. Windows付属の検索機能	57
Windows XP	58／Windows 7
2. 検索ソフトウェアを利用した検索	75
Grape File Search	75／JGREP
JGrepper	87

第3章 Word, Excel, テキストエディタで行う日本語データ処理の基礎

小西いずみ 95

1. Wordにおける文字列の検索・置換	96
スペースやタブなどの編集記号を表示する	97／もっとも基本的な文字列の検索・置換
「あいまい検索」の有用性	98／より高度な検索①—改行・タブなど「特殊文字」の検索
より高度な検索②—「ワイルドカード」の使用	99／100／書式情報を含む検索・置換
「文字種の変換」機能	101／一片仮名⇒平仮名、全角⇒半角などの置換
2. Excelにおける文字列の検索・置換とカウント	104
基本的な文字列の検索・置換	105／「ワイルドカード」を利用した検索
全角⇒半角、大文字⇒小文字の一括置換	106／Wordの併用、関数の利用
Wordの併用、関数の利用	106／107／入力データのバリエーションの確認のカウント
「フィルター」機能	108／関数を利用した検索とカウント
「データの集計—ピボットテーブル」の利用	109／110／111／112
3. テキストエディタを用いた日本語データ処理の基礎	115
エディタの環境設定	117／基本的な文字列の検索
表示など	118／文字列の置換機能の応用
用例をマーク、カウント	119／grep検索
用例を抽出する	120／複数のファイルにわたる検索
grep検索でのファイル指定	122／「正規表現」の使用①—より高度な検索
「正規表現」の使用②—前方・後方不一致検索	123／124／正規表現のいろいろ
4. おわりに	128

第4章 形態素解析

柏野和佳子 131

1. 形態素解析の必要性	131
2. ことばの単位	133
さまざまな単位…134／短めの単位の特徴…135／長めの単位の特徴…136／『現代日本語書き言葉均衡コーパス』で採用した単位…137	
3. 形態素解析処理の手法	137
規則に基づく手法…137／統計・確率に基づく手法…140	
4. 形態素解析ツール	141
JUMAN…145／茶筌…147／MeCab…156／「茶筌 + UniDic」と「MeCab + UniDic」用の「茶まめ」…158	
5. 形態素解析システム用の辞書	161
品詞分類…161／JUMANの辞書…166／茶筌の辞書…170／MeCabの辞書…172／UniDic(茶筌, Mecabの辞書)…174	
6. 形態素解析の活用のために	182

第5章 機械翻訳ソフト

杉浦滋子 187

1. はじめに	189
調査の概要…187／機械翻訳ソフトの性質…188／句読法と機械翻訳の精度…189／文字の表記…200／固有名詞…202／言語間の形態面での対応のずれ…203／言語間で形態面に対応する要素に見られる意味的なずれ…209	
2. 機械翻訳の実際	210
英語…210／韓国語…219／中国語…224	
3. 結び	230

第1章

荻野 綱男

ソフトウェア概説

1. はじめに

コンピュータを使うことは事実上ソフトウェア（ソフト）を使うことである。なぜならば、ソフトを入れ替えて、それによってさまざまな働きをさせることができコンピュータの本質だからである。

パソコンの場合も同様であり、どんなパソコンを使うかよりも、どんなソフトを使うかのほうがはるかに重要な問題になる。

ソフトには、いろいろな種類があるが、ここではアプリケーションソフトに限定して話を進めることにする。言いかえるとOSに関わる話はしないということである。OSとアプリケーションソフトの関係については第1巻第1章に記述してあるが、パソコンと一体になって動くのがOSであり、OS上で具体的なさまざまな処理を行うのがアプリケーションソフトである。

日本語学を中心とする言語研究者がパソコンを使う場合、どんな使い方をするかというと、大きく3つの使い方がある。1つは論文を書くときに使うということであり、ワープロソフトが典型である。2番目には、日本語に関するデータを処理するときに使うということであり、表計算ソフトや統計ソフトなどが典型である。そして3番目には各種調査のためにインターネットを使うということで、ブラウザソフトが典型である。

ワープロソフトが扱う対象は「文書」であり、表計算ソフトや統計ソフトが扱う対象は「データ」である。その意味で両者は性格が違うのだが、パソコンの中では文書もデータもファイルの中に格納されているので、ファイル