平成 29 年度
産業日本語研究会 報告書

「産業日本語」
Technical Japanese

平成 30 年 3 月

一般財団法人日本特許情報機構
特　許　情　報　研　究　所

序

市場のグローバル化を背景として日本企業の海外展開が進むのに伴い、特許文書を初めとした海外向けの産業・技術文書作成の機会が年々増加しております。欧米諸国はもとより、近年では東南アジア諸国等への展開に伴い、低コストで正確な多言語翻訳が、これまでにも増して強く求められているところです。

また、自然言語処理の発展とともに、特許情報の機械翻訳や検索など、産業・技術情報の利活用が効率化・高度化されてきていますが、より正確な出力を可能とするために、文書データの品質が重要視されております。

このような背景のもと、高度な文書処理を効率良く低コストで実施するためには日本語の改善からのアプローチも重要と考え、当財団では平成 19 年度から、産業・技術情報を人に理解しやすく、かつ、コンピュータ（機械）にも処理しやすく表現するための日本語（「産業日本語」と呼称します）について、知的財産、特許翻訳、情報工学等の専門家の皆様によるご協力を得て、検討を進めております。また、毎年、産業日本語研究会・シンポジウムを開催するなど、産業日本語にかかる日本語研究とその研究成果の普及を推進しております。そのほか、平成 25 年度に発行した「特許ライティングマニュアル」の改訂に向けて、これまでの産業日本語活動の研究成果を踏まえつつ検討を進めてきたところですが、今般、改訂版を本報告書と同時期に発行する運びとなりました。

本報告書は、平成 29 年度の活動成果をご紹介するものです。平成 29 年度は、前年度に引き続き、ライティング分科会、文書作成支援分科会、特許文書分科会のそれぞれの観点から検討を進めたほか、産業日本語研究会・シンポジウムを開催しました。

ライティング分科会では、主にビジネス文書を対象として、「伝わる文書」を作成するための日本的・欧米的アプローチを検討しました。

文書作成支援分科会では、情報を理解しやすく提示するために、仮想特許明細書を題材に、文書を構造化した図式表現を検討しました。

特許文書分科会では、特許文書の品質を可視化するための品質評価方法策定に向けた定性的・定量的評価方法を検討しました。

本報告書が、広く皆様のお役に立つことを期待しています。産業日本語研究会の成果が広く利用されることで、特許文書をはじめとする日本語文書の品質向上が図られ、機械翻訳の訳質向上や人工知能技術による文書処理の効率化・高度化などにつながり、グローバルな企業活動の一助となれば幸いです。

今後とも、当財団の産業日本語に関する活動に、ご支援、ご協力をお願い申し上げます。

平成 30 年 3 月

一般財団法人　日本特許情報機構
専務理事・特許情報研究所所長　小林　明

目　次

Ⅰ　産業日本語研究会について……5
1. はじめに ……7
2. 平成 29 年度の研究会活動の概要……7
3. 研究会の検討経過 ……8
4. 平成 29 年度　産業日本語研究会　世話人・委員名簿…… 17

Ⅱ　ライティング分科会報告書…… 19
1. 平成 29 年度　ライティング分科会　委員名簿…… 21
2. 分科会概要 …… 21
3. 検討内容 …… 23

Ⅲ　文書作成支援分科会報告書…… 31
1. 平成 29 年度　文書作成支援分科会　委員名簿…… 33
2. 検討内容 …… 33

Ⅳ　特許文書分科会報告書…… 41
1. 平成 29 年度　特許文書分科会　委員名簿…… 43
2. はじめに …… 44
3. 他の対象の品質の調査 …… 45
4. 特許文書の品質について …… 53
5. おわりに …… 67

Ⅴ　産業日本語研究会ワークショップ講演…… 69
1.「いかに情報をシェアするか
～オントロジー構築と共通語彙構築の経験から～」…… 71
2.「辞書記述の精緻化とシソーラス・コーパスの活用」…… 99
3.「ソフトウェア要求仕様書の品質
～ IEEE Std. 830-1998 より～」……139

I
産業日本語研究会について

Ⅰ　産業日本語研究会について

1. はじめに

1.1　用語の定義と目標

「産業日本語」とは、「産業・技術情報を人に理解しやすく、かつ、コンピュータ（機械）にも処理しやすく表現するための日本語」と定義している。Japioは、そのコンセプトを作りあげると共に、特許情報の専門機関として、産業日本語を特許情報へ応用することを中心に、平成19年度から研究を進めてきた（特許明細書等の特許情報への応用に関する活動を特に「特許版・産業日本語」と称してきた）。

機械翻訳や検索、情報管理をより高度化し、各種情報の利用性の高度化・効率化を図るためには、コンピュータによる言語処理技術を活用した高度な文書処理が不可欠である。他方、システム側の努力だけでは限界が明らかとなり、文書そのものの改善や、文書を記述する日本語の改善からのアプローチも重要である。そのため、「産業日本語」活動は、種々の言語処理技術を活用することによって、明瞭な日本語文の作成と高品質な翻訳文の低コストでの作成を目標としている。

2. 平成 29 年度の研究会活動の概要

2.1　検討体制

本年度は、昨年度に引き続き、産業日本語研究会世話人会のご指導のもと、ライティング分科会、文書作成支援分科会及び特許文書分科会でそれぞれ検討を進めたほか、産業日本語研究会ワークショップ及び産業日本語研究会・シンポジウムを開催した。

2.2　主な活動概要

今年度の主な活動概要は、以下のとおりである。

2.2.1　ライティング分科会

ライティング分科会（主査：佐野委員）を開催し、特許ライティングマニュアル[1]及び日本語マニュアルの会の活動成果（「日本人のための日本語マニュアル（暫定第１版）」[2]、以下、「日本語マニュアル」という。）を基に、言語の専門家以外にも（公共性の高い）産業日本語の書き方を（普遍財として）普及していくことを想定し、以下の論点について検討を進めた。

- 産業日本語ライティングの技術向上策
- 「日本人のための日本語マニュアル（暫定第１版）」の３章の検討と３章をベースとする「伝える日本語」ステージのライティング・マニュアル（試作版）の作成

[1] https://www.tech-jpn.jp/tokkyo-writing-manual/　(2018 年 2 月 28 日アクセス)
[2] http://ngc2068.tufs.ac.jp/nihongo/htdocs/　(2018 年 2 月 28 日アクセス)

2.2.2　文書作成支援分科会

文書作成支援分科会（主査：橋田委員）を開催し、昨年度に引き続き、人間による文書作成・読解の高精度化・効率化、及び、文書の検索・翻訳・要約の精度と文書データを用いた機械学習の質の向上を図るための文書の図式表現を議論した。特に、仮想特許明細書（特許請求の範囲、明細書及び図面）を例題として取りあげ、請求項、明細書、図面の対応関係を図式表現で記述することについて検討を進めた。その結果、図式表現の効果と今後の検討課題を確認した。

2.2.3　特許文書分科会

特許文書分科会（主査：谷川委員）を開催し、昨年に続き、「特許文書の品質とは」というテーマで議論を進めた。本年度は、「特許文書品質特性」「各特性の評価方法（定性的，定量的)」を中心に議論した。定性的、定量的な評価の観点について、検討を進めた結果、多数の評価項目を上げることができ、各品質特性から特許文書の評価が一定程度可能であると考えられた。今後、特許文書の評価をより精緻に実施することができるよう、その評価を必要とする場面を想定しつつ、より使いやすい観点となるようにさらなる考察を進める。

2.2.4　産業日本語研究会ウェブサイトの更新

本活動の内容や産業日本語研究会・シンポジウムの案内を、前年度に引き続き、産業日本語研究会ウェブサイトに掲載した[3]。

3.　研究会の検討経過

3.1　産業日本語研究会

産業日本語研究会は、産業日本語研究会・シンポジウムを 1 回、産業日本語研究会ワークショップを 1 回、ライティング分科会を 5 回、文書作成支援分科会を 5 回、特許文書分科会を 5 回、それぞれ開催した。

3.1.1　産業日本語研究会ワークショップ

日時：平成 29 年 12 月 19 日（火）14:00－16:30

プログラム：

(1) 開会挨拶

日本特許情報機構　小林　明

(2) 講演：「いかに情報をシェアするか

～オントロジー構築と共通語彙構築の経験から～」

国立情報学研究所　武田英明

[3] https://www.tech-jpn.jp/　(2018 年 2 月 28 日アクセス)

(3) 講演：「辞書記述の精緻化とシソーラス・コーパスの活用 」
国立国語研究所　柏野和佳子
(4) 講演：「ソフトウェア要求仕様書の品質～ IEEE Std. 830 -1998 より～ 」
東京女子大学　白銀純子
(5) 閉会挨拶
日本特許情報機構　守屋　敏道

3.1.2　ライティング分科会

3.1.2.1　第1回 ライティング分科会

日時：平成29年8月4日(金) 10:00～12:00
議題：
(1) 委員紹介
(2) ライティング分科会の活動計画（案）について
(3) 今年度の活動計画
(4) その他

3.1.2.2　第2回 ライティング分科会

日時：平成29年9月29日（金）16:00～18:00
議題：
(1) 「パラグラフ・ライティング・マニュアル（試案)」について
(2) 「具体事例に基づいたパラグラフライティング」の実践報告
(3) 次回議題について
(4) その他

3.1.2.3　第3回 ライティング分科会

日時：平成29年10月27日(金) 16:00～18:00
議題：
(1) 「具体事例に基づいたパラグラフライティング」の実践報告(続)
(2) 統括
(3) 次回議題について
(4) その他

3.1.2.4　第4回 ライティング分科会

日時：平成29年11月17日（金）16:00～18:00
議題：
(1) 「パラグラフライティング」のまとめ
(2) 次回議題について

3.1.2.5　第5回 ライティング分科会

日時：平成30年2月9日(金) 14:00～16:00

議題：

(1) 「パラグラフライティング」のまとめ

(2) ポスター案について

(3) 報告書案について

3.1.3　文書作成支援分科会

3.1.3.1　第1回 文書作成支援分科会

日時：平成29年8月3日(木) 14:00～16:00

議題：

(1) 委員紹介

(2) 産業日本語研究会の活動計画(案) について

(3) 今までの研究経過について

(4) 今後の活動計画

(5) その他

3.1.3.2　第2回 文書作成支援分科会

日時：平成29年 10月 25日 (水) 16:00～18:00

議題：

(1) 課題背景について(主査)

(2) 特許明細書の構造化の実行について(各委員)

(3) 用語整理の検討(資料紹介)

(4) 次回の検討内容について

(5) その他

3.1.3.3　第3回 文書作成支援分科会

日時：平成29年12月 14日 (木) 16:00～18:00

議題：

(1) 課題背景について(主査)

(2) 特許明細書の構造化の実行について(各委員)

(3) 用語整理の検討(資料紹介)

(4) 次回の検討内容について

(5) その他

3.1.3.4　第4回 文書作成支援分科会

日時：平成30年1月17日 (水) 16:00～18:00

議題：

(1) 課題報告：
　① 「解決するための手段」と「実施形態」との対応関係に関する、Concept Mapping 表示
　② 「リンクの端点が実際にはどこなのか」に関する検討
(2) 課題報告：テキストと図面との対応関係の表示方法、動的インタフェースの検討(各委員)
(3) シンポジウムのパネル内容について
(4) 次回の検討内容について

3.1.3.5　第 5 回 文書作成支援分科会

日時：平成 30 年 2 月 14 日（水）14:00～16:00
議題：
(1) 課題報告：「テキストの構造化の検討--図との対応関係の検討も含む」
(2) シンポジウムのパネル内容について
(3) 報告書について

3.1.4　特許文書分科会

3.1.4.1　第 1 回 特許文書分科会

日時：平成 29 年 8 月 2 日（水）16:00～18:00
議題：
(1) 委員紹介
(2) 産業日本語研究会の活動計画(案）について
(3) 今までの研究経過について
(4) 今年度の活動計画
(5) その他

3.1.4.2　第 2 回 特許文書分科会

日時：平成 29 年 10 月 4 日（水）15:30～17:30
議題：
(1) 他分野の評価対象の調査と紹介
(2) 評価特性の議論（特にパラメータ評価特性）
(3) 次回の検討事項について

3.1.4.3　第 3 回 特許文書分科会

日時：平成 29 年 11 月 15 日（水）16:00～18:00
議題：
(1) 委員分担による評価特性の議論
(2) 次回委員会の検討事項について

3.1.4.4　第 4 回 特許文書分科会

日時：平成 29 年 12 月 13 日（木）16:00～18:00

議題：

（1）委員分担による評価特性の議論

（2）次回委員会の検討事項について

3.1.4.5　第 5 回 特許文書分科会

日時：平成 30 年 2 月 7 日（水）16:00～18:00

議題：

（1）品質評価特性表の検討（続き）

（2）ポスターの検討

（3）報告書の作成について

3.2　平成 28 年度以前の検討経緯

平成28年度までの、「産業日本語」活動の検討経緯を以下に列記する。

＜平成 20 年度＞

平成19年度の検討結果を踏まえ、以下に掲げる仕様を策定した。

- 産業日本語共通基盤仕様＜第1版＞（多種多様な産業技術文書毎に策定される仕様に対し、共通の枠組みを規定）
- 特許版産業日本語＜第0版＞（特許法等の関連法規に定められた記載要件を適切に満し、人と機械処理の双方にとって明晰な特許関連文書を書くための産業日本語仕様）
- 日英機械翻訳産業日本語＜第0版＞（実用化されている日英機械翻訳システムが適切な英語訳文を自動生成できる日本語原文を書くための産業日本語仕様）
- 検索産業日本語＜基本仕様版＞（文の検索をベースとする次世代の文書検索システムに対応するための産業日本語仕様）
- 図式産業日本語＜基本仕様版＞（グラフィカルな表現機構を導入することによって、表現構造を明示化し情報伝達機能を強化する産業日本語仕様）

＜平成 21 年度＞

これまでの検討結果及び策定された仕様をもとに、特許明細書作成実務により密着した検討を行い、以下を取りまとめた。

- 特許明細書ライティングマニュアル第0版＜準備編＞（和文特許明細書を、日英機械翻訳用の翻訳原稿に書き換えるために必要な検討事項を洗い出し、書き換え規則を体系的に整理する下地）

- 日英機械翻訳産業日本語＜第0.1版＞（一般の技術文書でも広く共通に使われる表現に対応させて、特許明細書から抽出した文と、特許明細書に頻出する文とからなる試験文100文を選んで翻訳実験を行い、仕様の改定・改良を試行）
- 図式産業日本語・検索産業日本語（特許オントロジーに基づく特許文書の図式表現と、特許検索等での活用可能性に関する検討）

＜平成 22 年度＞

特許オントロジー検討小委員会と特許ライティング・タスクフォースを設置し、以下の設計・作成を実施した。

- 特許オントロジーの設計（対象範囲の拡大と特許分類(IPC、FI、Fターム)の扱い、及び、特許オントロジーの利用法の検討: 入力支援、検索、読解支援、翻訳での利用、について検討）
- 特許ライティングマニュアル（第0.5版）（発明者等が発明を日本語として明晰に表現することができるようにするための、また、現状の機械翻訳をできるだけ活用して特許明細書を効率よく英文化できるようにするためのマニュアル）

＜平成 23 年度＞

特許オントロジー検討小委員会と特許ライティング・タスクフォースの２つで実施した。いずれの検討体制も、平成22年度の体制を原則継続した。

活動内容は、特許版・産業日本語の全体で中心的なテーマである特許文書処理の高度化、体系化に向けて、過去の成果を整理するともに、さらに検討の深化を目指した。

- 特許オントロジー検討小委員会では、昨年度検討した請求項の構造化のためのオントロジーと特許処理への応用のためのオントロジーの利用をより具体的に展開するようにした。検討課題は、オントロジーの設計、特許分類体系とオントロジーの融合化、オントロジーの活用の3項目とした。
- 特許ライティング・タスクフォースは、仕様やマニュアルに関するこれまでの蓄積や前年度行った特許文章の分析作業を踏まえて、それらを整理することを中心とした。

＜平成 24 年度＞

特許版・産業日本語全体の方向付けを行うため、「作業調整グループ」を設置し、特許版・産業日本語ワークショップを開催した。また、新たに、特許実務者が利用出来る特許ライティングの支援環境を調査するために、支援環境の提供者と実務者(弁理士)が参加する特許ライティング支援システムに関するタスクを設置し、アドホックミーティングを開催した。これらの活動などにより、以下の成果が得られた。

- 特許版・産業日本語の言い換え規則として蓄積してきた事例を類別して、実務者に利用できるマニュアルの基礎とした。
- 図式クレームの表現により、特許出願書類で中核となる請求項文のライティングを

構造化レベルとオントロジー化レベルで記述する方法を示した。

- 特許ライティング支援システムタスクは、特許ライティングに係る IT 系の支援システム開発者とそれを利用する特許実務者（弁理士）をメンバーとして意見交換を行うことで、開発者は、特許文書の作成、診断、閲覧で試用できる 3 つの環境を提供し、特許実務現場の評価を行い、以下のような成果が得られた。
  - ・特許実務現場における有効性が確認できた。
  - ・実務者からの意見をもとに、開発者にとっての改良項目が明確になってきた。
  - ・利用場面を想定したシナリオも提示され、両者の協調的な開発も期待できた。

＜平成 25 年度＞

特許ライティングマニュアルのまとめを行うとともに、特許ライティング支援システム（PWSS）グループにおいて、特許ライティング支援システムの利用に関する検討を行った。また、36条ルール化検討グループ会議において、特許法第36条に関するルール化可能性の検討を行った。さらに「特許版・産業日本語ワークショップ」を開催し、各グループの検討経緯を踏まえ、特許版・産業日本語委員会にて議論を行った。これらの活動により、主に以下の成果が得られた。

- 平成24年度までの成果を踏まえ、特許文章の明晰化のための「言い換え規則」をルールとしてまとめた「特許ライティングマニュアル（初版）」を発行した。
- 産業日本語活動での研究に用いるための特許明細書サンプルとして、仮想特許明細書を独自に作成し、人手翻訳による英語への翻訳を行った。
- 特許ライティング支援システムの有効性を広く認識してもらうために、仮想特許明細書を用いて特許ライティング支援の活用プロセスの可視化を行った。
- 特許法第36条の各項目について、違反類型に該当するかを機械的にチェックするための条件を求めることが可能か否かをまとめた。

＜平成 26 年度＞

特許ライティングマニュアルの改訂を行うとともに、利用許諾なしで利用できるように仮想特許明細書の作成を行った。また、「構造化クレームを用いる請求項文ライティングマニュアル」の作成を行った。

36条ルール化検討グループ会議においては、拒絶理由通知書の収集・分析を行うとともに、特許法第36条第6項第2号および特許法第29条第1項柱書に対する違反の類型化の検討を行った。さらに「特許版・産業日本語ワークショップ」を開催し、各グループの検討経緯を踏まえ、特許版・産業日本語委員会委員による議論を行った。これらの活動により、主に以下の成果が得られた。

- 36 条ルール化検討グループ会議の検討

  表題会議を 4 回開催し、拒絶理由通知書の収集・分析を行うとともに、
  - ・明確性要件（特許法第 36 条第 6 項第 2 号）違反の類型化
  - ・産業上利用可能性（特許法第 29 条第 1 項柱書）違反の類型化

について検討を進めた。

- 特許ライティングマニュアルの普及及び改訂
  平成 25 年度に発行した「特許ライティングマニュアル（初版）」の普及活動を行った。また、初版から第 2 版への改訂に向けた検討を行った。
- 仮想特許明細書の作成
  利用許諾などの制限を受けずに調査・分析が行える特許明細書のサンプルとして、昨年度に引き続き、化学・機械・物理の各分野から 3 本の仮想特許明細書を作成した。
- 「構造化クレームを用いる請求項文ライティングマニュアル」の作成
  構造化クレームを用いる請求項文ライティングマニュアルについて、「構造化クレームを用いる請求項文ライティングマニュアル（第１版）」を取りまとめた。
- 木構造形式によるライティング支援の検討
  医療分野におけるカルテ等での活用を試みているグラフ形式に基づくライティング手法を特許分野への応用を試みるための検討を行った。

＜平成 27 年度＞

平成 26 年度の体制を一部引き継ぎ、「36 条ルール化検討グループ会議」を設置した。

また、特許版・産業日本語委員会による「平成 27 年度特許版・産業日本語ワークショップ」を開催した。これらの活動により、主に以下の成果が得られた。

- 特許法第36条違反の審査書類の収集・分析
  36 条ルール化検討グループ会議を開催し、新しい拒絶理由通知書の収集・分析を行うとともに、以下について検討を進めた。
  ・新しい拒絶理由通知書の収集
  ・段落解析技術の特許文書への応用可能性検討
  ・致命的な36条違反の分析
  ・36条違反の検出方法に基づく分類と特許審査基準との対応
- 特許ライティングマニュアルの改訂検討
  第 2 版発行に向けて、初版の改訂に向けた検討を行った。特許明細書へのルールなどの実例を付与するなど、特許ライティングにも活用できるような具体的な改訂内容の検討を進めた。
- 特許版・産業日本語ウェブサイトのリニューアル
  本活動の内容をウェブサイトからも理解しやすい形にするために、特許版・産業日本語ウェブサイトをリニューアルした。

＜平成 28 年度＞

産業日本語研究会世話人会のご指導のもと、産業日本語研究会に特許版・産業日本語委員会を統合し、ライティング分科会、文書作成支援分科会及び特許文書分科会を設置した。また、産業日本語研究会ワークショップ及び産業日本語研究会・シンポジウムを開催した。これらの活動により、主に以下の検討を行った.

● 産業日本語ライティングの技術向上の検討

ライティング分科会（主査：佐野委員）を開催し、「日本語マニュアルの会」の活動成果（「日本人のための日本語マニュアル（暫定第１版）」[4]、以下「日本語マニュアル」という。）を基に、言語の専門家以外にも公共性の高い産業日本語の書き方（特に、「伝える日本語」）を普及させることを想定し、以下の論点について検討を進めた。

・本分科会の方針（「日本語マニュアル」の内容検討とライティング規則充実）の確認
・「日本語マニュアル」の4章のライティング規則の扱いの検討
・「日本語マニュアル」の2章の「表す日本語」およびパラグラフの考え方の検討
・「日本語マニュアル」の3章の「『伝える日本語』への言い換え規則」の検討

● 産業日本語文書の構造化表現や図式化の検討

社会全体の業務改革（BPR）に不可欠なデータの連携・標準化を可能にする仕組みを議論した。特にAIに文書を理解させる手段としての構造化文書および図式表現（ダイアグラム）について議論し、それを支える用語管理方法について検討した。

有効性を検証するため、特許文書（特許請求の範囲および明細書）を例題として取りあげ、検討を行った。

● 特許文書の品質特性と評価シチュエーションの関係の検討

「特許文書の品質とは」というテーマで議論を進めた。既に確立されている「システム開発文書」の品質特性、および「ソフトウェアの品質モデル」を参考にしながら、「特許文書品質特性」として、３階層で構成される品質特性および評価シチュエーションのマトリクス表を作成し、各要素の検討を行った。

[4] http://ngc2068.tufs.ac.jp/nihongo/htdocs/ （2018年2月28日アクセス）

4. 平成 29 年度　産業日本語研究会　世話人・委員名簿

4.1　世話人名簿

（五十音順・敬称略）

	氏　名	所　属
代表世話人	井佐原　均	豊橋技術科学大学　情報メディア基盤センター センター長・教授
世話人	潮田　　明	国立研究開発法人産業技術総合研究所　人工知能研究センター
世話人	隅田英一郎	国立研究開発法人情報通信研究機構 先進的音声翻訳研究開発推進センター 副研究開発推進センター長
世話人	辻井　潤一	国立研究開発法人産業技術総合研究所　人工知能研究センター 研究センター長／東京大学 名誉教授
顧　問	長尾　　眞	公益財団法人国際高等研究所 所長／京都大学名誉教授
世話人	橋田　浩一	東京大学大学院　情報理工学系研究科 ソーシャル ICT 研究センター　教授
世話人	柏野和佳子	大学共同利用機関法人人間文化研究機構　国立国語研究所 音声言語研究領域　准教授
世話人	横井　巨人	一般財団法人日本特許情報機構　特許情報研究所 調査研究部長

4.2　委員名簿

（五十音順・敬称略）

	氏　名	所　属
委　員	安彦　　元	ミノル国際特許事務所　所長・弁理士
委　員	石井　岳史	旭化成株式会社　研究・開発本部　知的財産部 知財リエゾン第二グループ　課長
委　員	石崎　　俊	慶應義塾大学環境情報学部　名誉教授
委　員	猪野真理枝	東京外国語大学オープンアカデミー講師／ 翻訳家／語学教材制作家
委　員	岩永　勇二	平田国際特許事務所　パートナー　弁理士
委　員	岩山　　真	株式会社日立製作所　研究開発グループ システムイノベーションセンタ 知能情報研究部　主任研究員
委　員	烏　　日哲	国立国語研究所　研究系・日本語教育研究領域　プロジェクト 非常勤研究員／一橋大学国際教育センター非常勤講師
委　員	江原　暉将	元・山梨英和大学　教授

委　員	黒川　　恵	日本弁理士会／阿部・井窪・片山法律事務所　弁理士
委　員	久保田真司	パナソニックＩＰマネジメント株式会社 グローバル出願部品質強化担当参事
委　員	熊野　　明	東芝デジタルソリューションズ株式会社 ICT インフラサービスセンター　ソフトウェア開発部
委　員	栗田　朋香	テルモ株式会社　知的財産部　主任
委　員	古賀　勝夫	株式会社クロスランゲージ　相談役
委　員	佐野　　洋	東京外国語大学大学院　総合国際学研究院　教授
委　員	新森　昭宏	株式会社インテック　生産技術部　理事
委　員	武田　英明	大学共同利用機関法人情報・システム研究機構 国立情報学研究所　情報学プリンシプル研究系　教授
委　員	田中　成彦	一般財団法人工業所有権協力センター 調査業務センター　調整部／情報システム部　次長
委　員	谷川　英和	IRD 国際特許事務所　所長・弁理士
委　員	冨田　修一	株式会社知財コーポレーション 代表取締役社長
委　員	橋田　浩一	東京大学大学院　情報理工学系研究科 ソーシャル ICT 研究センター　教授
委　員	藤田　卓仙	東京大学大学院　医学系研究科医療品質評価学講座 特任研究員
委　員	間瀬　久雄	株式会社日立製作所　研究開発グループ システムイノベーションセンタ　知能情報研究部　主任研究員
委　員	的場　成夫	有限会社夢屋　代表取締役　弁理士
委　員	山本ゆうじ	秋桜舎　代表
委　員	横井　巨人	一般財団法人日本特許情報機構　特許情報研究所 調査研究部長
委　員	白土　博之	一般財団法人日本特許情報機構　特許情報研究所 研究企画課長
委　員 （事務局）	荻野　孝野	株式会社日本システムアプリケーション　言語処理グループ 主任研究員
オブザーバ	小林　　明	一般財団法人日本特許情報機構　専務理事／ 特許情報研究所　所長
事務局	三橋　朋晴	一般財団法人日本特許情報機構　特許情報研究所 研究管理課長
事務局	塙　　金治	一般財団法人日本特許情報機構　特許情報研究所　研究管理部
事務局	三吉　秀夫	株式会社日本システムアプリケーション

Ⅱ
ライティング分科会報告書

Ⅱ　ライティング分科会報告書

1. 平成 29 年度　ライティング分科会　委員名簿

（五十音順・敬称略）

	氏　名	所　属
委　員	石崎　　俊	慶應義塾大学環境情報学部　名誉教授
委　員	猪野真理枝	東京外国語大学オープンアカデミー講師／ 翻訳家／語学教材制作家
委　員	烏　　日哲	国立国語研究所　研究系・日本語教育研究領域　プロジェクト 非常勤研究員/一橋大学国際教育センター非常勤講師
委　員	熊野　　明	東芝デジタルソリューションズ株式会社 ICT インフラサービスセンター ソフトウェア開発部
委　員	古賀　勝夫	株式会社クロスランゲージ　相談役
委　員 （主　査）	佐野　　洋	東京外国語大学大学院　総合国際学研究院　教授
委　員	冨田　修一	株式会社知財コーポレーション 代表取締役社長
委　員	間瀬　久雄	株式会社日立製作所　研究開発グループ システムイノベーションセンタ　知能情報研究部　主任研究員
委　員	横井　巨人	一般財団法人日本特許情報機構　特許情報研究所 調査研究部長
委　員	白土　博之	一般財団法人日本特許情報機構　特許情報研究所 研究企画課長
委　員	荻野　孝野	株式会社日本システムアプリケーション 言語処理グループ　主任研究員

2. 分科会概要

2.1　活動概要

- 産業日本語ライティングの技術向上について検討する
- 特許ライティングマニュアル及び日本語マニュアルの会の活動成果を引継いで，言語の専門家以外にも（公共性の高い）産業日本語の書き方を（普遍財として）普及していくことを想定している
- 「日本人のための日本語マニュアル（暫定第 1 版）」の 3 章の検討と 3 章をベースとする「伝える日本語」ステージのライティング・マニュアル（試作版）を作成する

■分科会実施状況

	開催日時	検討内容
第一回	平成 29 年 8 月 4 日（金）（10 時～12 時）	今年度の活動計画（「日本人のための日本語マニュアル（暫定第 1 版）」の 3 章の規則について検討すること）を確認し，同マニュアルの第 3 章の内容確認を行った。
第二回	平成 29 年 9 月 29 日（金）（16 時～18 時）	日本語パラグラフについて具体事例をもとに分析（各委員が実施，本分科会で報告する）した。4 人の委員が報告を行い，報告内容について検討した。
第三回	平成 29 年 10 月 27 日（金）（16 時～18 時）	日本語パラグラフについて具体事例をもとに分析（各委員が実施，本分科会で報告する）した。3 人の委員が報告を行い，報告内容について検討した。
第四回	平成 29 年 11 月 17 日（金）（16 時～18 時）	パラグラフ・ライティングに関するいくつかの概念・観念を整理した。文書（ドキュメント）と文章（テキスト），文章と論理的な思考，支持の方法論（合理的な表明（態度），粗筋・論調の立て方），文章と段（パラグラフ）の区別など概念範囲を議論した。
第五回	平成 30 年 2 月 9 日（金）（14 時～16 時）	「伝える日本語」ステージのパラグラフ・ライティングについて検討し，文書と文章の区分を行い，合理的な表現態度と文章構造の関係，文章表現を支える筋書きの特徴について整理した。

■「日本人（ビジネスマン）のための日本語（ビジネス文書）マニュアル（暫定第 1 版）」（ [1]）

ダウンロードサイトのスナップショットを図 1 に示す。

URL ： http://ngc2068.tufs.ac.jp/nihongo/htdocs/

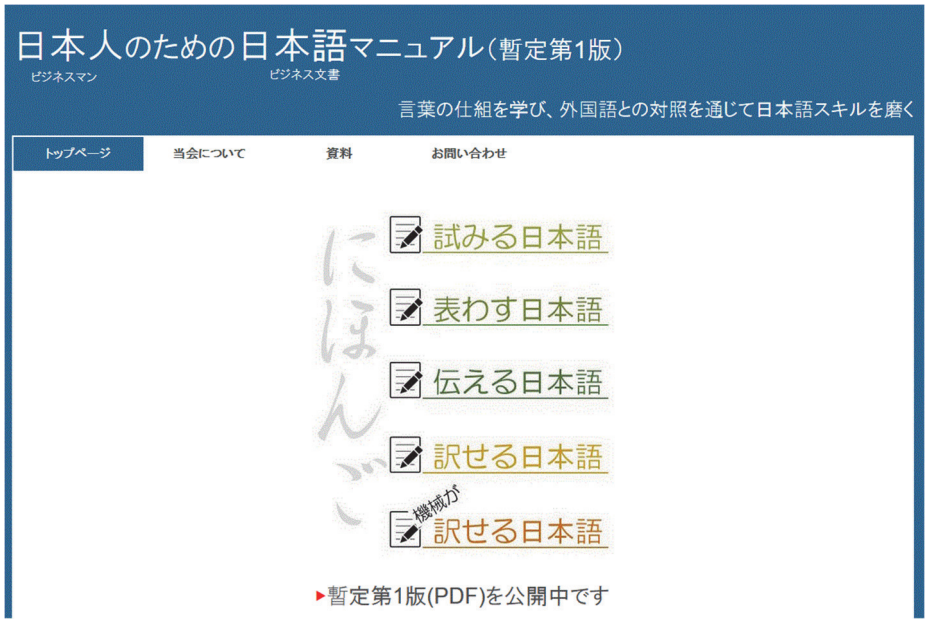


図 1 ダウンロードサイト

3. 検討内容

3.1　「日本人のための日本語マニュアル」第 3 章の内容検討（第一回）

本分科会の検討方針を確認した。まず，「日本人のための日本語マニュアル」（第 3 章）を下地として「伝える日本語」段階のパラグラフ・ライティングについて検討すること，そして，この検討内容と 3 章の内容を参照して「伝える日本語」段階のビジネス・ライティングマニュアル（試作版）の作成することの 2 点である。

本方針に対する意見として，第 3 章の内容表現は難解であること，全体を俯瞰する必要があること，パラグラフ・ライティング規則のトップダウン的な記述スタイルから，具体事例を示してからパラグラフ・ライティング規則を解説したり説明したりする表現に変えたほうがよいことなどが提案された。また，第 3 章を参照するとパラグラフの構造が不明瞭であることや，（曖昧になっている）欧米型パラグラフの構造と日本型パラフラフの構造を明確に区分する必要があることも議論となった。

議論のポイントを以下に示す。

- 主張の言語表現を使った支持方法をその表現も併せて具体化するべきだ
- 日本語の表現を用いた事実や意見の文型を提示する
- 英語表現では感覚や知覚表現と認識表現が語形的にも文型的にも違う
- 上記に対する日本語の語形表現や文型表現を具体的に示すべきだ
- 因果性（演繹推論や帰納推論）の説明は，理解が難しいこともあって避けたほうがよい

こうした議論から日本語表現を基本とするパラフラフライティングの方法論を視界に入れる必要性を認識した。具体的には，「だ，である」が is でないこと，「なる」が become で

ないこと，「する」が do でないことなど，表現を支える思考様式の違いを含めて説明することが求められる。

同時に，文章表現の方法論を企業内で広めることを考慮すると，簡潔性を軸としたポイントを取り出すことこと，技法上のポイントと考え方のポイントを分けることのなど工夫が必要であることも認識した。このことは，体系的に作文を学ぶ指南書も必要だが，ニーズ指向で，ピンポイント的に要求に応える指南書もまた必要であるということである。以下の点が議論された。

- 読み手を意識することのポイントを挙げる
- ライティングマニュアルの対象読者を具体化するべきだ
- 対象読者のニーズも特定する必要がある
- 書き方のニーズ（「深く詳しく」「広く浅く」「目的別」「一般的」など）を絞り込む
- （日本語母語話者としての）日本人の思考様式と表現方法を具体的に挙げる
- 体系だった説明と具体用例を使った説明の双方を挙げる

こうした議論を踏まえ，本分科会の作業として，パラグラフ単位の説明を既存の（実例の）文章を使って試みること，その中で「企画書」（未来の予測），「報告書」（過去の分析）を取り上げ，表現の多様性を確認することなどを確認した。作業仮説として，英語では（多分），主張があって説得し，対立によって幾多のポイントが明らかになり，その結果（いわゆる，議論の）「落としどころ」を見出す。一方，日本語では（恐らく），主張に対する納得を求め，沢山のポイントを共有し，その結果「落としどころ」を見出すと思われる。このように結果となる着地点へのアプローチが違う。アプローチが違うので論調や論調を担う形式が異なると考えられる。

別視点―「訳しやすい日本語はいい日本語なのか」という観点からも議論を行った。否定的な意見が出て，同時に，そのことを表現の具体的な方法論を含めて具体的に明らかにするべきであることを確認した。

教本としてのマニュアルは，その内容を簡単にすることで，読ませるよりも参照させるスタイルの教本にするべきとの意見があった。具体的には，直接書き方を指南するか（技能型，ダイレクトメソッド），日本語の特性や特徴を理解させて指南するか（座学型，納得メソッド）を文章量などの観点から検討するべきだろう。

3.2 パラグラフ・スタイルの具体的な検討（第二回）

各委員が文章表現の具体事例の分析報告を行った。

- 論文や技術報告書のパラグラフ・スタイルについての書き方の具体例を用いた説明
- 通知書の例と日本語教育の例を取り上げた説明
- ある技術雑誌の巻頭言の例の説明と，その巻頭言の一部の表現の詳細分析

（「日本人のための日本語マニュアル」（第 3 章）で提案される）道標文と前方支持文，話題文，後方支持文のすべてが，ひとつのパラフラフに整う必要がないことが議論された。以下に論点を示す。

- 筋書きと，起パラグラフの構造や承パラグラフの構造，結パラグラフの構造は違うので

はないか

- 筋書きと，パラグラフ構造の関係や筋道のステップと知識構造の関係は何か（情報構造とその連鎖，知識累積と概念構造の変化など形式化が可能か）
- 道標文だけのメタ・パラグラフを想定してよいのではないか

こうした議論から，話題文が必ずパラグラフには必要であるという説明も見直すことも必要であること，筋書の各ステージに応じたパラグラフ－例えば，話題導入や背景，概念定義などを表現するパラグラフ，説明する表現のパラグラフ，説得する表現のパラグラフなどの区分―があることを明確にした。

3.3　パラグラフ・スタイルの具体的な検討（第三回）

第二回に引き続き，各委員が文章表現の具体事例の分析報告を行った。以下に主な論点を挙げる。

- 文書，文章，節，小節といった文書形式構造と，パラグラフの関係（シナリオ，筋書き，アウトライン）がある
- 文書の話題（トピック）と文章の話題，ひとつのパラグラフの話題がある。各話題の観念を定めるべき（話題文の話題の限定範囲を明確にするため）
- 目的によってシナリオや筋書き（アウトライン）を選択するべき（「起 - 承 - 結」（共感戦略）という筋書きもあるし，「結 - 承 - 結」（説得戦略）という筋書きもある）
- 文書タイプを想定する（特許分野から拡張するとすれば，技術分野とか工業分野など。文章タイプは企画文書とか報告文書など）

3.4　パラグラフ・ライティングとその関係概念（第四回）

文書と文章の違いを具体的に検討した。

文書と文章は違うこと，すなわち，文書は形式と文章構造を持ち，文章はその伝達目的別で区分される。日本語スタイル（共感スタイル）のパラグラフに，中心となる文が必ず１つあるという主張は正しいかどうか検討することも確認した。この議論は，主題（文書全体の話題），話題（文章全体の話題），主旨と中心（ひとつのパラグラフの話題）の観念上の区別を進めることにつながる。

日本語表現を基礎とするパラグラフ・ライティングの必要性もその輪郭が明確になってきた。分科会の議論では，この視点に沿って，共感論調と意志論調の二つの合理的な態度の表明法があることが提案された。文書作成技法として，この二つを並列させて説明することがよいだろう。

3.5　「伝える日本語」ステージのパラグラフ・ライティング（第五回）

第 4 回の分科会では，パラグラフ・ライティングに関するいくつかの概念や観念を整理した。文書（ドキュメント）と文章（テキスト）の違い，文章とパラグラフの区別，論理的な思考とは何かと表現の仕方の関係，話題と支持の仕方や粗筋・論調の立て方などである。こうした議論を踏まえてひとつの案として「パラグラフ・ライティング書き方（簡易版）」

（図 2）と「パラグラフ・ライティング考え方（簡易版）」（図 3）をまとめた。特徴は以下である。

- 文書と文章を概念的に分け、書き手の態度を文書の書き方の指針に取り入れた
- 共感を通じて納得してもらうことと，意志を持って説得することを合理的な表明の方法とした
- 表明の道筋としてアウトラインを取り入れた（問題になった起・承・結は論じ方で，これをパラフラフの類型に使うことを止めた）
- パラグラフ（形式）とパラグラフ・ライティング（技能）を分けた
- 論じ方を類型化して，二つの表明の仕方（帰納的な推察で根拠から結論を導く方法と，演繹的な推察で根拠から結論を導く方法）に対応させた
- 話題文と支持文の役割を限定した。話題文は制限で，支持文は推論である。（パラグラフの意味も定義した）
- 支持文を使った表明は，表明の仕方の選択と論拠の担保の技法に関連させるようにした

二つの表明の仕方，すなわち思考の方法に深く関わる。この二つの区分は，必ずしも正確ではないものの直観的な理解のために，前者を日本的アプローチと言い，後者を欧米的アプローチと称してもよいだろう。以下（表 1）のように整理できる。なお、表現の特徴を表 2 に挙げる。

表 1 共感技法と意志技法

	日本的アプローチ（共感技法）	欧米的アプローチ（意志技法）
合理的な表明	共感を通じて納得してもらう	意志をもって説得する
合理的な心構え（共感と意志）	共感による納得は，共通項を探すことであり（調和），多視点で客観的な事実や主観的な信念を担保づけること	意志による説得は，他の選択肢を排除することであり（対立），単視点で客観的な事実や主観的な信念を裏付けること
思考の方法	結論からその根拠を推し量る，または根拠を積み上げて結論に至る	根拠から結論を導く，または結論を先行させ，根拠を後から理屈付ける
書き手の態度	・根拠（共有知識や一般事実）で正しさを主張する ・信念（理性的な情感や情意）で価値を主張する	
表明の際の留意（健全な思考）	・根拠（共有知識）を明確にして表現する ・公（おおやけ）の見解と私的意見を明確に区別する ・安易な判断と断定は思考を止めるので避ける	・根拠（一般事実）を明確にして表現する ・追認された出来事と経験を明確に区別する ・安易な判断と断定は思考を止めるので避ける

表 1から分かるように書き手の態度や表明の際の留意は変わらない。なお，パラグラフ・ライティングの考え方と書き方をまとめた資料も作成した。この資料（段作文指南）は，言語現象分析ではなく，（思考活動の）約束事としての「段（パラグラフ）」の規格化を通じて，ことばの専門家以外にも（公共性の高い）産業日本語の書き方を（普遍財として）普及していくことを目指す。同時に，本分科会は，基本枠組みの理解ならびに表現不備の点，追加や修正などの検討を行うとともに，目的別に，取りあげる項目を簡素化したり，説明手段を工夫したりするなどの点を追求し，成果とする予定である。

■資料１　パラグラフ・ライティング（書き方，簡易版）

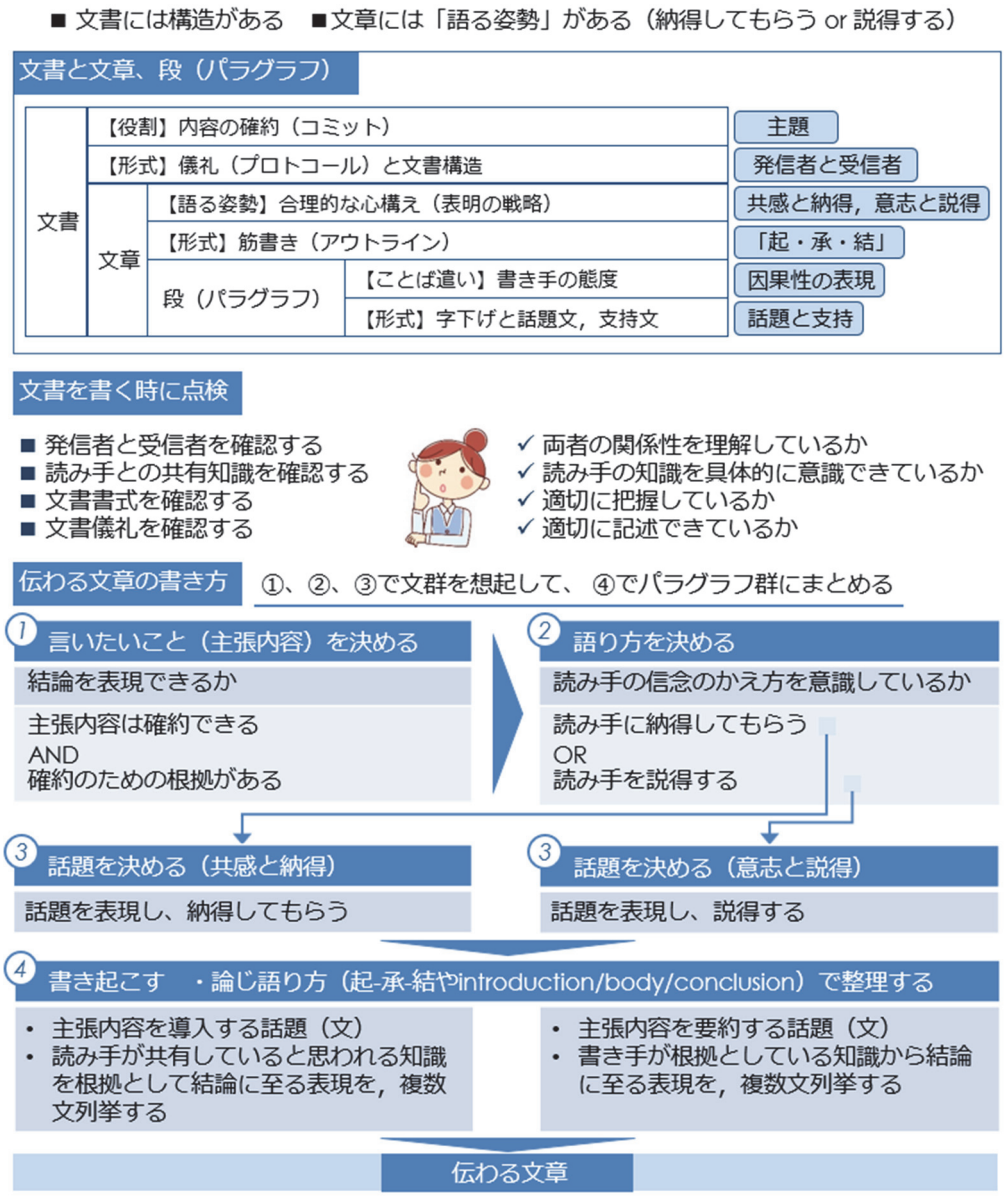


図 2 パラグラフ・ライティングの書き方（簡易版）

表 2 表現の特徴

	日本的アプローチ（共感技法）	欧米的アプローチ（意志技法）
合理的な表明	共感を通じて納得してもらう	意志をもって説得させる
表現（と目標）	「何々は（どんな）だ」，「何々が～に（どう）だ」，「何々が（どう）なる」 （事実等の根拠を積み上げていくことにより，前提を共有していくことで，判断のみでなく，それに至る過程までを含めて共有する。）	「何（々）が何（々）をどう動かす」，「誰（々）が（何（々）をどうする」 （判断を主張し，それを事実等の根拠を持って裏付けることで，反論の余地をできるだけ排除して，相手を説得する。）

■資料２　パラグラフ・ライティング（考え方，簡易版）

図 3 パラグラフ・ライティングの考え方（簡易版）

■資料３　パラグラフ・ライティング教本（案）

（http://ngc2068.tufs.ac.jp/nihongo/ を参照）

■資料４　ライティングプロセス（参考）

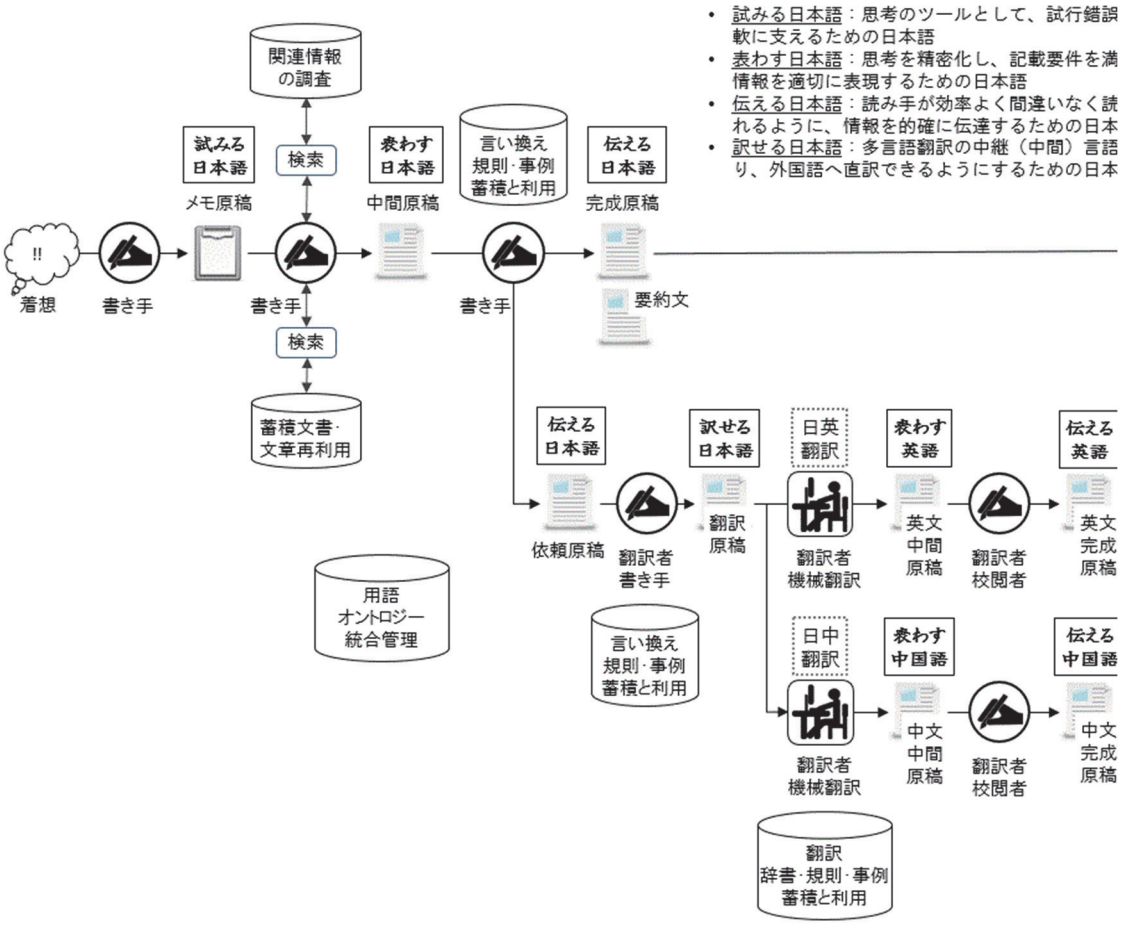


図 4 ライティングプロセス（参照資料から引用）

参照資料

[1] 日本語マニュアルの会, “日本人のための日本語マニュアル（暫定第１版）,” 8 2016. [オンライン]. Available: http://ngc2068.tufs.ac.jp/nihongo/htdocs/.

Ⅲ
文書作成支援分科会報告書

Ⅲ　文書作成支援分科会報告書

1. 平成 29 年度　文書作成支援分科会　委員名簿

（五十音順・敬称略）

	氏　名	所　属
委　員	岩山　　真	株式会社日立製作所　研究開発グループ システムイノベーションセンタ 知能情報研究部　主任研究員
委　員	江原　暉将	元・山梨英和大学　教授
委　員	新森　昭宏	株式会社インテック　生産技術部　理事
委　員	武田　英明	大学共同利用機関法人情報・システム研究機構 国立情報学研究所　情報学プリンシプル研究系　教授
委　員 （主　査）	橋田　浩一	東京大学大学院　情報理工学系研究科 ソーシャル ICT 研究センター　教授
委　員	藤田　卓仙	東京大学 大学院医学系研究科 医療品質評価学講座 特任研究員
委　員	山本ゆうじ	秋桜舎　代表
委　員	横井　巨人	一般財団法人日本特許情報機構　特許情報研究所 調査研究部長
委　員	白土　博之	一般財団法人日本特許情報機構　特許情報研究所 研究企画課長

2. 検討内容

2.1　検討の方針

たとえば「ロボットは東大に入れるか」プロジェクトで明らかになったように、現在の人工知能は複雑な問題文を理解できないため東大等には合格しないけれども、平均よりレベルが高い大学には十分合格し得る。つまり、平均的な大学受験生は文書読解能力に関して人工知能に劣る[新井・尾崎 17]。それは受験生が限られた試験時間の中で効率的に点数を稼ぐための受験テクニックを使っているからだという意見もあるが、社会人になっても業務の時間は限られているからやはり同様のテクニックを使うので文書読解能力は高まらないのではないか。また、米国の大学では 4 年間の在学中に批判的思考(critical thinking)等の能力がほとんど向上しないことが明らかになっている[Arum&Roksa 11]が、日本の大学教育の効果が米国のそれを大きく上回るとはあまり期待できない。社会人でも、契約書や論文や法律など、ある程度複雑な文書の作成・読解に日常的に携わりその能力を高める機会に恵まれているのはごく一部だから、大人の平均的な文書作成・読解能力が受験生や大学生のそれを大きく越えるとも思えない。

社会の運営は、契約や法律等の文書による情報共有と合意形成に基づく。したがって、文書を作成・読解する人々の能力が平均的に低いならば、その作成・読解を支援して情報共有と合意形成の精度と効率を高めることにより、社会全体の生産性を向上させられる余地が大きい。それがスマートソサエティを実現するための主な必要条件のひとつだろう。

しかし、たとえば複雑な文書の作成・読解に日常的に携わるのはほんの一握りの人々に過ぎず、文書の作成・読解に関する教育をその他の多くの人々に日常的に課するのは不可能だろう。それよりむしろ、音声言語に由来する線状性の制約ゆえにわかりにくく作りにくい従来形式の文書にとらわれず、文書において意味構造を作りやすくわかりやすいように明示することによって下記の実現を図る方がはるかに現実的かつ有益と考えられる。

- 人間が高い精度と効率で文書を作成・読解できるようにする
- 文書の検索や翻訳や要約の精度と、文書データを用いた機械学習の質を高める

文書を構成する言語表現には構造的曖昧性や語彙的曖昧性がある。そのため、表出および理解の両面で誤解が生じ、情報交換に障害を来たす。特に特許文書は、技術用語が多用され、文も長く複雑である。そのため、上記 2 つの曖昧性が顕在し、「分かりにくい」という問題が指摘されている。図式表現は 2 つの曖昧性をできるだけ少なくした情報交換手段を提供することを目的としている。ここで、情報交換は人間同士の間での情報交換に加えて人間と機械との情報交換も含む。

特許請求項は物事が抽象的に（上位概念化されて）記載されることが多く、一見しただけではそれぞれ何を意味するのか明確でない場合も多い。通常、明細書を参照することで、請求項の構成要素と明細書（実施例）中の対応関係を把握することになるが、そのためには全体（又はある程度の文章のかたまり）を読むことを要するため、時間を必要とする。

また、請求項だけでみても、内容把握が容易でないという問題がある。例えば独立請求項において記載された構成要素は、上述のように抽象的である一方で、下位請求項（引用請求項）では、当該構成要素が下位概念化された形で記載されることが一般的である。この時、上位概念で表された構成要素と下位概念で表された構成要素の関係（包含関係など）の把握に時間を要することがある。さらに、請求項/明細書と図面との対応関係の把握という問題もある。

図式表現では基本的にノードおよびノード間のリンクから構成されるグラフの形のデータ構造を明示するが、下記のようにさまざまな制約や拡張や変換が想定される。

- リンクは始点と終点の間の 2 項関係を表現する。始点が表わすものがその 2 項関係の第 1 項、終点が表わすものが 2 項関係の第 2 項である。
- ノードの内容はテキストとは限らず、映像(静止画や動画)でも良い。映像には時空間的な全体部分を含むが、それをグラフの代わりの図式表現と見なすこともできる。
- ノードがグラフを含むことがある。そのようなノードにより、量化や否定の作用域、発言や信念の内容などを表現できる。
- 部分グラフをノードに縮退させることにより、大きく複雑なグラフを簡単なグラフに変換し、複雑な意味構造の概要を簡単な意味構造として捉えることができる。

2.2　得られた知見

具体的な特許文書のいろいろな部分の図式表現を実際に分科会の各委員が作ってみたところ、その図式表現からもとのテキストの内容を読み取ることが概ねできたことから、任意のテキストの意味内容を図式表現が存在する可能性が高いと考えられる。それに加えて、図式表現の利用によって期待される以下のような効果を確認できた。

- 図式表現により、文書の部分の間の関係の視認性が高まり、内容把握に寄与する。
- 通常のテキストでは係り受け等の曖昧性が生じ得るが、図式表現ではその問題は生じにくい。図式化は文章の不明瞭性の発見にもつながる。
- テキストと図面との対応付けを行うことで、特に物の発明の場合に、構成要素をグラフィカルに認識でき、内容把握が容易になる。

2.2.1　「特許請求の範囲」以外の低可読性

「特許請求の範囲」の可読性が低いという問題だけでなく、実施例など「発明の詳細な説明」にも範囲を広げて分析を行うことで、「特許請求の範囲」以外の部分の低可読性の問題を再認識することができた。

「発明の詳細な説明」中に書かれる文章は一見、通常の技術文書に書かれる文章のように見え、技術者にとって可読性が高いように思える。しかし、特許請求項の記述の一部を流用していたり、特許特有の言い回しを不必要かつ不自然に使っていたりする部分があり、その部分は特許請求項の低可読性の問題をそのまま受け継ぐことになる。また、曖昧な用語を使って長文で記述されている部分があり、その部分は慎重に読まないと意味を取り違えるリスクがある。

たとえば、Japio 仮想特許明細書（「照明装置および照明装置付き自転車」）の「第二の実施形態」には、以下のような記述がある。

> 第一の実施形態では、基板３３の外周方向において点滅等を発生するための間隔を介して複数のＬＥＤ・・・を使用しているが、第二の実施形態では、…基板３３の外周方向において間隔を詰めて複数のＬＥＤ・・・を配置することができる。

ここで、「外周方向において」という表現が 2 回出現しており、ＬＥＤを配置する際に「外周方向を向いて設置する」ということを意味していると考えられる。しかし、この場合における「において」という表現は明らかに特許特有の言い回しであり、結果として不自然かつ意味不明瞭な文章になっている。このような問題は、今回、上記仮想特許明細書の実施例の文章を構造化して図式により表現する作業の過程で明らかになった。

2.2.2　構造的曖昧性の解消

通常の 1 次元の文書と異なり、図式表現は樹状あるいはネットワーク状の構造を明示するので、構造的曖昧性が生じにくい。

例えば、「前記の存在表示機能部材(20)は、前記の前照灯(11)が照らす方向とは別の方向へ配置された側方表示機能部材(21)および後方表示機能部材(22)を備える。」という言語表現には、「側方表示機能部材(21)」と「後方表示機能部材(22)」という 2 つの部材が記述されているが、「前記の前照灯(11)が照らす方向とは別の方向へ配置された」のは(21)のみなのか、

それとも(21)と(22)の両方なのか曖昧である。しかし、下のような図式表現を用いることで(21)と(22)の両方であることが明確になる。

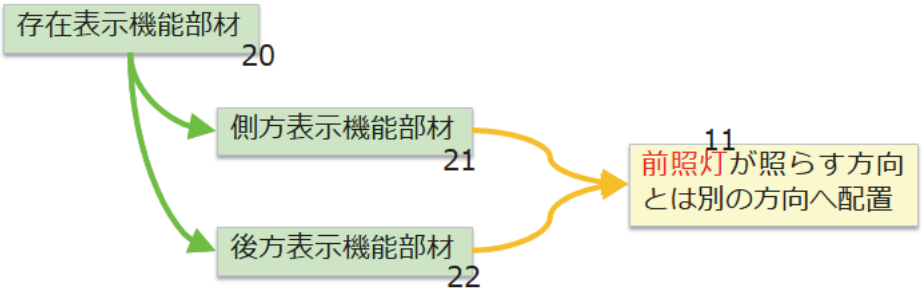


図 1

2.2.3　語彙的曖昧性の解消

日本語の格助詞には語彙的曖昧性がある。例えば、「存在表示機能部材(20)は、前照灯(11)の光軸とほぼ直角をなす中心軸を周回するように配置されている」という表現は「中心軸を周回する」という表現を含むが、この「を」は「存在表示機能部材(20)」が「中心軸」に対して何かの働きかけをするわけではない。つまり「目的格」ではなく、「中心軸の周りを周回する」という意味であり、英語であれば“around”のような前置詞で表現される意味である。しかし、意味別に色分けし必要に応じてラベル付けしたリンクを用いる下記のような図式表現とすることで格助詞の曖昧性が解消できる。

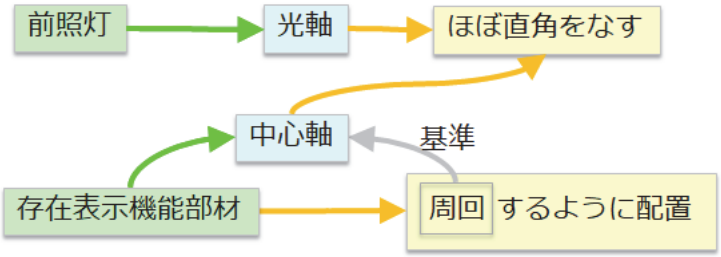


図 2

同様に、内容語の語彙的曖昧性もオントロジーを用いることによって解消する。

2.2.4　関係の明示の程度

上記仮想特許明細書の段落００３９において、「基板取付台２８」はモータの上方に位置すると記されているが、「存在表示機能部材２０」との関係は明示されていない。

> 【００３９】
> 前記の存在表示機能部材２０について詳しく説明する。
> 存在表示機能部材２０は、図５および図６に示すように、円筒状の存在表示機能ケース２４内の底部にモータホルダ２５を配置し、そのモータホルダ２５に弾性部材２６を介して回転駆動装置としてモータ２７を保持固定する。モータ２７の回転軸２７ａは、前照灯１１の光軸とほぼ直角をなすように構成している。モータ２７の上方には、円板状の基板取付台２８が、存在表示機能ケース２４の周壁から内側に突出した

回転体受部２９にて支持され、モータ２７の回転軸２７ａに直交する面で回転可能に配置されている。さらに、基板取付台２８は、モータ２７の回転軸２７ａによって回転駆動される構成である。

このため、段落００３９の記載を元に図式表現を作成すると下図のようになる。

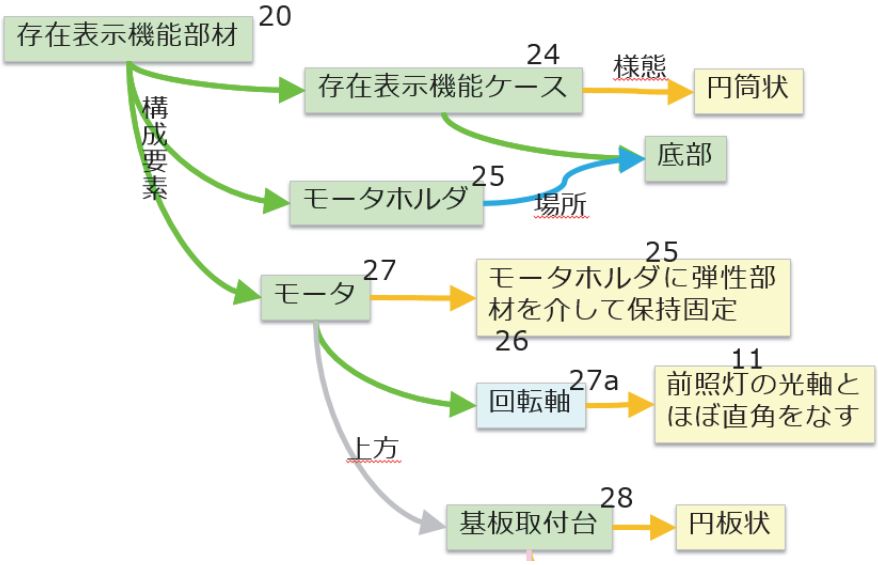


図 3

しかし、上記仮想特許明細書の図 5 (下にその一部を示す)によれば、「基板取付台２８」は「存在表示機能部材２０」の構成要素の１つと考えられる。

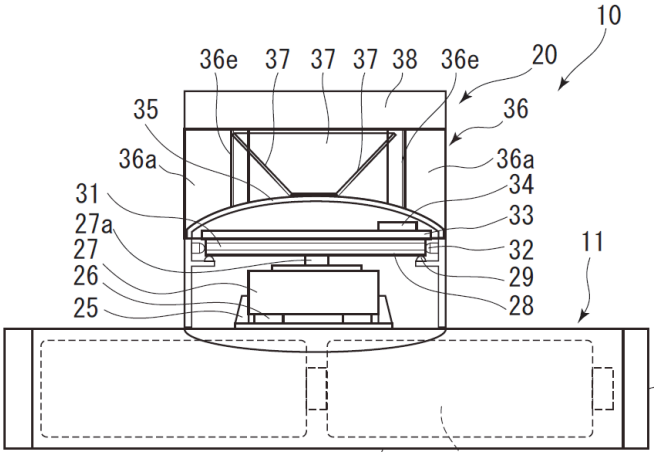


図 4

したがって、下図に示すように、「基板取付台２８」と「存在表示機能部材２０」の間にも構成要素を意味するリンクが存在し、上記仮想特許明細書の図 5 の意味構造にも含まれるはずである。テキスト文書に明示されないこのような関係も図式表現には含めるべきだろう。

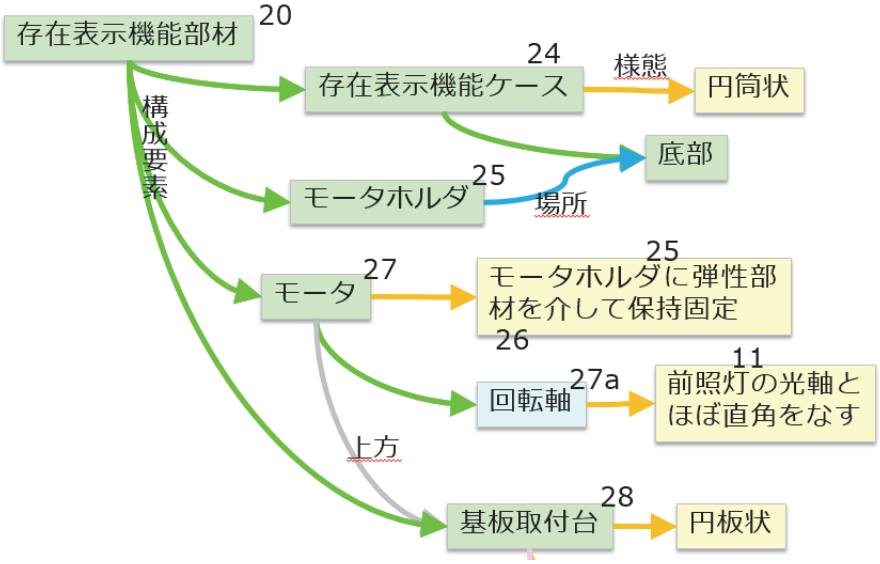


図 5

2.3　検討課題

ソフトウェアに関しては、形式的仕様記述によって仕様を厳密に表現することができ、矛盾の早期発見やプログラムの自動生成が可能になる。しかし、形式的仕様記述が難しいことも指摘されている。自然言語による仕様記述に潜む曖昧性を形式的仕様記述で明らかにすることの方に意義があるとの主張もある。

特許明細書の図式表現や構造化もソフトウェアの形式的仕様記述と似た側面を持つと思われるが、過度の厳密性を求めず、人間にとっての理解容易性に主眼を置くことで、記述の容易性と厳密性の間でバランスが取れる可能性がある。とすると、すべての内容語の意味を捉えるようなオントロジーは不要ということになる。今後は、文書の「適度な厳密化」のための表現方法を確立することが一般的な課題と思われる。

以下に具体的な検討課題を列挙する。

- ノードとリンクのラベル、ノードの粒度、リンクの端点と向きをどうすべきかは、少なくとも初心者にとっては判断が難しい。その判断を容易にするためには明確な定義が必要である。
- 照応・共参照を簡便に表現する方法が必要である。図式表現において各事物を１個のノードとして表現する方法と複数回出現させる方法が考えられる。前者はグラフが冗長性を含まない点では望ましいが、粒度が細かく複雑になり過ぎるきらいがあるので、後者と併用する必要があると考えられる。
- 図式の粒度が細かすぎると煩雑で却って理解を妨げる恐れもあるので、表示の粒度をダイナミックに調整することによって文書の意味構造の理解を支援するユーザインタフェースが必要である。
- 図面の細かい部分との対応を取るためには、図面データに複数レイヤーを設ける、メタデータを埋め込む、などの方策が必要と考えられる。

文書の図式表現に関する検討に加えて、特許文書内の概念を整理するために、AAMT（ア

ジア太平洋機械翻訳協会）が作成したシンプルな用語のデータ形式である UTX 用語形式 http://www.aamt.info/japanese/utx/ で用語データを作成し、概念の関係性（物理的包含関係、上位・下位概念）を含むオントロジーをシンプルに記述できるように UTX を拡張した。この成果を上述の図式表現に関する検討の結果と統合することも今後の課題である。

参考文献

[新井・尾崎 17] 新井 紀子・尾崎 幸謙: デジタライゼーション時代 に求められる人材育成. NIRA Opinion Paper, 31, 2017. http://www.nira.or.jp/pdf/opinion31.pdf

[Arum&Roksa 11] Richard Arum and Josipa Roksa: *Academically Adrift: Limited Learning on College Campuses*. University of Chicago Press, 2011.

Ⅳ
特許文書分科会報告書

Ⅳ　特許文書分科会報告書

1. 平成 29 年度　特許文書分科会　委員名簿

（五十音順・敬称略）

	氏　名	所　属
委　員	安彦　　元	ミノル国際特許事務所　所長・弁理士
委　員	石井　岳史	旭化成株式会社　研究・開発本部　知的財産部 知財リエゾン第二グループ　課長
委　員	岩永　勇二	平田国際特許事務所　パートナー　弁理士
委　員	黒川　　恵	日本弁理士会／阿部・井窪・片山法律事務所　弁理士
委　員	久保田真司	パナソニックＩＰマネジメント株式会社 グローバル出願部　品質強化担当参事
委　員	栗田　朋香	テルモ株式会社　知的財産部　主任
委　員 （主査）	谷川　英和	IRD 国際特許事務所　所長・弁理士
委　員	的場　成夫	有限会社夢屋　代表取締役　弁理士
委　員	横井　巨人	一般財団法人日本特許情報機構　特許情報研究所 調査研究部長
委　員	白土　博之	一般財団法人日本特許情報機構　特許情報研究所 研究企画課長

2. はじめに

2.1　昨年度までの取り組み

昨年度から「特許文書分科会」を立ち上げ、特許文書の「品質」という根本について、議論を開始した。特許文書の「品質」というものを明確にせずに、明確な特許文書の作成ルールを作成したり、特許文書を評価したり、特許文書をチェックするツールを開発したりすることは困難であるためである。

そして、昨年度、以下に示す特許文書品質特性のツリー構造の素案を作成した。ただし、以下の素案に対して、さらに詳細な検討が必要性であり、より納得性の高い特許文書品質特性を検討していくこととしていた。

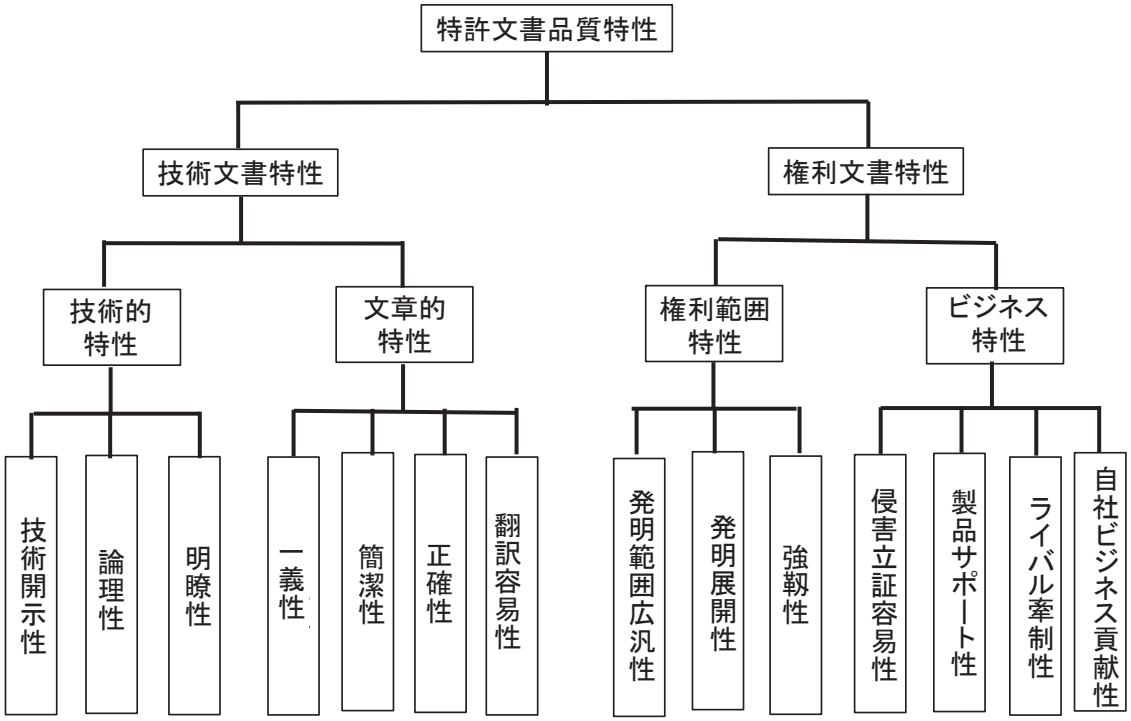


図１　**特許文書品質特性**（昨年度の素案）

2.2　特許文書分科会の目的等

特許文書分科会の現在の目的は、特許文書の「品質」を明らかにすることである。そのため、昨年度の議論をベースに、特許文書品質特性について検討を深めるとともに、各特性についての評価方法についても議論を進めた。なお、特性の評価の観点については、将来的な利用性を高めるために、定性的な評価と定量的な評価の２つの観点から検討した。

3. 他の対象の品質の調査

特許文書の品質を検討するのに先立ち、他の技術文書等について品質特性を分析したものを調査し、特許文書の品質特性を検討していく上での参考とすることとした。対象としては、特許文書の品質特性は一定の細分化が必要であると想定されたことから、一定の粒度に細分化して検討したものを調査した。この調査の結果、内容を確認することができた「システム開発文書品質モデル」と「システム・ソフトウェア製品の標準品質モデル」について分科会で確認した。

3.1　「システム開発文書品質モデル」（1）

「システム開発文書品質モデル」は、システムを開発するために作成する文書を対象として，文書の評価，文書の改善の目安にすることを目的として、システム開発文書品質研究会（ASDoQ[5]）が 2015 年 11 月に公開したものである。現在は、バージョン 1.0 が、ASDoQ のウエブサイト[6]で公開されている。開発文書には、要求仕様書、設計書、テスト仕様書、バグ管理票、打ち合わせ議事録などが含まれる。

環境変化の中で開発文書に求められる役割も変化しており、その課題として、（a）品質説明力の向上、（b）設計中心のソフトウェア開発、（c）システムの統合化、（d）開発拠点のグローバル化の 4 つの項目が挙げられていることから[7]、システム開発文書品質モデルはこのような点を考慮して作成されたものと考えられる。

表 1　開発文書品質の課題と環境変化の中で求められる開発文書の役割

開発文書品質の課題	環境変化の中で求められる開発文書の役割
品質説明力の向上	文書に根拠や理由を明晰かつ論理的に示す
設計中心のソフトウェア開発	設計や分析の検討を文書上で重ねる
システムの統合化	安全性や信頼性などを文書で検証する
開発拠点のグローバル化	翻訳に耐えうる明晰な日本語を記述する

「システム開発文書品質モデル」は、第 1 階層の品質特性、第 2 階層の品質副特性、第 3 階層の測定項目の 3 階層で構成されている。第 1 階層を最上位層にして、第 2，3 階層へと下位展開していくよう構成されている。

[5] ASDoQ：Association of System Documentation Quality
[6] システム開発文書品質研究会，「システム開発文書品質モデル Ver.1.0 を発行しました」，2015 年 5 月，URL: http://asdoq.jp/news/index.php?page=article&storyid=47　(2018 年 2 月 28 日アクセス)
[7] 塩谷敦子，「システム開発文書品質の課題と研究」，2014年11月URL: https://www.ipa.go.jp/files/000043976.pdf (2018年2月28日アクセス)

表 2　システム開発文書品質モデルを構成する 3 階層の内訳

階層：名称	説明
第 1 階層：品質特性	文書品質の大分類を示す
第 2 階層：品質副特性	各品質特性に含まれる小分類としての副特性を示す
第 3 階層：測定項目	品質特性・品質副特性の測定を行う項目

第 1 階層の品質特性は、(1)完全性、(2)論理性、(3)理解容易性、(4)可読性、(5)規範適合性の 5 つの特性に分類されている。「完全性」は、システム開発文書として特徴的な品質であり、「論理性」、「論理性」、「理解容易性」、「規範適合性」は、多くの実用的な文書に共通する品質となる。

表 3　システム開発文書品質モデル第 1 階層の品質特性の内訳

第 1 階層：品質特性	説明
(1)完全性	開発に必要十分な情報が記載されていること
(2)論理性	論理的に整合がとれていること
(3)理解容易性	理解しやすいこと
(4)可読性	読みやすいこと
(5)規範適合性	記述が文法や規則に則していること

第 2 階層の品質副特性は、第 1 階層の品質特性をそれぞれ 2 つ又は 3 つに細展開して、詳細化する構成となっており、第 2 階層内を 14 の特性に分類している。

表 4　システム開発文書品質モデル第 2 階層の品質副特性の内訳

第 1 階層：品質特性	第 2 階層：品質副特性	説明
完全性	合目的	読み手と目的を明示している
		目的に合致した内容を記述している
	正確	記述内容が正しい
	妥当	記述内容が妥当である
論理性	無矛盾	論理的な衝突（矛盾）がない
	一貫	論理展開が合理的で一貫している
	構造	内容の整理が合理的・体系的である
理解容易性	非曖昧	一意に解釈できる（一意性）
		動作または状態を特定できる（具体性）
	関係	各情報間の関係が明確である
可読性	簡潔	短文で，かつ簡潔に記述している
	統一	表記・表現方法および表現上の視点が統一されている
	表記工夫	内容の理解と解釈を助けるために，表記上の工夫がある

規範適合性	文法適合	言語の文法に則している
	記法適合	各種の設計表記法（状態遷移表，UML など）や標準記法などに則している
	基準適合	組織で定められた基準または標準に則している

第 3 階層には、測定項目を設けており、第 2 階層の品質副特性のそれぞれについて具体的な測定項目を例示している。システム開発文書品質の測定を行う際には測定対象とする文書等に応じて、カスタマイズすることが必要となる。

表 5　システム開発文書品質モデル第 3 階層の測定項目の内訳

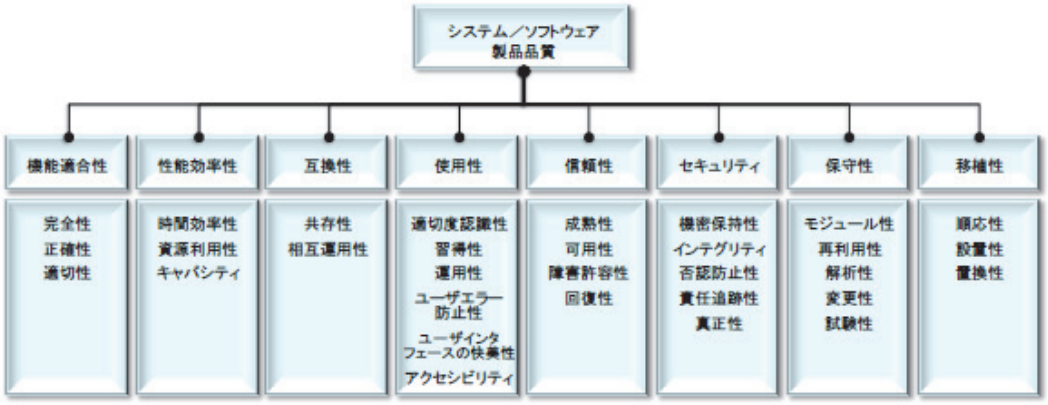
品質特性	品質副特性	測定項目
完全性	合目的	・文書の読み手（開発技術者，テスト技術者，発注者など）の明示 ・文書の使用範囲（提出先，配布先，開示範囲）の明示 ・文書の読み手に要求する条件（開発経験，保有スキルなど）の明示 ・開発の目的の明示（開発のスコープを含む．例：要求仕様書に開発の目的を書く） ・開発文書の目的の明示（例：実装に加えて将来の改訂開発を想定した設計書） ・記述単位（章・節・項・段落・図・表）の目的の明示 ・関連情報（工程，入力文書など）の明示
		・記述内容が文書の目的に合致 ・個々の記述単位(章・節・項・段落・図・表）の内容が各目的に合致 ・目的を達成するために必要な，情報，条件および項目の記載 ・目的を達成するために，理解しておくべき方針や意図の記載（例：設計書には設計方針を書くことを推奨する） ・記述に漏れが無い ・目的に適合する内容のみの記載 ・読み手の条件（開発経験，暗黙知など）を満たせば，読み手が目的を達成できる内容
	正確	・技術が正しく，実現可能
	妥当	・開発プロジェクトが許容する資源の範囲で開発可能 ・TBD を解消する決定時期・決定方法の明記
論理性	無矛盾	・項目間の矛盾がない ・条件間の矛盾がない
	一貫	・上位と下位と，原因と結果とが対応 ・理由または根拠の明記 ・記述間や文書間の整合性がとれ，追跡可能
	構造	・全体像を明示した後に詳細が記述されている（目次、全体構造の提示など） ・記述単位の見出しと内容とが合致 ・粒度が揃っている ・階層が合理的（例：木構造，is-a，has-a などを適切に使い分ける） ・内容に関する無駄な重複がない
理解容易性	非曖昧	・複数の解釈が成り立たない 例 1：複数の解釈が成り立つ接続語，否定表現などが利用されている （「～し（帰結・追加)」「すべて～でない（全部・部分)」） 例 2：係り受け（修飾する・される）の関係が明確でない
		・動作や状態などを特定 例：動作（「処理する」「制御する」など)，状態（「正常」「異常」など)， 性質(「速い」「大量の」など) ・値や範囲などを定義 ・専門用語や略語などの用い方を定義

	関係	・参照先や引用元を正しく明示 ・既知の知識に関連付く手がかりの記述 ・宣言や定義には，その理由や根拠の記述
可読性	簡潔	・一文一義 ・修飾語と被修飾語の距離が短い ・できるだけ肯定表現を使用し，二重否定は使用しない ・記述に重複がない
	統一	・文体（例：ですます調，である調など），用語および表記法の統一 ・視点の統一（例：主語をシステムに統一して記述する）
	表記工夫	・箇条書きまたは図表を適切に使用 ・インデント，改行および空行の使用による読みやすさの向上 ・読点の適切な使用による文章構成の明確化 ・読みやすいレイアウトやフォント種の使用 ・不必要なカタカナ表現を使用しない
規範適合性	文法適合	・誤字脱字がない ・主語と述語と，かつ，目的語と述語とが対応 ・接続詞，助詞などを正しく使用
	記法適合	・箇条書きのルールに適合 ・図の記載ルールに適合
	基準適合	・テンプレートに準拠 ・適用すべき基準または標準に準拠

3.2 「システム・ソフトウェア製品の標準品質モデル」(2)

3.2.1 システム・ソフトウェア製品の標準品質モデル

経済産業省　ソフトウェアメトリクス高度化プロジェクト　プロダクト品質メトリクス WG の「システム／ソフトウェア製品の品質要求定義と品質評価のためのメトリクスに関する調査報告書」によれば、システム・ソフトウェア製品の標準品質モデルを以下の図 2 に示すように規定している。



システム／ソフトウェア製品の品質（System and Software Product Quality）モデル[8]

図 2

システム･ソフトウェア製品の標準品質モデルにおける各特性の意味は、以下である。

（１）機能適合性(Functional Suitability)

製品やシステムが、定められた利用状況下で定められて包含されたニーズを満たす機能を提供する度合を示す。副特性として次がある。

完全性(Functional completeness)・・・機能がユーザの目的、定められたタスクをカバーしている度合

正確性(Functional correctness)・・・製品やシステムが必要な精度で正確な結果を与える度合適切性(Functional appropriateness)・・・機能が定められたタスクや目的の遂行を円滑に行う度合

（２）性能効率性(Performance efficiency)

定められた利用状況下で利用される資源量の性能の度合を示す。副特性として次がある。

時間効率性(Time behaviour)・・・要求を満たすために機能を実行するときのシステムの応答時間、処理時間及び処理能力の度合

資源利用性(Resource utilisation)・・・要求を満たすために機能を実行するときの使用した資源の量や種類の度合

キャパシティ(Capacity)・・・要求を満たすにあたり製品やシステムのパラメータの最大許容値

（３）互換性(Compatibility)

製品、システム、コンポーネントが他の製品、システムコンポーネントと情報を変換できる度合、また、同じハードウェアやソフトウェア環境を共有し、要求される機能を実行する度合を示す。副特性として次がある。

共存性(Co-existence)・・・他の製品へ有害な影響を与えずに、他の独立した製品と共通の環境や資源を共有して要求機能を効果的に実行する度合

相互運用性(Interoperability)・・・2つ以上のシステム、製品やコンポーネントが情報を交換し、その情報を利用できる度合

（４）使用性(Usability)

製品やシステムが、定められたユーザにより、定められた使用状況下で効果的、効率的、満足度が達成される度合を示す。副特性として次がある。

適切度認識性(Appropriateness recognisability)・・・製品やシステムがユーザニーズに適しているかどうかを、ユーザが認識できる度合

習得性(Learnability)・・・定められたユーザにより、定められた使用状況下で満足性、安全性、効果性、効率性のある製品やシステムの使い方の学習が達成される度合

運用性(Operability)・・・運用や管理の労力に係る度合

ユーザエラー防止性(User error protection)・・・システムがユーザを誤操作することから保護する程度
ユーザインタフェースの快美性(User Interface aesthetics)・・・ユーザインタフェースがフレンドリーで満足できるインタラクションを可能にする程度
アクセシビリティ(Accessibility)・・・製品やシステムが定められた使用状況下で目標を得る上で、幅広い層の特徴や能力を持つ人々により利用される度合

（５）信頼性(Reliability)
システム、製品やコンポーネントが制限時間内で定められた状況の下で機能を実行する度合を示す。副特性として次がある。
成熟性(Maturity)・・・システムが通常の運用の下で信頼性のニーズを満たす度合
可用性(Availability)・・・システム、製品やコンポーネントが必要とされるときに運用、接続できる度合
障害許容性(Fault tolerance)・・・システム、製品やコンポーネントがハードウェアやソフトウェア障害が存在する中で運用できる度合
回復性(Recoverability)・・・障害時に製品やシステムがデータを回復し、システム状態を再構築する度合

（６）セキュリティ(Security)
人やシステムによる読み込み・修正等が不当にアクセスされることなく情報やデータが保護されている度合を示す。副特性として次がある。
機密保持性(Confidentiality)・・・製品やシステムが、許可されたもののみがアクセスできるようデータを保証する度合
インテグリティ(Integrity)・・・システム、製品やコンポーネントがコンピュータプログラムやデータへの修正に対して許可されていないアクセスを防止する度合
否認防止性(Non-repudiation)・・・イベントやアクションがのちに拒否することができないよう、イベントやアクションが起こされたことが証明される度合（ディジタル署名等）
責任追跡性(Accountability)・・・エンティティのアクションが唯一のエンティティであると証明できる度合
真正性(Authenticity)・・・リソースや事項の身元が要求されるものであることを証明できる度合

（７）保守性(Maintainability)
製品やシステムが保守担当により修正するにあたっての効果性、効率性の度合を示す。副特性として次がある。
モジュール性(Modularity)・・・ひとつのコンポーネントの変更が他のコンポーネントへ最小の影響ですむよう、システムやコンピュータプログラムが別々のユニーク

なコンポーネントで構成される度合

再利用性(Reusability)・・・資産が複数のシステムや他の資産を構築する際に利用できる度合

解析性(Analysability)・・・製品やシステムの一部を変更したり変更された部分の特定や障害の原因や欠陥のために製品を診断したり、一つ以上の部分の変更に当たり影響を評価する際の効果性、効率性の度合

変更性(Modifiability)・・・製品やシステムが欠陥の発生や既存の製品品質の低下がなく、効果的、効率的に変更できる度合（ISO/IEC 9126シリーズでは変更性、安定性）

試験性(Testability)・・・システム、製品やコンポーネントのためにテスト基準を確立し、基準が満たされたかどうか定めるために実行する際の効果性、効率性の度合

（８）移植性(Portability)

システム、製品やコンポーネントが、あるハードウェア、ソフトウェアや運用、利用環境を他へ移行されるにあたっての効果性、効率性の度合を示す。副特性として次がある。

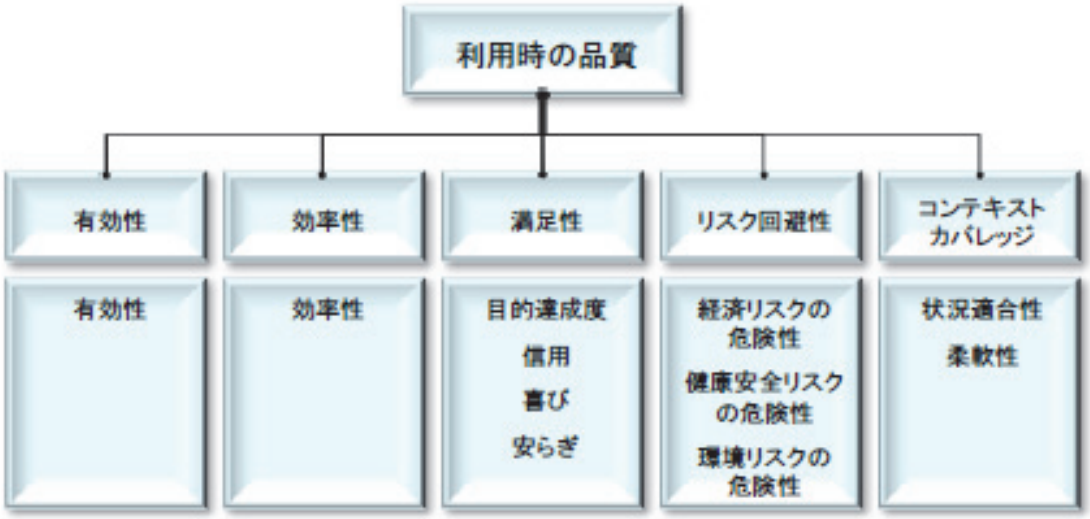
順応性(Adaptability)・・・製品やシステムが、異なるあるいは進化したハードウェアやソフトウェア、あるいは運用や利用可能な他の環境に効果的かつ効率的に順応できる度合

設置性(Installability)・・・製品やシステムが定められた環境に正しく設置（インストール）されたり撤去される（アンインストール）されたりする際の効果性、効率性の度合

置換性(Replaceability)・・・製品が同一の目的、環境で他のソフトウェア製品に置換（リプレース）される度合

3.2.2　利用時の標準品質モデル

経済産業省　ソフトウェアメトリクス高度化プロジェクト　プロダクト品質メトリクスWG の「システム／ソフトウェア製品の品質要求定義と品質評価のためのメトリクスに関する調査報告書」によれば、システム・ソフトウェア製品の利用時の標準品質モデルを以下の図３に示すように規定している。



利用時の品質(Quality in Use)モデル[5]

図 3

利用時の品質の各特性の意味は、以下である。

（1）有効性(Effectiveness)

利用者(ユーザ）が指定された目標を達成する上での正確さ、完全性の度合。

（2）効率性(Efficiency)

利用者(ユーザ）が目標を達成する際に、正確さと完全性に費やした資源の度合。

（3）満足性(Satisfaction)

製品、システムが定められた利用状況下で利用された時のユーザニーズに対する満足の度合。副特性として次がある。

目的達成度(Usefulness)・・・システムを利用する目標（ふるまいや最終結果）に対して実際に得た結果への満足の度合

信用(Trust)・・・製品、システムが想定されたふるまいをする能力の度合

喜び(Pleasure)・・・個人のニーズを遂行することから喜びを得る度合

安らぎ(Comfort)・・・身体的安らぎに対する満足度合

（4）リスク回避性(Freedom from risk)

製品やシステムが経済的状況、生活、健康、環境への潜在的リスクを軽減する度合。副特性として次がある。

経済リスクの危険性(Economic risk mitigation)・・・利用状況下で経済的状況、運用

効率、商業的所有物、評判、他の資源への潜在的リスクを軽減する度合
　健康安全リスクの危険性(Health and safety risk mitigation)・・・利用状況下で人への潜在的リスクを軽減する度合
　環境リスクの危険性(Environmental risk mitigation)・・・利用状況下で環境や資源への潜在的リスクを軽減する度合

（５）コンテキストカバレッジ(Context coverage)
　定められた利用状況や、想定外の状況で有効性、効率性、安全性、満足性を持って利用される度合。副特性として次がある。
　状況適合性(Context completeness)・・・要求で定められた利用状況下で有効性、効率性、安全性、満足性など利用される度合
　柔軟性(Flexibility)・・・要求で定められた利用状況以外で有効性、効率性、安全性、満足性を持って利用される度合

4. 特許文書の品質について

　品質の定義が容易ではない対象である技術文書、システム・ソフトウェア製品等の品質モデルを参考にして、特許文書の品質について考察した。
　つまり、特許文書の品質を検討する対象として、（１）特許文書品質特性、（２）特許文書品質の各特性の評価方法、（３）特許文書の品質評価の場面、の３つについて検討していくこととした。

4.1　特許文書品質特性

　特許文書の品質をできるだけ客観的に評価するために、比較的評価しやすい特性（項目と言っても良い）に分けることとした（図４参照)。

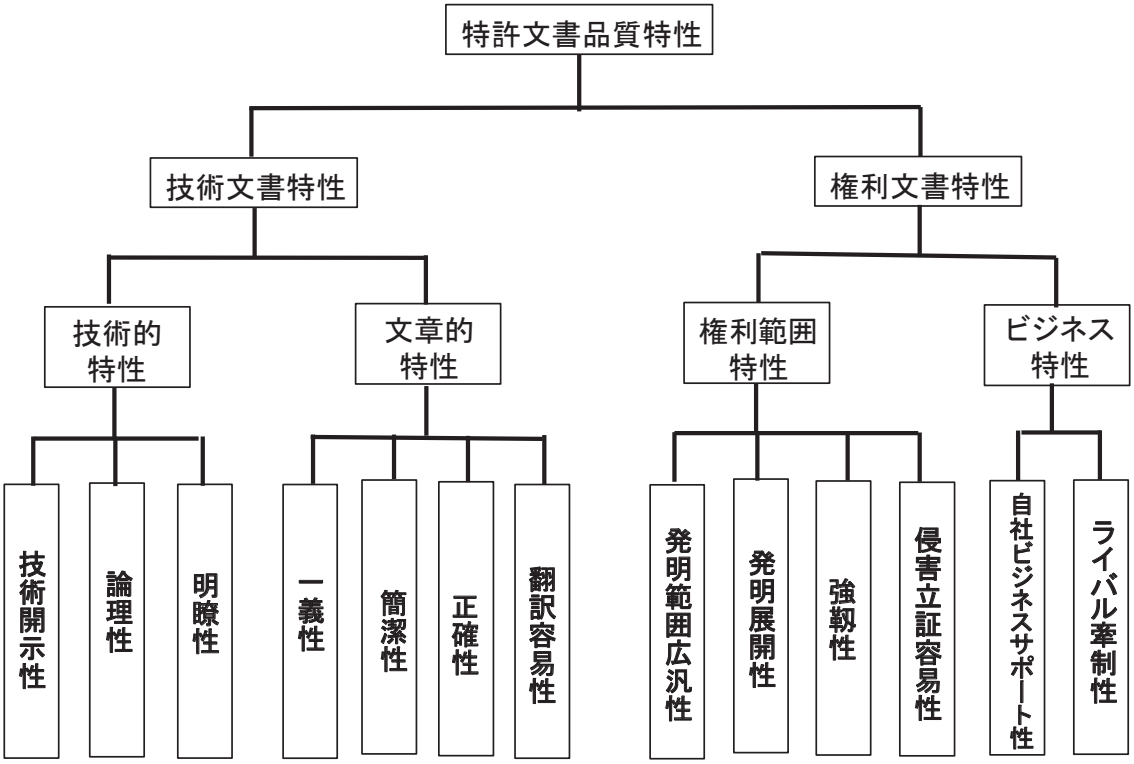


図 4　特許文書品質特性

以下、特許文書品質特性の定義等について説明する（図 3.1.2 参照）。

特許文書は、技術を普及していくための文献としての役割と権利を公示するための権利書としての役割を持つ。そこで、特許文書品質の大分類は、「技術文書特性」「権利文書特性」とした。技術文書特性は、技術を第三者に伝える場合の伝えやすさである。権利文書特性は、権利書としての役割の果たしやすさである。

また、「技術文書特性」を「技術的特性」と「文章的特性」に分けた。「技術的特性」は、技術が的確に記載されている度合いである。「文章的特性」は、文章としての適切な度合いである。

また、「技術的特性」を「技術開示性」「論理性」「明瞭性」に分けた。「技術開示性」は、発明が十分に説明されている度合いである。言い換えれば、特許法が要求しているサポート要件、実施可能性を満足している度合いである。「論理性」は、論理的である度合いである。「明瞭性」は、発明の技術説明が明瞭である度合いである。

また、「文章的特性」は、「一義性」「簡潔性」「正確性」「翻訳容易性」に分けた。「一義性」は、特許文書を構成する文章が多義的に解釈できない度合いである。「簡潔性」は、特許文書を構成する各文が簡潔な文である度合いである。「正確性」は、特許文書に誤記・技術的誤りが出現しない度合いである。「正確性」は、誤記が多いほど低くなる、と考えられる。「翻訳容易性」は、翻訳のし易さの度合いである。

また、「権利文書特性」を「権利範囲特性」「ビジネス特性」に分けた。「権利範囲特性」

とは、権利範囲に関する特性である。「ビジネス特性」は、権利を利用する場合に関係する特性である。

また、「権利文書特性」を「発明範囲広汎性」「発明展開性」「強靭性」「侵害立証容易性」に分けた。「発明範囲広汎性」は、発明の本質が抽出されており、無用な限定が無い度合いである。言い換えれば、権利範囲の広さの度合いである。「発明展開性」は、発明が十分に展開されている度合いである。「強靭性」は、拒絶、無効になりにくい度合いである。「侵害立証容易性」は、侵害の立証が容易である度合いである。

さらに、「ビジネス特性」を「自社ビジネスサポート性」「ライバル牽制性」に分けた。「自社ビジネスサポート性」は、自社製品をカバーしている度合いである。「ライバル牽制性」は、ライバルを牽制できている度合いである。

表 6　特許文書品質特性の定義

品質特性			品質特性の説明
大分類	中分類	小分類	
技術文書特性	技術的特性	技術開示性	発明が十分に説明されている度合い。特許法が要求しているサポート要件、実施可能性を満足している度合い。
		論理性	論理的である度合い
		明瞭性	発明の説明が明瞭である度合い
	文章的特性	一義性	多義的に解釈できない度合い
		簡潔性	簡潔な文の度合い
		正確性	誤記・技術的誤りが出現しない度合い
		翻訳容易性	翻訳しやすい度合い
権利文書特性	権利範囲特性	発明範囲広汎性	発明の本質が抽出されており、無用な限定が無い度合い。権利範囲の広さの度合い
		発明展開性	発明が十分に展開されている度合い。
		強靭性	拒絶、無効になりにくい度合い
		侵害立証容易性	侵害の立証が容易である度合い
	ビジネス特性	自社ビジネスサポート性	自社製品をカバーしている度合い
		ライバル牽制性	ライバルを牽制できている度合い

4.2　特許文書特性の評価方法

各特性から品質を評価することで、特許文書の品質を評価しようとするものである。品質を評価する際には、定量的な評価ができればより客観的な評価が可能になると考えられるが、定量的な評価を網羅的に行うことは必ずしも容易でないことから、定性的な評価と定量的な評価の両面について検討した。以下に検討の際に上がった評価観点を列挙しているが、議論中の内容であり、今後精査していく予定である。

4.2.1　定性的な評価方法

（1）技術開示性

技術開示性について、以下が定性的な評価を構成する項目の例である。

・請求項の発明が、発明の詳細な説明に記載されたものであるか（36 条 6 項 1 号、サポート要件）。
・発明が当業者が実施可能な程度に記載されているか（36 条 4 項 1 号、実施可能要件）。
・一般名称、専門用語が正しい意味で使用されているか。独自の用語・造語は、定義が記載されているか。
・発明の内容を正しく反映しているか。
・専門用語や独自用語、略語などは、一般的な名称と併記されているか。
・請求項に記載の発明特定事項（発明の本質）の目的と効果が記載されているか。
・「どのような発明特定事項か」「その結果どのような効果があるのか」が明確に把握できるか。
・実施形態の補足的説明（○○であってもよい等）が記載されているか。
・明細書の説明を補足するための図面、図面参照番号、フローチャートが適切に作成されているか。

（２）論理性

論理性について、以下が定性的な評価を構成する項目の例である。
・背景・従来技術・課題・解決手段・効果等のストーリーの筋が通っているか。
・クレームと明細書のストーリーとの技術的因果関係が明瞭であるか。
・請求項の順および各請求項間の関係性が明確に把握できるか。
・構成の説明順（上下関係）は、適切であるか（通常、大枠から細部へが好ましい）。
・請求項に記載の必須の発明特定事項（発明の本質）の目的と効果が明確に把握できるか。
・明細書の説明を補足するための図面、図面参照番号、フローチャートが適切に作成されているか。
・明細書の文章全体の構造が明確に把握できるか。
・クレームツリーの枝の順と明細書本文の説明順との整合性がとれているか。

（３）明瞭性

明瞭性について、以下が定性的な評価を構成する項目の例である。
・必要な説明が離れた箇所にある場合、それを示唆する表現が使用されているか（例：○○については後述する）。
・請求項の順および各請求項間の関係性が適切であるか。
・（場合によっては）発明特定事項の見出しが記載されているか（例：シャフト、コイルなど）。
・当業者の知識レベル（先行技術）を踏まえて発明特定事項をグルーピングしているか。
・発明特定事項は、当業者が理解できる基準を採用しているか。
・一般名称、専門用語が正しい意味で使用されているか。独自の用語は、定義が記載されているか。
・参照番号は適切につけられているか（上下階層に誤解を生じさせないか）。

・発明の名称は、発明の内容を反映した名称となっているか。
・明細書の見出しは、当業者の知識レベル（先行技術）を踏まえた名称を使用しているか。
・それぞれの実施形態、構成要件が容易に区別できるか。
・一般名称、専門用語が正しい意味で使用されているか。独自の用語は、定義が明確に記載されているか。わかりにくい表現について、補助的な説明文が設けられているか（例：すなわち、つまり、等による言い換え）。
・専門用語や独自用語、略語などは、一般的な名称と併記されているか。
・明細書の説明を補足するための図面、図面参照番号、フローチャートが適切に作成されているか。
・箇条書きや表組による表現を適切に使用しているか（例：実験データ等がある場合）。
・図面が何を説明するためのものか理解できるか。
・図面が、どの実施形態を示すものか明確か。図面が、どのような使用状態を示すものか明確か。
・図面中の指示・行為を示す要素（行為の方向を示す矢印や参照番号の引き出し線）が明確か。
・形状を正しく把握できる範囲で線の数を減らした、スッキリとした図面であるか。
・発明の理解を助けるアングルで描画されているか。
・説明の目的によっては、テクニカル図面ではなくデフォルメした図面を使用されているか（適度に拡大・縮小されている）。
・画像を掲載する場合、画像の内容が判別／判読できるものであるか（過度の縮小・画素数の間引きによる劣化は避けること）。必要に合わせて、説明上で強調したい箇所が明示されている（トリミングされている）ことが望ましい。〃

（４）一義性
　一義性について、以下が定性的な評価を構成する項目の例である。
・多義的な用語が使用されていても明細書中で一義に特定されていれば多義的とはしない

（５）簡潔性
　簡潔性について、以下が定性的な評価を構成する項目の例である。
・用語の統一性が図れているか。
・冗長用語、冗長表現（例：制御を実行する）が使用されているか・

（６）正確性
　正確性について、以下が定性的な評価を構成する項目の例である。
・用語が統一されているか。
・明細書中の符号と図面の統一が統一されているか。
・単数と複数が明確化されているか。
・先行詞が明確化されているか。

（７）翻訳容易性

翻訳容易性について、以下が定性的な評価を構成する項目の例である。

・主語があるか

・長文や複文が多いか

・助詞の誤使用が多いか

なお、翻訳容易性について、対象別に定性的な評価を構成する項目を挙げれば以下である。

［対象１］形式・書式

①クレームの階層の視覚的記述

②構成要素毎にブロック記述

③カンマ等の適切な区切り

④（一文でだらだら記述しない）

⑤発明カテゴリー毎に独立項、従属項のブロック化

⑥従属項の広さを段階的に狭くし、実施例、図面の開示順と整合

［対象 2］文章

①主体の明確化、主語と述語のペアー

②修飾と被修飾の最短配置

③先行修飾詞の一致、明確化

④単文以外での、目的語を省略しない

⑤重文・複文時に、主語と述語との関係性、態の一貫性の統一

［対象 3］単語・用語

①単数と複数の明確化

②先行詞の一致、明確化

⑨複合詞の回避（複数の漢字名詞を連続直結）

④構成要素名が同じ場合でも、「同一」か否かの区別

⑤カタカナ用語の明確化（英語のカッコ書き）

⑥冠詞を意識した別メモ（初出時に、a, one, a plurality of, 無冠詞等）

［対象 4］その他

①日本語特有の「方向」、「側」、「対向」、「部分」などの英訳時に不明確になり易い頻出パターンの本当の意図を確認

②企業内の方言、業界用語の回避、置換、「載置」、「押圧」等

（８）発明範囲広汎性

発明範囲広汎性について、以下が定性的な評価を構成する項目の例である。

・自社実施品がクレームインされているか。

・上位概念の用語が用いられているか。

・明細書中に、発明特定事項の定義が限定的に記載されていないか。

・明細書中に、発明特定事項の作用効果が限定的に記載されていないか。

（９）発明展開性

発明展開性について、以下が定性的な評価を構成する項目の例である。

・クレーム文言が明瞭でありクレーム文言と実施例との間に齟齬がないか。

・独立項が多いか。

・クレームに包含される例が豊富であるか。

（１０）強靭性

強靭性について、以下が定性的な評価を構成する項目の例である。

・クレーム文言が明瞭でありクレーム文言と実施例との間に齟齬がないか。

・クレームに包含される例が豊富であるか。

（１１）侵害立証容易性

侵害立証容易性について、以下が定性的な評価を構成する項目の例である。

・クレーム文言及び実施例の用語が、通常の語義や業界において用いられている用語と齟齬がないか。

・一般に、物クレームは高く、方法（製法）クレームは低い。

・発明特定事項を実施する主体が複数になっていないか。

・クレームに包含される例が豊富であるか。

（１２）自社ビジネスサポート性

自社ビジネスサポート性について、以下が定性的な評価を構成する項目の例である。

・自社製品・自社サービスに対応した　特許請求の範囲、図面、フローチャート　が存在するか。

・自社製品・自社サービスに対応した用語が使用されているか・

・自社製品・自社サービスに対応した登録商標が、明細書中に記載されているか。

・収益を守るための請求項が記載されているか。

（１３）ライバル牽制性

ライバル牽制性について、以下が定性的な評価を構成する項目の例である。

・他社製品・他社サービスと同一または近似した特許請求の範囲、図面、フローチャートが存在するか。

・他社製品・他社サービスと同一または近似した用語が使用されているか。

・先行技術の記載欄に、ライバル製品に関連した特許番号やそれに関連する記載があるか。

・他社製品・他社サービスに対応した登録商標が、明細書中に記載されているか。

4.2.2　定量的な評価方法

定量的な評価方法の検討結果を、各特性に対して、主として、「数」と「割合」の２つの

観点でまとめる。

（１）技術開示性

技術開示性について、以下が定量的な評価を構成する項目の例である。

［数］

・実施の形態の数
・実施例の数
・図面数
・造語の定義の数
・例示文の数

［割合］

・実施例でサポートされている請求項の割合
・図面で開示されている実施例の割合
・明細書本文に記載した効果の内、根拠となる記載の有るものの割合
・請求項に規定した数値範囲をカバーするデータの度合い
・明細書本文に記載した効果の内、裏付けデータ（実施例・比較例）の有るものの割合
・請求項に規定された数値範囲の内、裏付けデータ（実施例・比較例）の有るものの割合

（２）論理性

論理性について、以下が定量的な評価を構成する項目の例である。

［数］

・クレームツリーの枝にぶら下がっている請求項に番号飛びの箇所の数

［割合］

・発明特定事項と効果が組として記載されている請求項の割合
・請求項の発明特定事項の内、明細書に例示が記載されているものの割合

（３）明瞭性

明瞭性について、以下が定量的な評価を構成する項目の例である。

［数］

・造語の定義数
・例示文の数
・定義が記載されている略語の数

［割合］

・請求項に規定の造語の内、本文に定義が記載されているものの割合
・請求項の発明特定事項の内、本文に例示が記載されているものの割合
・図面の内、明細書本文中に説明があるものの割合
・図面の構成要素の内、明細書本文中に説明があるものの割合
・明細書の説明中の見出しの文書全体の中の割合

（４）一義性

一義性について、以下が定量的な評価を構成する項目の例である。

［数］

・多義的な用語の数

・多義的な文の数

［割合］

・単文率

・複文率

・同一文章での同一助詞の使用率

・単数と複数の混在率

・主体の欠落率

・能動態の使用率（受動態の使用率）

・先行詞の欠落率

・多義用語の使用率

・多義的な用語が使用されていても明細書中で一義に特定されていれば多義的とはしない

（５）簡潔性

簡潔性について、以下が定量的な評価を構成する項目の例である。

［数］

・長文の数

・複文の数

・重複記載の数

［割合］

・冗長用語の使用率

・冗長修飾詞の使用率

・能動態の使用率

（６）正確性

正確性について、以下が定量的な評価を構成する項目の例である。

(数)

・誤記・誤字・脱字の数

・技術的誤りの数

(割合)

・クレームと実施例との用語の統一率

・図面用語、参照符号の統一率

・実施例を通した、用語の統一率

・単数と複数の混在率

・先行詞の欠落率

・誤記の発生率
・不適切句読点の使用率
・不適切用語の使用率
・不適切文法の使用率

（７）翻訳容易性
　翻訳容易性について、以下が定量的な評価を構成する項目の例である。
［数］
・主語の無い文の数
・長文の数
・複文の数
・助詞の誤使用の数
［割合］
・クレームの階層による視覚的記述率
・構成要素毎のブロック記述率
・適切句読点の使用率
・発明カテゴリー毎に、独立項、従属項のブロック化率
・従属項の広さの段階的記述率
・古い専門用語、業界用語の使用率
・企業内の方言の使用率
・長い複合詞の使用率
・独特なカタカナ用語の使用率
・日本語特有の「方向」、「側」、などの不明確になり易い用語の使用率
・難解用語の定義の記述率

（８）発明範囲広汎性
　発明範囲広汎性について、以下が定量的な評価を構成する項目の例である。
［数］
・独立項の発明特定事項数
・独立項の文字数
・独立項の発明特定事項に対する修飾数
・格成分数
・１つの請求項について、課題や効果の数が３つ以上あると、権利範囲が狭く解釈されるおそれが出てくる。

（９）発明展開性
　発明展開性について、以下が定量的な評価を構成する項目の例である。
［数］

・クレーム数
・独立項数
・クレームツリーのネストレベル（上位レベル、中位レベル、下位レベルの深さ）
・コンビネーション・サブコンビネーションの数
・明細書中に作成した num（クレームの数＋クレームアップ可能な明細書中の構成の数）
・明細書及び図面における実施例数
・数値限定発明において、意義が記載された数値
・測定値がパラメータとなったクレームにおいて、測定方法の記載の有無

（１０）強靭性
強靭性について、以下が定量的な評価を構成する項目の例である。
［数］
・背景技術の先行文献数
・明細書中に作成した num（クレームの数＋クレームアップ可能な明細書中の構成の数）
・明細書及び図面における実施例数

（１１）侵害立証容易性
侵害立証容易性について、以下が定量的な評価を構成する項目の例である。
［数］
・格成分数

（１２）自社ビジネスサポート性
自社ビジネスサポート性について、以下が定量的な評価を構成する項目の例である。
［数］
・実施報償に関する情報、製品の売上情報等の自社固有の管理情報

（１３）ライバル牽制性
ライバル牽制性について、以下が定量的な評価を構成する項目の例である。
・契約情報等

4.3　特許文書の品質評価の場面

出願前、出願後・権利化前、権利化後に至るまで、特許文書の品質を評価する場面は種々あり得る。評価するタイミングによって注目する観点が異なると考えられることから、タイミングごとの評価方法についても検討が必要である。

（１）出願前の特許文書の品質評価の場面
まず、「出願前の代理人による品質保証」のための特許文書の品質評価がある。この評価は、特許文書の作成補助者の上司や代理人となる特許事務所所長等による評価である。

また、「特許事務所評価」のための特許文書の品質評価がある。この評価は、出願人による評価である。

また、「所員評価」のための特許文書の品質評価がある。この評価は、特許事務所所長や特許文書の作成補助者の上司等による評価である。

さらに、「発明者育成」のための特許文書の品質評価がある。この評価は、知財担当や発明者の上司等による評価である。

（２）出願後・権利化前の特許文書の品質評価の場面

まず、「外国出願時・国内段階移行」のための特許文書の品質評価がある。この評価は、出願人による評価である。

また、「中間処理時」の特許文書の品質評価がある。この評価は、出願人または代理人による拒絶理由の内容等を踏まえた評価である。

また、「他社特許の権利化可能性分析」のための特許文書の品質評価がある。この評価は、主として競合企業による評価である。

また、「特許事務所評価」のための特許文書の品質評価がある。この評価は、出願人による評価である。

また、「所員評価」のための特許文書の品質評価がある。この評価は、特許事務所所長や特許文書の作成補助者の上司等による評価である。

さらに、「発明者育成」のための特許文書の品質評価がある。この評価は、知財担当や発明者の上司等による評価である。

（３）権利化後の特許文書の品質評価の場面

まず、「権利行使段階」の特許文書の品質評価がある。この評価は、権利者による評価である。

また、「クロスライセンスのための特許抽出」のための特許文書の品質評価がある。この評価は、権利者による評価である。

また、「特許権の棚卸し」のための特許文書の品質評価がある。この評価は、権利者による評価である。

また、「他社特許の自社ビジネスへの影響分析」のための特許文書の品質評価がある。この評価は、競合企業等による評価である。

また、「他社技術の導入」のための特許文書の品質評価がある。この評価は、競合企業による評価である。

また、「特許事務所評価」のための特許文書の品質評価がある。この評価は、出願人による評価である。

また、「所員評価」のための特許文書の品質評価がある。この評価は、特許事務所所長や特許文書の作成補助者の上司等による評価である。

また、「発明者育成」のための特許文書の品質評価がある。この評価は、知財担当や発明者の上司等による評価である。

4.4　本年度の議論結果のまとめ

以下の表 7 が本年度の議論結果のまとめの表である。本表に示すように、本年度では、「特許文書品質特性」「各特性の評価方法（定性的，定量的）」について、特に議論した。定性的、定量的な評価の観点について、検討を進めた結果、多数の評価項目を上げることができ、各品質特性から特許文書の評価が一定程度可能であると考えられる。今後、特許文書の評価をより精緻に実施することができるよう、その評価を必要とする場面を想定しつつ、より使いやすい観点となるようにさらなる考察が必要である。

表7　本年度の議論結果のまとめ

小分類	品質特性の説明	評価方法(定性的)	パラメータ(定量的)	
			数量	割合
技術開示性	発明が十分に説明されている度合い。	・特許請求の範囲において、全ての実施形態が含まれるように記載されているか？ ・請求項に記載の発明特定事項（発明の本質）について、明細書中に十分な記載があるか？	・実施の形態の数、実施例の数 ・図面数	・実施例でサポートされている請求項の割合 ・請求項の数値範囲をカバーするデータの割合
論理性	論理的である度合い	・背景・従来技術・課題・解決手段・効果のストーリーの筋が通っているか？ ・明細書の文章全体の構造が明確に把握できるか？	・クレームツリーの枝にぶら下がっている請求項に番号飛びが存在する場合の飛びの数	・発明特定事項と効果の組が記載されている割合 ・請求項の発明特定事項の例示の記載率
明瞭性	発明の説明が明瞭である度合い	・一般名称、専門用語が正しい意味で使用されているか？ ・造語や略語は、定義が記載されているか？	・造語や略語の定義の数	・請求項の造語の定義率 ・請求項の発明特定事項の例示の記載率
一義性	多義的に解釈できない度合い	・定義の無い多義的な用語・文が使用されていないか？	・多義的な用語の数 ・多義的な文の数	・単文率または複文率 ・能動態または受動態の使用率
簡潔性	簡潔な文の度合い	・用語の統一性が図れているか？ ・冗長用語、冗長表現（例：制御を実行する）が使用されているか？	・長文、複文の数 ・重複記載の箇所の数	・冗長用語、冗長修飾詞の使用率 ・能動態の使用率
正確性	誤記・技術的誤りが出現しない度合い	・用語が統一されているか？ ・明細書中の符号と図面の統一が統一されているか？	・誤記、誤字、脱字の数 ・技術的誤りの数	・用語の統一率 ・誤記、誤字、脱字の発生率
翻訳容易性	翻訳しやすい度合い	・主語があるか？ ・長文や複文が多いか？	・主語の無い文、長文、複文の数 ・助詞の誤使用の数	・長い複合詞の使用率 ・日本語特有の不明確になり易い用語の使用率
発明範囲広汎性	権利範囲の広さの度合い	・上位概念の用語が用いられているか？ ・明細書中に、発明特定事項の定義が限定的に記載されていないか？	・独立項の発明特定事項数、文字数、発明特定事項に対する修飾数、格成分数 ・1つの請求項についての課題や効果の数	
発明展開性	発明が十分に展開されている度合い。	・独立項が多いか？（独立項が多いと、訂正要件を満たす可能性が高くなる。） ・クレームに包含される例が豊富であるか？	・クレーム数 ・クレームツリーのネストレベル（深さ）	
強靱性	拒絶、無効になりにくい度合い	・クレーム文言が明瞭でありクレーム文言と実施例との間に齟齬がないか？ ・クレームに包含される例が豊富であるか？	・背景技術の先行文献数 ・実施例数	
侵害立証容易性	侵害の立証が容易である度合い	・発明特定事項を実施する主体が複数になっていないか？ ・クレームに包含される例が豊富であるか？	・格成分数	
自社ビジネスサポート性	自社製品をカバーしている度合い	・自社製品・自社サービスに対応した特許請求の範囲、図面、フローチャートが存在するか？ ・収益を守るための請求項が記載されているか？	・実施報償に関する情報、製品の売上情報等の自社固有の管理情報	
ライバル牽制性	ライバルを牽制できている度合い	・他社製品・他社サービスと同一または近似した特許請求の範囲、図面が存在するか？ ・他社製品・他社サービスと同一または近似した用語が使用されているか？	・契約情報（ライセンス契約を締結している等）	

5. おわりに

特許文書分科会では、特許文書の作成の根本に立ち返って、特許文書の「品質」とは何か、をテーマに議論を続けている。

「特許文書品質モデル」の構築に当たって、「特許文書品質特性」と「特許文書品質の各特性の評価方法」「特許文書の品質評価の場面」の３つの観点から議論を進めている。

今後、特許文書品質の各特性の評価方法、特許文書の品質評価の場面のさらなる検討、および特許文書品質モデルの普及に向けた方策の検討等を行っていきたい。

参考文献

(1)［ASDoQ　15］システム開発文書品質研究会（ASDoQ：Association of System Documentation Quality）:『システム開発文書品質モデル Ver.1.0』, 2015.05, URL: http://asdoq.jp/common/fckeditor/editor/filemanager/connectors/php/transfer.php?file=/uid000003_E69687E69BB8E59381E8B3AAE383A2E38387E383AB5F3135303531312E706466　(2018 年 2 月 28 日アクセス)

(2)［SMHP　10］　経済産業省　ソフトウェアメトリクス高度化プロジェクト　プロダクト品質メトリクスWGの『システム／ソフトウェア製品の品質要求定義と品質評価のためのメトリクスに関する調査報告書』,2011.03, URL:http://www.meti.go.jp/policy/it_policy/softseibi/metrics/product_metrics.pdf#search=%27%E3%82%BD%E3%83%95%E3%83%88%E3%82%A6%E3%82%A7%E3%82%A2+%E5%93%81%E8%B3%AA%27（2018年2月28日アクセス）

V
産業日本語研究会ワークショップ講演

平成 29 年 12 月 19 日に開催した産業日本語研究会ワークショップでは、文章の品質や作成能力という点で産業日本語活動と共通する取り組みをしている外部講演者より、3 件の講演をいただいた。以下に当日の講演内容を掲載する。

１．「いかに情報をシェアするか
～オントロジー構築と共通語彙構築の経験から～」

いかに情報をシェアするか
～オントロジー構築と共通語彙構築の経験から～

武田英明
国立情報学研究所
takeda@nii.ac.jp

社会に役立つオントロジー・語彙

- 社会に役立つオントロジー・語彙を提供するにはどうしたら良いか。
  - オントロジー・語彙の構築の仕方
    - 農作業オントロジー構築の経験から
  - オントロジー・語彙を維持の仕方
    - 共通語彙基盤での経験から

農作業オントロジー構築の経験から

農作業名称の標準化

■ **現状**

農業ITシステムの普及により農業現場では管理や記録など作業を効率的に行うことが可能となった。そして、膨大なデータが発生しており、そのデータは統計・分析により作業の最適化、予測に重要な資料として期待されている。

■ **問題点**

基準となる情報が標準化されていないため、ベンダーは独自のデータ項目を定義して対応しており、これがデータの連携、統合に妨げとなる。

http://www.toukei.maff.go.jp/dijest/kome/kome05/kome05.html

■ **本研究の目的**

農作業名称における既存語彙体系について考察し、その考察に基いて農作業名称の標準語彙となる農作業基本オントロジーを構築する。

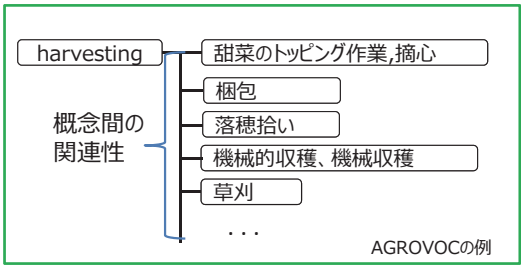
農作業オントロジーの構築 – シソラスとオントロジー

■ **シソーラス (Thesaurus)**

単語を同義語や意味上の類似関係、包含関係などによって分類する体系。(AGROVOC)

効率的な関連語検索が可能。

上位/下位階層の関係性が明確に定義されていないため同じ階層の概念間,すなわち兄弟関係の概念の関連性も定義されていない。

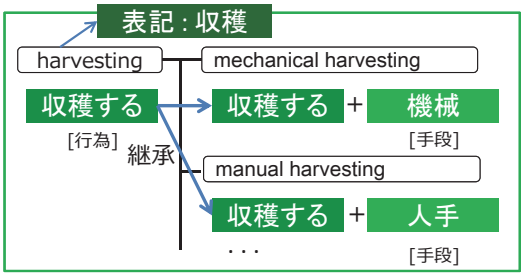


■ **オントロジー(Ontology)**

対象の概念と概念間の相互関係を定義する体系。

- 上位クラスが持つ属性を下位が継承
- 概念と表記を分離

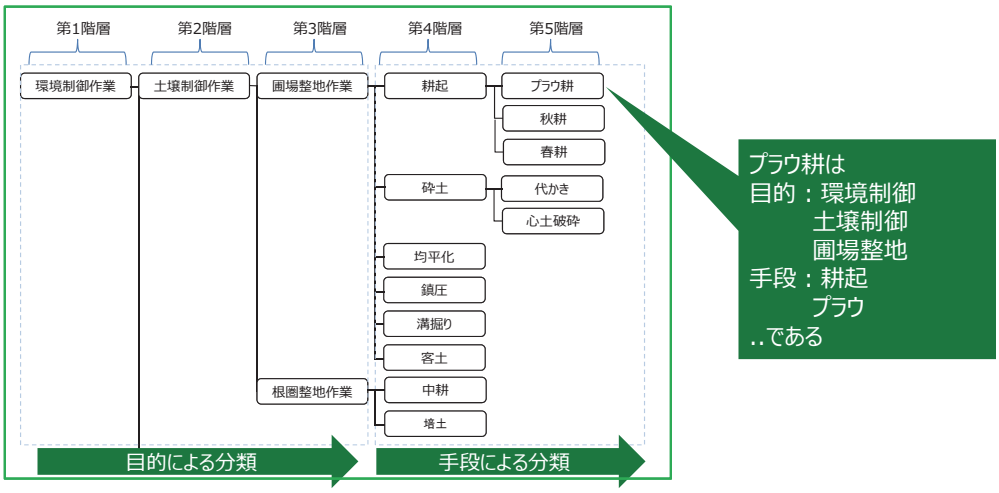
➡ 農業分野におけるドメインオントロジーを構築する



農作業オントロジーの構築 – 農作業の基本概念と構造化(1/2)

■ **農作業の構造概念**

「農作業」を「作物生産作業」、「環境制御作業」、「管理作業」、「間接作業」に分類し上位概念とする。そして、属性として「目的」と「手段」を用意し2層構造とし、分類基準とする。



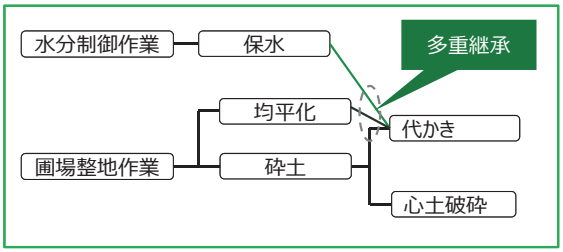
「環境制御作業」の例

農作業オントロジーの構築 – 農作業の基本概念と構造化(2/2)

■ **概念間の多重継承**

農作業の中では複数の目的で行われる作業が存在しており、それを概念の定義に適用させるために多重継承を考慮する必要がある。

例)「代かき」 目的 「砕土」
「保水」
「均平化」



■ **概念と表記**

同じ作業に対して複数の表記が存在する場合に対応する必要がある。

- ひらがなと漢字による表記　例)「播種」と「は種」、「代かき」と「代掻き」など
- 英語と漢字による表記　例)「ハードニング」と「硬化」など

農作業オントロジーの構築 – 農作業基本オントロジー(1/2)

■ **AAO : Agricultural Activity Ontology**

農作業における基本概念185語を網羅したオントロジー

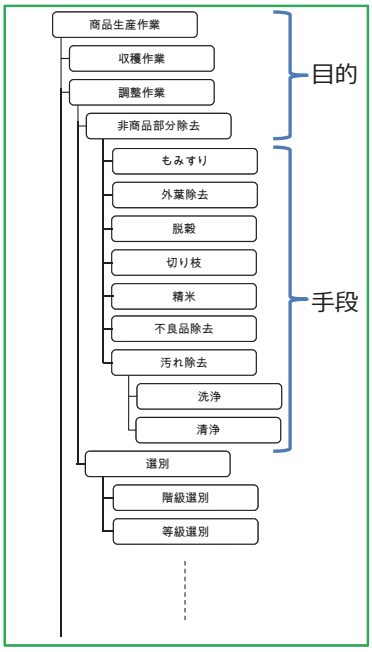
■ **特徴**

「目的」と「手段」による2層構造
概念の多重継承と複数の表記に対応

■ **公開**

"CAVOC : Common Agricultural VOCabulary"
`http://www.cavoc.org` でAAO ver.0.94を公開中

- 検索
- 名前空間

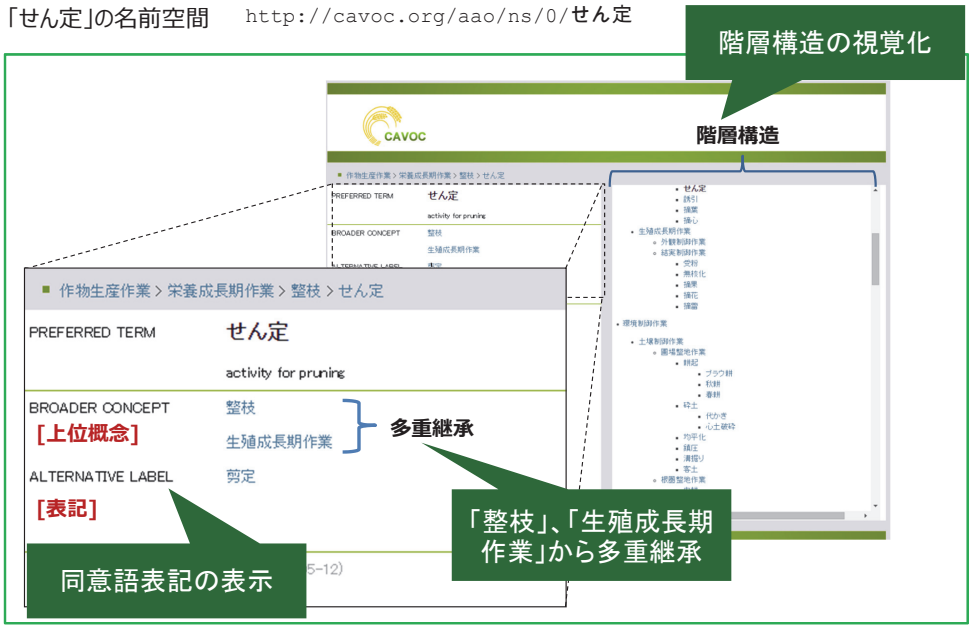


「商品生産作業」の例

農作業オントロジーの構築 – 農作業基本オントロジー(2/2)

■ **www.cavoc.org**

「せん定」の名前空間　http://cavoc.org/aao/ns/0/せん定

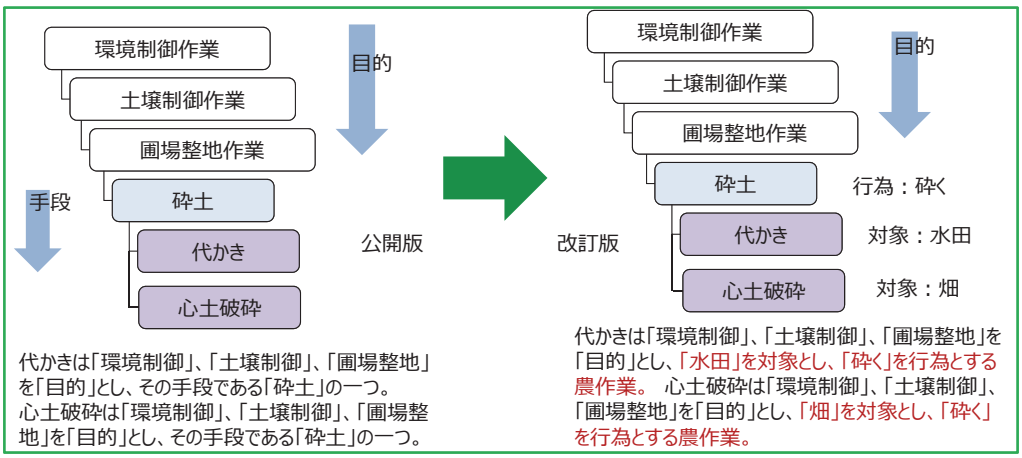


今後の課題 (1/5)

■ **属性の具体化**

公開版(ver 0.94)では「目的」と「手段」による2層構造となっており、農作業の概念を定義するために属性の数が十分ではない。農作業名称の多様性と規模に対応するために属性を拡張する必要がある。

改訂版では属性として「作物」、「目的」、「行為」、「対象」、「手段(=道具)」、「時期」を用い、より明確な定義を行い、分類を行う。



農作業オントロジーの設計(3)

■ **別名の表記(同意語)**

農作業名称は同じ作業なのに様々な別名が存在する。
1) 作業名に含まれる漢字が常用漢字ではない場合：は種 = 播種
2) 英語の作業名と日本語の作業名が混在する場合：ハードニング = 硬化
3) 最後に同じ作業でありながら判断の基準によって名称が異なる場合：秋耕 = 荒起こし
は種 = 種まき

■ **記述論理に基づく構築**

記述論理を用いて表現

PREFERRED TERM	**せん定**
	せんてい
	Pruning
BROADER CONCEPT	草姿制御作業
	生殖成長作業
ALTERNATIVE LABEL	剪定
	[行為] 切る
	[対象] 枝

せん定 ⊒ 農作業
⊓∀ 目的.(草姿制御 ⊔ 生殖成長)
⊓∀ 行為.切る
⊓∀ 対象.枝 (1)]

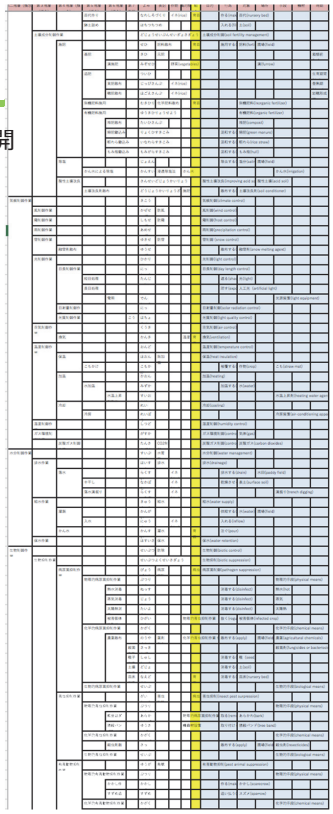
せん定 ≡ 剪定

農作業基本オントロジー

農作業基本オントロジー(AAO, Agriculture Activity Ontology)の公開

Ver.	公開日	語数	属性数	階層の最大深さ
0.94	2015/05/12	185	2	6
1.00	2015/11/02	301	7	6
1.10	2016/02/12	330	7	7
1.31	2016/04/22	355	8	7

[http://www.cavoc.org/aao/]



オントロジー構築手順

- 設計プロセス
  - 0th Step: プロジェクト形成
  - 1st Step: サーベイ
  - 2nd Step: データ分析
  - 3rd Step: 第1バージョン構築
  - 4th Step: 形式化（記述論理導入）
  - 5th Step: 専門家評価

設計プロセス
- 0th Step: プロジェクト形成-

- SIPプロジェクトの中のサブプロジェクト
  - メンバー
    - 農研機構2名
    - NII 2名

設計プロセス
- 1st Step: サーベイ -

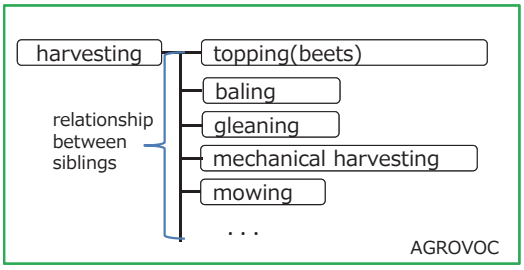
- 既存の語彙、タクソノミー、オントロジーの調査
  - Agrovoc: FAO作成. この分野で最も著名な語彙
    - 国際的
    - 継続的メンテナンス
    - 機械可読 (LOD)
  - Agropedia
    - 日本語
    - 説明あり
  - 農水省ガイドライン
    - 公式
    - 農業統計

AGROVOC

AGROVOC はFood and Agriculture Organization(FAO)が用意するシソーラスで、約32000個の用語で、農業、漁業、食品、環境、関連分野をカバーする。

- **シソーラス Thesaurus**

  AGROVOC は単語を同義語synonym, 狭義/広義narrower/broader, 関連語related relationshipで繋げている。



稲作等、日本の農業の特徴的な部分はそれほどサポートされていない。

- **economic activities**
  - agriculture
    - agricultural practices
      - **agronomic practices**
        - beating
        - bundling
        - handling
          - bulk handling
          - fish handling
          - **food handling**
          - loading
          - mechanical handling
          - pneumatic handling
          - pumping
          - stacking
          - unloading
            - fish unloading
        - harvesting
          - baling
          - gleaning
          - haymaking
          - **manual harvesting**
          - mechanical harvesting
          - mowing
          - picking
          - tapping
            - resin tapping
          - topping (beets)
          - windrowing
        - lifting
        - threshing
          - **manual threshing**
          - **mechanical threshing**
      - animal husbandry
        - animal husbandry methods

- economic activities
  - 農業
    - agricultural practices
      - agronomic practices
        - beating
        - 結束
        - 運搬
          - 大量運搬、大量取り扱い、ばら荷扱い、バラ積取扱
          - 水産物処理
          - food handling
          - 載荷
          - 機械運搬
          - 空気による流動化操作
          - ポンプ揚水
          - 積み上げ
          - 荷おろし
            - 水産物水揚げ
        - 収穫
          - 梱包
          - 落穂拾い
          - 乾草作り
          - manual harvesting
          - 機械的収穫、機械収穫
          - 草刈
          - 葉ぬき
          - タッピング
            - 樹脂採取、乳液を採取する切付け
          - 甜菜のトッピング作業、摘心
          - 地干し列、地干し刈り
        - 根菜掘上げ
        - 脱穀
          - manual threshing
          - mechanical threshing
      - 畜産業
        - 家畜飼養法

大項目	中項目	小項目（例示） 【ユーザーによって拡張（組み替え）可能】
1 種子等予措	1 選種	種もみの選種
	2 消毒	種子消毒
	3 催芽	浸種
	4 その他	球根予措、種子コーティング
2 育苗	1 床土・苗床作り	
	2 は種	苗箱播種　等
	3 施肥	
	4 かん水	
	5 換気	
	6 間引き	
	7 防除	苗床消毒
	8 移植	ずらし、間引き移植
	9 接ぎ木	
	10 その他	苗代管理、除草、こもかけ等
3 耕起整地	1 耕起・砕土	荒起し、秋田起し、代かき、ハロー等
	2 整地	
	3 畝立て	
	4 温室、ビニールハウスの本ぼ床土作り・床土入れ換え	
	5 その他	畦塗り
4 基肥施肥	1 運搬	
	2 配合	
	3 施肥	
	4 養液栽培における定植時に行う養液作りや施肥	
	5 その他	緑肥すき込み、秋落ちを防ぐための客土、施肥溝掘り等
5 は種・定植（本圃）	1 は種・覆土	直まき、種まき
	2 苗取り・運搬	
	3 定植・移植	
	4 補植	
	5 その他	
6 追肥施肥	1 運搬	
	2 配合	
	3 施肥	
	4 養液栽培における定植後の養液の濃度管	
	5 その他	
7 除草	1 中耕	
	2 除草剤散布	
	3 土入れ・土寄せ	
	4 草刈り・下刈り	
	5 マルチング	
	6 その他	
8 病害虫防除（雑草以外）	1 化学的防除（農薬散布）	
	2 化学的防除（土壌消毒）	
	3 物理的防除	
	4 生態的防除	
	5 被害茎等の除去及び焼却	
	6 その他	
9 管理	1 圃場（けい畔を含む）及び農道の整備・修復	畦畔除草、土壌改良材散布、圃場周り除草、防風林枝払い　等
	2 施設の整備・修繕	ハウス修繕、機械整備、果樹棚の補修、マット交換・掃除 保温施設の組立て、取り壊し、暖房器具の組立て、設置、取り壊し
	3 かん排水	かん水、落水、水管理、落水溝掘り等
	4 整枝	せん定、支柱立て、誘引、摘芯、摘果等
	5 生育調節（整枝以外）	麦踏み、間引き、ホルモン剤処理、受粉、袋かけ等
	6 防風・防霜	こもかけ、防風ネット設置　等
	7 保温・換気	
	8 見回り	生育調査、作柄見回り　等
	9 収穫後の株（根）等の除去	
	10 その他	鳥獣害対策、融雪剤散布　等
10 収穫	1 収穫	稲刈り、刈取り、摘取り、掘取り等
	2 結束	ラッピング　等
	3 運搬	もみ運搬　等
	4 稲架け	稲架の組立て、取り壊し
	5 脱穀	
	6 その他	稲わら処理、除袋、果実着色、貯蔵前予措等
11 調製	1 乾燥・調製	乾燥、選別　等
	2 もみすり	もみ殻処理、もみ及び玄米の運搬　等
	3 収穫物の水洗い	
	4 外葉除去・切断	
	5 切り枝・整形	
	6 その他	抜き取り、下葉落とし、脇芽つみ、水上げ、泥染め等
12 出荷	1 選別、包装、荷造り作業	計量、選別、袋詰め、箱詰め　等
	2 出荷前の貯蔵作業	予措、貯蔵　等
	3 搬出、出荷作業	
	4 その他	
13 生産管理労働	1 集会出席（打合せ等）	ミーティング　等
	2 技術習得	職員指導、講習会受講　等
	3 簿記記帳	作業記録等の記帳、パソコンデータ入力　等
	4 その他	
14 間接労働	1 自給肥料の生産	
	2 建物、自動車及び農機具の修繕	作業機械の準備、修繕　等
	3 購入資材等の調達	農業購入、肥料購入、燃油購入　等
	4 水利賦役	
	5 その他	

農業ITシステムで用いる農作業の名称に関する個別ガイドライン（試行版）

設計プロセス
- 2nd Step: データ分析

大項目	中項目	小項目（例示）【ユーザーによって拡張（組み替	2014ｙ農場_移植	2014ｙ農場_乾田直播	2014ｙ農場_湛水直播	岡山y農場(移植・直播)	滋賀f農場(移植・乾田直播・湛水直播)	アグリささもと（移植・乾田直播・湛水直播）	耕谷アグリ(移植・直播)	八幡(移植)
1 種子等予措	1 選種	種もみの選種	芒取・塩水選							
	2 浸種									
	3 消毒		種子消毒							
	4 催芽						催芽			
	5 その他	球根予措			種子コーティング		根菜準備：コーティング；ハウス準備			
2 育苗	1 床土・苗床作り									
	2 ﾀﾈ播		育苗箱播種：その他播種			苗箱播種	播種	播種作業		
	3 施肥									
	4 かん水									
	5 換気						（このあたりの項目は細かすぎて使わない？）			
	6 間引き									
	7 防除（土壌消毒）									
	8 移植	ずらし、間引き移植								
	9 接ぎ木									
	10 その他	苗代管理、除草、こもかけ等	育苗管理（複数の中項目を含む概念）；苗出し（ﾊﾟﾚｯﾄ）				育苗器管理：苗出し；ハウス育苗管理；苗箱管理	育苗	ハウス内片づけ・整地；播種準備・播種・育苗	
3 耕起整地	1 耕起・砕土	荒起し、秋田起し、代かき、ハロー等	ブラウ耕；1番耕起；2番耕起；3番耕起；スタブル耕；代かき	ブラウ耕；1番耕起；2番耕起；3番耕起	ブラウ耕；1番耕起；2番耕起；3番耕起；スタブル耕；代かき	耕転；耕転３回目；耕転４回目；ブラウ耕；代掻き	ブラウ耕；荒代（かじり）；代かき；ディスク戻し	バーチカルハロー；ロータリー耕転；２番ロータリー耕転；ブラウ耕；荒代掻き；仕上げ代掻き	ブラウ耕；耕起；代かき	ハロー；耕起；代掻き
	2 畝立て									
	3 温室、ビニールハウスの木ぼ床土作り・床土入れ換え									
	4 その他（ここに入る項目が多い？）	畦塗り	整地（あるいは管理の1か？）；畔塗・修繕	整地；レーザーレベラー；畔塗・修繕	整地；畔塗・修繕	畦塗り	レーザーレベラー；畦塗り	レイザーレベラー；畦塗り	均平作業；　自社圃場畔塗	
4 基肥施肥	1 運搬									
	2 配合									
	3 施肥		肥料散布；堆肥散布	肥料散布；堆肥散布	肥料散布；堆肥散布	堆肥散布；元肥	施肥；土改剤散布	堆肥；元肥；土壌改良材	肥料散布	肥料散布
	4 養液栽培における定植時に行う養液作りや施肥									
	5 その他	秋落ちを防ぐための客土、施肥溝掘り等								
5 ﾀﾈ種・定植（本圃）	1 ﾀﾈ種・覆土	直まき、種まき		鎮圧；播種	播種	直播作業	湛水直播；乾田直播	鎮圧作業；乾田直播；鉄コーティング直播；不耕起播種機		
	2 苗取り・運搬						直播補助；田植え補助		田植え（補助）	
	3 定植（移植）	田植え、植穴（ﾎﾟﾄ）作り等	田植え		田植え	田植	成苗移植（移植）；田植え	田植作業	田植え（オペレーション）；　田植え（直播）	田植え（移植）
	4 [illegible]									

大項目	中項目	小項目（例示）【ユーザーによって拡張（組み替	2014ｙ農場_移植	2014ｙ農場_乾田直播	2014ｙ農場_湛水直播	岡山y農場(移植・直播)	滋賀f農場(移植・乾田直播・湛水直播)	アグリささもと（移植・乾田直播・湛水直播）	耕谷アグリ(移植・直播)	八幡(移植)
	4 追肥の配合							（2の配合と同じ？）		
	5 養液栽培における定植後の養液の濃度管理や補充など									
	6 その他	施肥溝掘り等								
7 除草	1 中耕									
	2 除草剤散布		除草剤散布	出芽前除草剤；出芽後除草剤；除草剤散布	除草剤散布	一発剤散布；除草剤散布	本田除草剤	除草剤		除草剤散布1回目
	3 土入れ・土寄せ									
	4 草刈り・下刈り		手取除草	手取除草	手取除草	草抜き			草刈（トラクター）；本田草取り	草引き；モアリ
	5 マルチング	ビニールマルチ、敷わら、敷草等								
	6 その他						アイガモ管理			
8 防除（雑草以外）	1 化学的防除（農薬散布）								農薬散布	殺虫剤散布
	2 化学的防除（土壌消毒）									
	3 物理的防除	かんれいしゃ砕り等								
	4 生態的防除	組皮はぎ、バンド誘蛾、誘蛾灯管理等								
	5 被害茎等の除去及び焼却									
	6 その他		病害虫防除（中項目より大きい概念）	病害虫防除	病害虫防除	防除１回目；防除２回目		病虫害防除		播種時防除；防除1回目；防除2回目
9 管理	1 圃場（けい畔を含む）及び農道の整備・修復	（いろいろな作業がここに入る可能性があるので、具体的な例示が必要ではないか）	畦畔除草（除草の4か？）	畦畔除草	畦畔除草	畦草刈	整地・補修；保全修理；番堀；畦畔草刈り；畦畔除草剤散布	スライドモアー道路側刈り	畦畔草刈	
	2 施設の整備・修繕	ハウス修繕、機械整備、果樹園の補修、マット交換・排除 保温施設の組立て、取り壊し、暖房器具の組立て、設置、取り除し							農舎清掃	
	3 かん排水	かん水、落水、水管理、落水溝掘り等	溝掘り；水管理；入水；中干し	溝掘り；水管理；入水；中干し；暗渠・明渠；用排水整備	水管理；入水；中干し；暗渠・明渠；用排水整備	入水；中干終了；落水；排水対策	溝切り；本田水管理	排水作業；暗渠排水；弾丸暗渠；暗渠明渠	作溝；水管理	
	4 製材	せん定、支柱立て、誘引、摘芯、摘果等								
	5 生育調節（裂枝以外）	麦踏み、間引き、ホルモン剤処理等								
	6 防風・防霜	こもかけ								
	7 保温・換気									
	8 見回り	生育調査、作柄見回り								
	9 収穫後の株（根）等の除去		枯草焼却	枯草焼却	枯草焼却					
	10 その他	鳥獣害対策								
10 収穫	1 収穫	稲刈り、刈取り、摘取り、掘取り等	稲刈り	稲刈り	稲刈り	収穫；直播・刈取乾燥（調製の1を含む）；刈取乾燥調整		コンバイン刈り取り	稲刈（オペレーション）	収穫

設計プロセス
- 3rd Step: 第1版の提案 -

■ **階層関係の明確化**

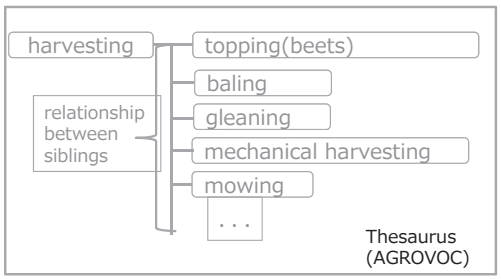
階層関係は人間が理解しやすい形式であるが、いろいろな関係性が混ざることが多い。

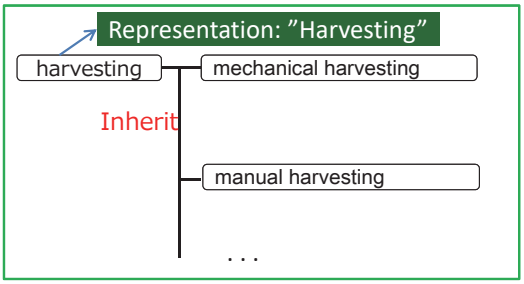
➡ ここでは、isa関係のみに限定する。

■ **同意語の許容**

一つの概念がいろいろな用語で指されることがある（作物の違いなど）。

➡ 概念と表示される語の分離





最上位階層	第二階層（概念）	第二階層（表記）	第3階層（概念）	第3階層（表記）	第4階層（概念）	第4階層（表記）	第5階層（概念）	第5階層（表記）	多重継承
作物生産作業	繁殖作業		種子繁殖作業		採種				
					予措		選種		
							催芽		
					は種				
			栄養繁殖作業		苗取り				
					挿し木				
					接ぎ木				
					取り木				
					株分け				
					分球				
	栄養成長調節作業		草丈制御作業						
			移植		定植		鉢上げ		
					捕植				
			間引き						
			ハードニング（硬化）						
	生殖成長調節作業		外観制御				玉なおし（スイカ）		
			結実制御		受粉				
					無核化				
			開花制御						
			整枝		せん定				作物生産作業＞栄養成長調節作業
					誘引				
					摘果				
					摘葉				
					摘蕾				
					摘心				
環境制御作業	土壌の制御作業		圃場整地作業		耕起		ブラウ耕		
							秋耕		
							春耕		
					砕土		代かき		
							心土破砕		
					均平化				
					鎮圧				
					溝掘り				水分環境制御＞排水
					客土				環境整備作業＞栄養の制御作業
			根圏整地作業		中耕				
					培土				

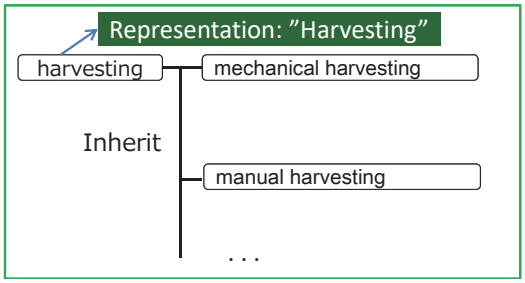
整枝
他の項目は「制御」これだけ性質が異なる。栄養成長調節が目的の場合もあり

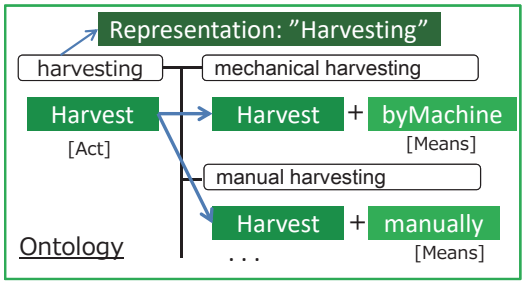
「圃場整地作業」は圃場全体、「根圏整地作業」は作物周辺に限定

設計プロセス

- 4th Step: Introduction of Description Logics -

- Consideration of the structure
  - Discovery of logical structure
  - Reformation of the structure by Description Logics
    - Use of a property for each is-a relation
      - Introduction of a new property
      - Is-a hierarchy of a property value
    - Re-arrangement of classes





設計プロセス

- 5th Step: 専門家による評価と追加-

- 専門家への評価依頼
  - 作物ごとの専門家(農研機構)
  - 農業情報システム開発会社
- Feedback
  - 一部の概念の階層の変更
  - 多くの新規用語
    - 作物固有の用語
    - 地域固有の用語

Version History

Version	Date	Concepts
1.41	2017-01-05	410
1.33	2016-09-23	374
1.31	2016-04-22	367
1.10	2016-02-12	330
1.00	2015-11-02	301
0.94	2015-05-12	185

学んだこと

- 既存語彙のサーベイと批判
  - メリットとデメリットの理解
  - ターゲットを決める
- データ・ドリブンアプローチ
  - 抽象的な議論を避ける
- 両方からの少人数の専門家によるチーム編成
  - 建設的議論
- 核を作り、それを拡張する
  - その後、AIの専門家へ
  - その後、ドメインのより専門家へ
- コミュニケーション大事

CAVOC
Common Agricultural Vocabulary

CVO : 農作物語彙

農研機構 NARO 国立研究開発法人 農業・食品産業技術総合研究機構

NII 大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 National Institute of Informatics

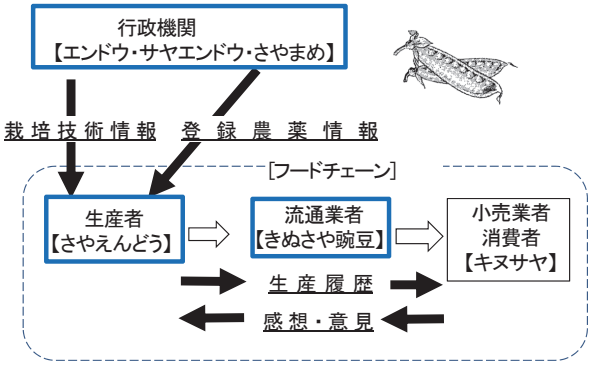
26

農作物語彙体系（CVO）構築_背景・目的

【背景】
・フードチェーンでの食品の安全性確保の重要性の高まり。
　＝＞フードチェーンの各段階で収集管理されたデータを相互利用できるシステムの要望。
・フードチェーンの多様なキープレーヤーは同一農作物のデータに異なる項目名を付けて管理
　＝＞データの相互利用を阻害する要因の一つ。

【目的】
データ項目名が示す農作物概念の共通化を進めシステム間のデータ連携を促進するため**農作物語彙体系**
（Crop Vocabulary; CVO）を構築・公開

【前提条件】
フードチェーンの上流
（生産者＝＞卸売業者）が対象
特に農薬情報の流通に着目して構築



農作物語彙体系（CVO）構築_ 特徴

1. 花き・樹木を除いた農薬情報に含まれる農作物が主な対象

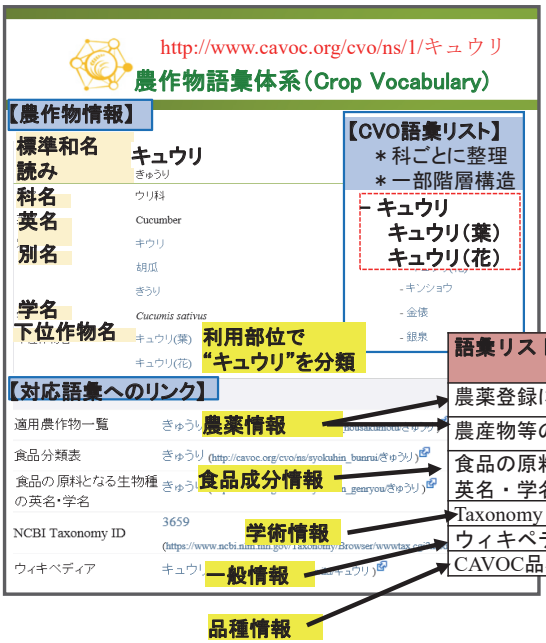
2. 農作物を標準和名、学名、よみ、別名、英名等で表現

3. CVOの農作物名（標準和名）は
　1）植物学的分類に基づいた栽培植物、
　2）利用部位、用途、特性等により栽培植物を分類した集合　を示す。
　　例1；飼料用トウモロコシ、成熟トウモロコシ
　　例2；キュウリ（葉）、キュウリ（花）

4. CVOの農作物語彙は外部語彙リストとマッピングされる。
　農薬・食品成分・学術・品種情報
　＝＞CVOを基盤とした語彙リストの連携が実現

5. 農研機構OB・大学教授による精査（作物・果樹・野菜専門家）
　現在も継続中

農作物語彙体系（CVO）構築_ 公開



CVOは固有の名前空間を持ち
各概念はその上で定義

語彙リスト	発行機関	対応語彙数
農薬登録における適用農作物名	農林水産省	972語
農産物等の食品分類表	厚生労働省	743語
食品の原料となる生物種の英名・学名	文部科学省	274語
Taxonomy Database	米NCBI	972語
ウィキペディア	ウィキメディア財団	768語
CAVOC品種リスト	CAVOC	69語

データ・ドリブンアプローチ

農薬登録における適用農作物名（農林水産省）

農作物名称ガイドライン（農水省，2017）

CVO (tentative)

農作物名称ガイドライン（農水省，2016）

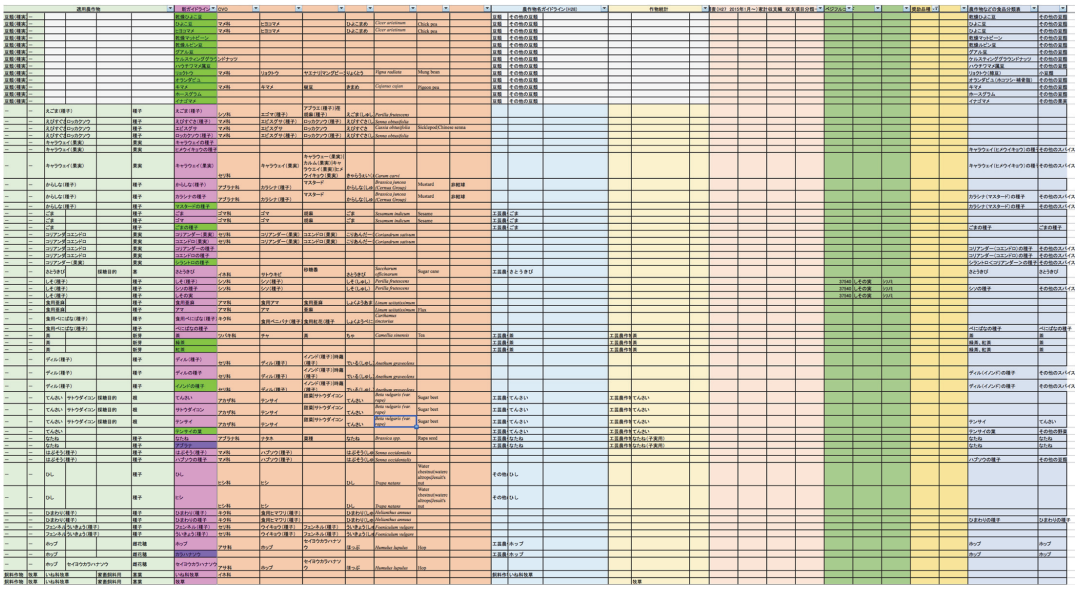
農業統計（農水省）

家計調査（総務省）

ベジブルコード

奨励品種（農水省）

食品の原料となる生物種の英名・学名(文部科学省)



CVO (Crop Vocabulary)

CVO has links to other standards in food industry, agricultural chemical industry.

CVO

農薬登録における適用農作物名
(農林水産省)

食品の原料となる生物種の
英名・学名(文部科学省)

SIP_CVO					適用農作物一覧						厚生労働省		文科省 食品の原料となる生物種の英名・学名				品種登録データベース		
農作物名	科名	学名	英名	属性	大グループ名	中グループ名	作物名	作物名で言まれる形名	カッコ書き	備考(収穫部位)	食品分類表 食品名	食品分類表 食品分類	食品番号	食品名 (()内は生物名)	生物種の英名	学名	登録番号	出願番号	農林水産植物の種類
カリフラワー	アブラナ科	*Brassica oleracea (Botrytis Group)*	Cauliflower		野菜類	–	カリフラワー			花蕾及び花茎	カリフラワー	カリフラワー	06054, 06055	カリフラワー	Cauliflower	*Brassica oleracea Botrytis*			
カリフラワー	アブラナ科	*Brassica oleracea (Botrytis Group)*	Cauliflower		野菜類	–	カリフラワー	はなやさい		花蕾及び花茎									
キャベツ	アブラナ科	*Brassica oleracea (Capitata Group)*	Cabbage		野菜類	–	キャベツ			結球した葉	キャベツ	キャベツ	06061～064, 06333	キャベツ	Cabbage	*Brassica oleracea capitata*			
ブロッコリー	アブラナ科	*Brassica oleracea (Italica Group)*	Broccoli		野菜類	–	ブロッコリー			花蕾及び花茎	ブロッコリー	ブロッコリー	06263, 06264, 06354	ブロッコリー	Broccoli	*Brassica oleracea Italica Group*			
スティックセニョール	アブラナ科	*Brassica oleracea (Italica Group)*		品種	野菜類	–	茎ブロッコリー	スティックセニョール		花蕾及び花茎	スティックセニョール(茎ブロッコリーブロッコリーとケールを交配したもの)	その他のあぶらな科野菜					3871	4594	Brassica oleracea
茎ブロッコリー	アブラナ科	*Brassica oleracea (Italica Group)*		品種群？	野菜類	–	茎ブロッコリー			花蕾及び花茎	スティックセニョール(茎ブロッコリーブロッコリーとケールを交配したもの)	その他のあぶらな科野菜							

農産物等の食品分類表
(厚生労働省)

品種登録デーベース
(農水省)

31

CVO構築

- 作物 =/= 種
- 作物 = 種
  - 種 + 食用/非食用
    - + 利用部位
    - + 成熟/未成熟
    - + 栽培方法
  - + 品種

名前をどうする?

- 最小限の関与
  - それは実際に使われているか？（市場で？農家で？）
  - 適切な名前の付与Give the suitable names
    - 種名（総称）
    - 種名（利用部位）
- 専門家からのフィードバック

共通語彙基盤での経験から

共通語彙基盤(IMI)

- 電子行政分野におけるオープンな利用環境整備に向けた政府のアクションプランの一環
  - 「世界最先端IT国家創造宣言・官民データ活用推進基本計画」[1]
- データに用いる文字や用語を共通化することで情報の共有や活用を円滑に行うための基盤

語彙の設計方針

- 日本語の語彙とする
- 意味や構造は明確になるようにする
- 業界を越えての情報交換を可能にする
- 用途に応じて適した物理形式によるデータ交換・共有が可能とするため、特定の物理形式に依存しないようにする
- 一部の項目を利用することや、項目を追加すること、他の語彙との併用が可能となるようにする
- 継承などにより、コア語彙など基本的な語彙を最大限活用できるようにする
- 既存システムの変更をすることなしに、データ交換の際に利用できるようにする
- 順次導入が可能なようにする

IMI共通語彙とは

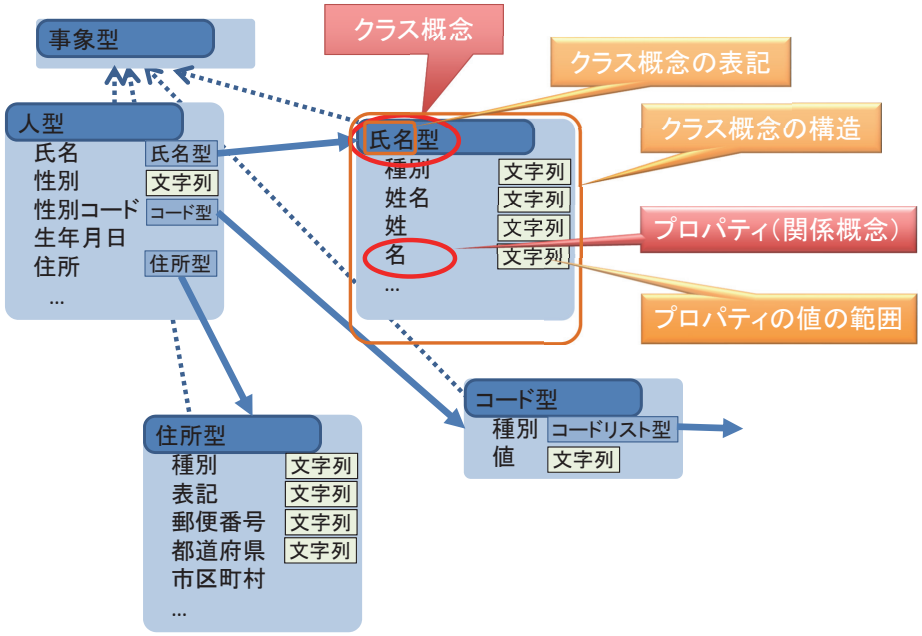
- 構造化概念辞書
  - 概念辞書
    - 概念の表記としての用語
      - 各項目は概念であって用語でない。
  - 構造化辞書
    - 概念は相互につながっていて、その組み合わせ（構造で意味を表現する

用語の種類

用語は「クラス用語」と呼ばれる用語と「プロパティ用語」と呼ばれる用語の2種類からなる。どちらも言葉として意味をもっているので「用語」と呼ぶが、その構造は異なる。

- **クラス用語（クラス概念）**
  - ある事柄を表現する用語。自身の名前とプロパティ用語を構成要素として持つ。
- **プロパティ用語（関係概念）**
  - 事柄の性質や属性を表現する用語および事柄と事柄の関係を表現する用語。自身の名前と使われる事柄（クラス用語）、指し示すもの（クラス用語あるいは値の型）から構成される。

IMI共通語彙の構造



クラス用語

- 項目名、説明、継承、及び、プロパティ用語の集合で構成される
- 各プロパティに対して、カーディナリティ(出現数)とクラス用語における意味の説明をもつ　※「クラス用語」とは、項目名のことではないことに注意

人

識別子:
ic:人型

継承:
ic:実体型

説明:
人の情報を表現するためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:ID	ID	ic:ID型	0..n	人に割り振られたIDを記述するためのプロパティ用語
ic:氏名	氏名	ic:氏名型	0..n	氏名を記述するためのプロパティ用語
ic:性別	性別	xsd:string	0..1	性別の表記を記述するためのプロパティ用語
ic:性別コード	性別コード	ic:コード型	0..1	性別コードを記述するためのプロパティ用語
ic:生年月日	生年月日	ic:日付型	0..1	生年月日を記述するためのプロパティ用語
ic:死亡年月日	死亡年月日	ic:日付型	0..1	死亡年月日を記述するためのプロパティ用語
ic:住所	住所	ic:住所型	0..n	現住所を記述するためのプロパティ用語
ic:本籍	本籍	ic:住所型	0..1	本籍を記述するためのプロパティ用語
ic:連絡先	連絡先	ic:連絡先	0..n	連絡先を記述するためのプロパティ用語

プロパティ用語

- クラス用語を説明するためのクラス用語や値を表すための用語
- 項目名、データ型、及び説明から構成される
- データ型はクラス用語の型かXMLスキーマで定義される基本型をとる

項目名	データ型	説明	定義域
氏名	ic:氏名型	人の氏名	人

- クラス用語の子要素として使用される

人

識別子:
ic:人型

継承:
ic:実体型

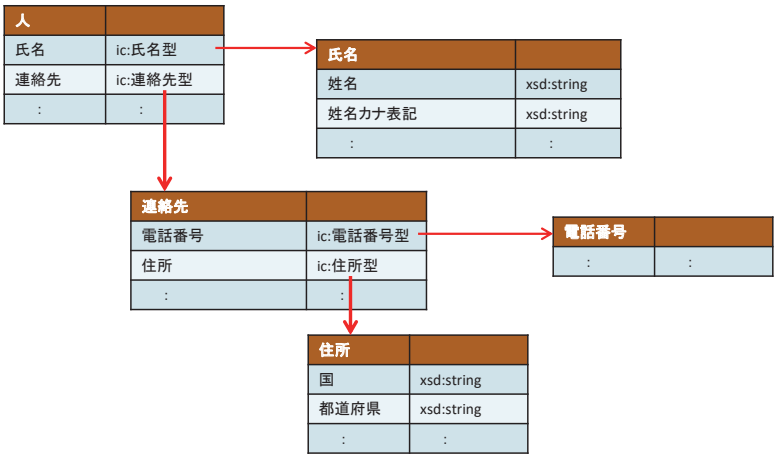
説明:
人の情報を表現するためのクラス用語

プロパティ:

識別子	項目名	値型	回数	説明
ic:ID	ID	ic:ID型	0..n	人に割り振られたIDを記述するためのプロパティ用語
ic:氏名	氏名	ic:氏名型	0..n	氏名を記述するためのプロパティ用語
ic:性別	性別	xsd:string	0..1	性別の表記を記述するためのプロパティ用語
ic:性別コード	性別コード	ic:コード	0..1	性別コードを記述するためのプロパティ用語

用語の構造

- プロパティ用語のデータ型としてクラス用語を指定することができる。その場合、データ型であるクラス用語がまたプロパティ用語をもつので、図のような概念的な構造をもつ



クラス用語の継承

- 他のクラス用語を継承して新たなクラス用語を作成することができる
- 継承したクラス用語は、継承元のすべてのプロパティをもつ
- 継承したクラス用語は継承元にないプロパティを追加することができる
- 継承したクラス用語は継承元のクラス用語としても使用できる。例えば、データ型が「組織」となっているプロパティには、「業務組織」の値を指定することができる。

共通語彙基盤 コア語彙2.4.1

To the extent possible under law, Ministry of Economy, Trade and Industry and Information-technology Promotion Agency, Japan has waived all copyright and related or neighboring rights to Infrastructure for Multilayer Interoperability (IMI) Core Vocabulary 2.4.1. This work is published from: Japan .

赤い取り消し線が引かれた用語およびプロパティは将来の版では使用できなくなるため、使用を推奨しません。

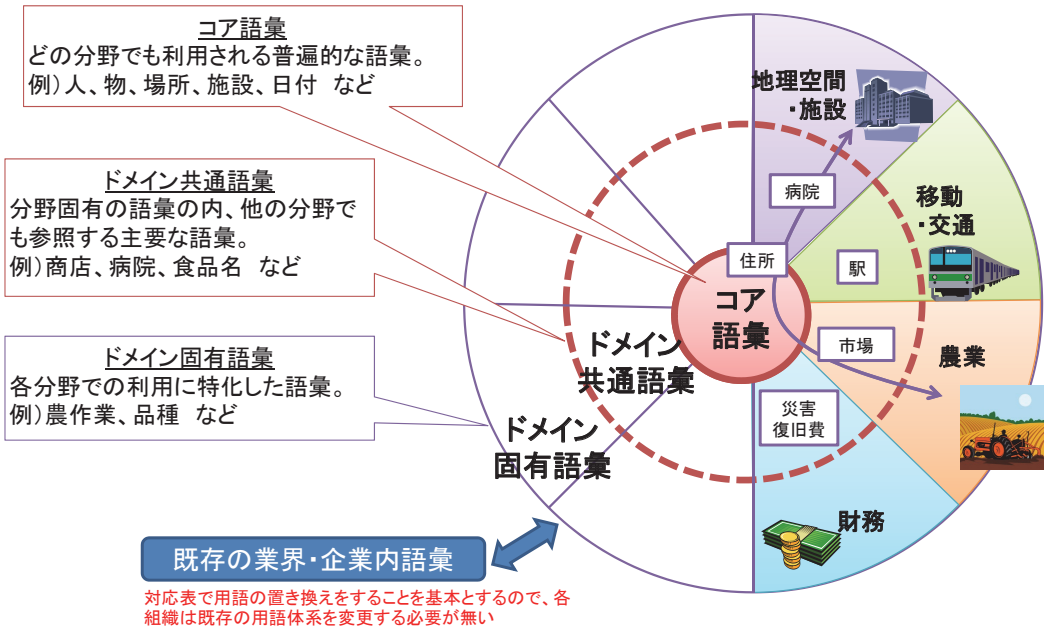
=>プロパティ用語一覧

クラス用語一覧

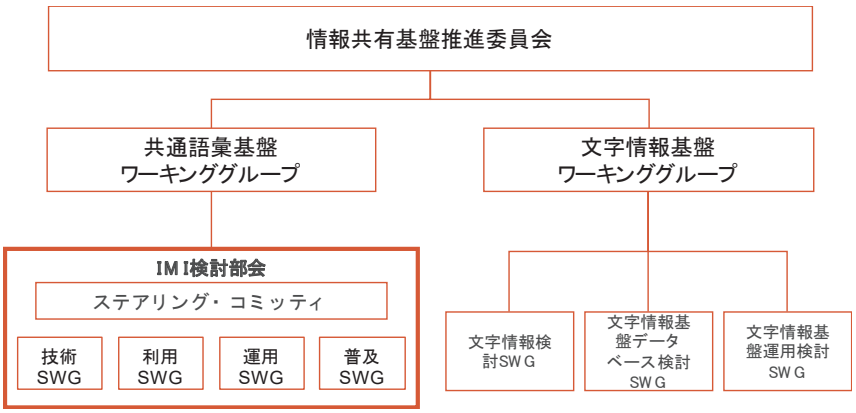
人型	名称型	氏名型	住所型
連絡先型	電話番号型	組織型	業務組織型
法人型	人数型	構成員型	組織関連型
場所型	座標型	ID型	ID体系型
製品型	製品個品型	地物型	土地型
施設型	施設関連型	建物型	駐車場型
設備型	イベント型	活動型	関与型
~~測定値型~~	数量型	容量型	面積型
重量型	長さ型	~~時間型~~	金額型
価格型	実体型	状況型	日時型
日付型	期間型	期間スケジュール型	イベントスケジュール型
定期スケジュール型	詳細スケジュール型	詳細スケジュール規則型	コード型
単位コード型	コードリスト型	アクセス型	アクセス区間型
概念型	事物型	参照型	記述型
サービス型	文書型	対象型	制約型
コード制約型	範囲制約型	期間制約型	

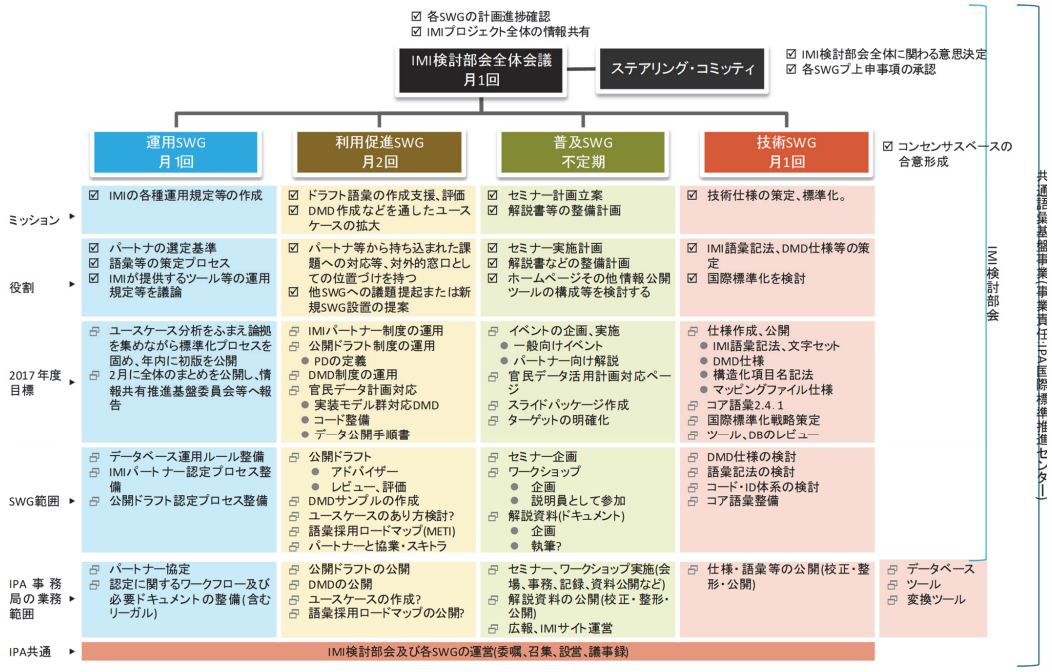
IMI共通語彙基盤の構造

- 情報交換のための用語連携インタフェースは、社会全体で使われる用語であるコア語彙、共通語彙及びドメイン固有語彙から構成される。



語彙基盤のマネージメント





いくつかのタスク

- プロジェクトの維持
- コア語彙のメンテナンス
- ドメイン語彙、応用語彙の構築

コア語彙のメンテナンス

- データの提供
  - 多様フォーマット
    - 共通語彙記法
    - RDF (JSON-LD, turtle)
    - XML
    - 構造化項目名
    - HTML
- バージョンアップ
  - メジャー／マイナー／マイナー・マイナー
  - バージョンアップ手順の制式化
    - 提案 – 意見募集 – 決定 – 周知
  - バージョンアップ手順のシステム化

ドメイン語彙、応用語彙の構築

- ドメイン語彙の構築
  - あるドメインの共通語彙。誰かがこれをコア語彙の同様に維持。
  - 今の所、未構築
- 応用語彙
  - 特定の用途のための語彙。
  - 誰でも提案できる。ただし、IMIパートナーとなることでIMIサイト上にPD(Public Draft)語彙として置くことができる。
    - パートナーとIMIとの協同作業で語彙を提案

IMIパートナー一覧

2017年5月現在のIMIパートナーです(五十音順)。

株式会社アスコエパートナーズ	ASUKOE 私の声で明日を変える
一般社団法人XBRL Japan	XBRL JP
一般社団法人オープン・コーポレイツ・ジャパン	OCJ OpenCorporates Japan
公益財団法人　九州ヒューマンメディア創造センター	HMC
Code for Tokyo	
バイザー株式会社	VISOR
ビッグデータ＆オープンデータ・イニシアティブ九州	BODIK Since 2013
一般社団法人 ユニバーサルメニュー普及協会	UM International Universal Menu Association

まとめ
オントロジー・語彙の構築と維持

- 構築する
  - オントロジー・語彙の専門家とそのドメインの専門家との協調プロセスをいかに作るか
- 維持する
  - 実は大変。
  - オントロジー・語彙は変化する。それをどう維持するか。
  - 体制づくり

２．「辞書記述の精緻化と
シソーラス・コーパスの活用」

2017/12/19

辞書記述の精緻化と
シソーラス・コーパスの活用

柏野　和佳子

大学共同利用機関法人 人間文化研究機構
国立国語研究所
National Institute for Japanese Language and Linguistics
NINJAL

辞書記述の精緻化

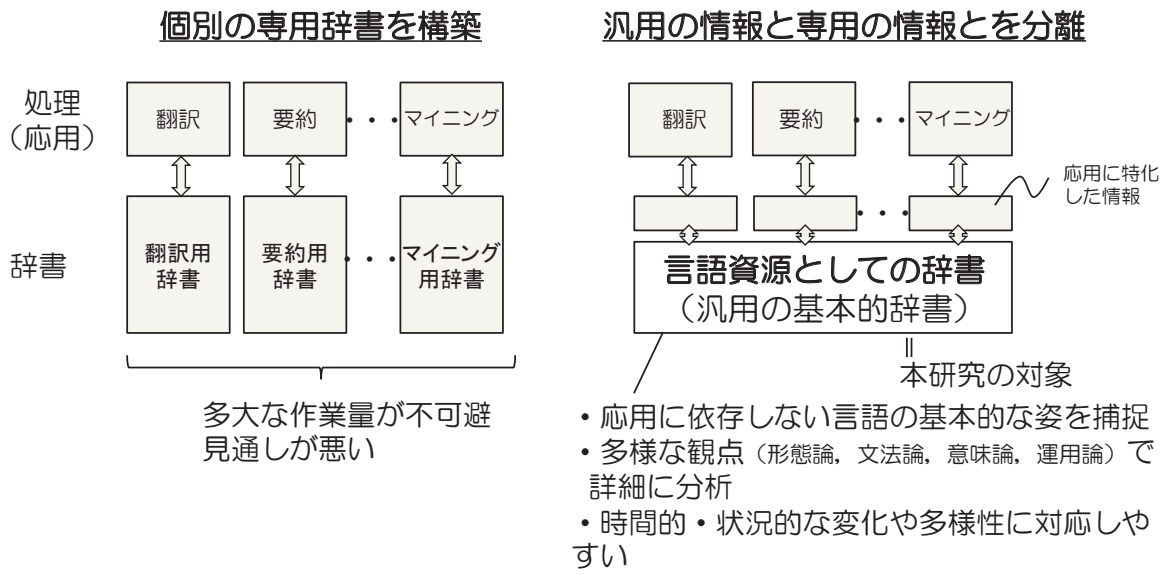
柏野和佳子 (2016)コーパスに基づく辞書記述の精緻化の研究.東京工業大学総合理工学研究科博士論文.

2

背景

- 「言語資源としての辞書」の構築の必要性

  汎用の基本的辞書の整備が応用上も言語学上も重要



3

これまでの問題点と本論文の目的

国語辞典　電子化したものを様々な言語処理に適用した例は多い

計算機用辞書　例：IPAL(1987～1996)，EDR(1993)，NTT日本語語彙体系(1997)

問題点

編集者の知識や内省への依存度が高い

① 国語辞典　→　母語話者利用を暗黙の前提としているため重要な情報が不足

② 計算機用辞書　→　これまで以上に精度の高い解析・生成処理を実現するためにはまだ情報が不足

本論文の目的

(1) 辞書記述の骨格となる見出し語の選定・語義区分・多義語の構造の記述を客観化・明示化して行う方法を提案する。

(2) 解析・生成処理の精度向上のために，コーパスに基づく新たな辞書情報の記述方法を提案する。

(3) 均質的・効率的な辞書構築を作業ベースで行えるように辞書記述を手順化する。

4

辞書に記載すべき情報と本論文で着目する情報

見出し語
見出し語レベルの≪メタ情報≫**頻度情報**，≪形態情報≫ 表記，
≪統語情報≫語種，品詞，活用型など，≪意味情報≫**多義の構造**

語義1
≪意味情報≫語義の記述1

語義1レベルの≪メタ情報≫
頻度情報など・・・

語義1レベルの≪形態情報≫
表記など・・・

語義1レベルの≪統語情報≫
コロケーションなど・・・

語義1レベルの≪意味情報≫
使用域など・・・

語義2　・・・
≪意味情報≫語義の記述2

5

辞書に記載すべき情報と本論文で着目する情報

用例を収集し，「見出し語」を選ぶ。

見出し語
見出し語レベルの≪メタ情報≫**頻度情報**，≪形態情報≫ 表記，
≪統語情報≫語種，品詞，活用型など，≪意味情報≫**多義の構造**

用例を分析し，「多義の構造」を記述する。

語義1
≪意味情報≫語義の記述1

用例を分析し，「区分」する。

語義1レベルの≪メタ情報≫
頻度情報など・・・

語義1レベルの≪形態情報≫
表記など・・・

用例を分析し，《各情報》を記述する。

語義1レベルの≪統語情報≫
コロケーションなど・・・

語義1レベルの≪意味情報≫
使用域など・・・

語義2　・・・
≪意味情報≫語義の記述2

6

本論文の着眼点と 提案する辞書記述の要件

(1) 辞書記述の骨格における客観化・明示化の提案

要件1-1 見出し語の選定基準

→ 対象とする語の取捨選択の主観への依存性を低減

要件1-2 語義の区分

→ 区分の主観への依存性を低減，文法・用法による多義性の解消が可能

要件1-3 多義語の意味関係の構造

→ 多義性の仕組みに根差した多義性の解消が可能

(2) 辞書に記述すべき新規な情報の提案

要件2-1 多義性をもつコロケーション（名詞句＋格助詞＋述語）の意味の分離

→ コロケーションへの修飾情報の記述により多義性の解消が可能

要件2-2 語の使用域と位相

→ 文脈に相応しい語選択が可能

要件2-3 語の使用頻度の経年変化

→ 語の流行り廃りなどの情報により文脈や目的に応じた語選択が可能

7

コーパスに基づくことによる問題解決の見通し

(1)客観的・明示的になる

頻度情報や実例に基づき，辞書情報の抽出と記述とが客観的・明示的に行えるようになる。

(2)網羅的になる

生きた用例を可能な限り網羅的に収集できるようになる。

(3)具体的になる

資料の幅が広がり，語の多様性を具体的に捉えることができる。ジャンルあるいはレジスタなどと呼ばれる，言葉の使用域が異なる場合，それを具体的に辞書記述に反映できる(例えば，話し言葉的か，書き言葉的か，フォーマルか，インフォーマルか)。

8

要件1-3 多義語の意味関係の構造①

用例より区分間の意味関係を認定し，明示化

本論文での適用対象：「名詞の下位区分間の意味関係」　（『IPAL名詞辞書』）

1 区分間の関係を，**メタファー(隠喩)，メトニミー(換喩)，シネクドキー(提喩)**のいずれで捉えられるか判定し，記号と区分番号とを用いて関係を記述する。

2 拡張の方向性は，歴史的な派生順序にできるだけ沿うように与える。（『日本国語大辞典』(小学館，1976年)を適宜参照）

3 可能な限り単純明瞭に記述するために，3タイプの比喩の融合と捉えられる場合も，いずれのタイプにできるだけ特定して記述する。

4 比喩によって捉えきれない多義については，できるだけパターン化した文章でその多義がどう捉えられるかを記述する。

Wakako Kuwahata, Minako Hasimoto, Fumihiro Aoyama (1996) Senses of Polysemous Nouns : Building a Computational Lexicon of Basic Japanese Nouns. *Proc. of The 16th International Conference on Computational Linguistics (COLING1996)*, pp.1082-1085.
柏野和佳子，本多啓（1998）IPAL名詞辞書における多義構造の記述．『情報処理学会論文誌』 39(9)，pp.2603-2612.

9

比喩に基づく意味の拡張関係

– シネクドキー（提喩）　カテゴリ全体とその一部という包摂関係によって意味を拡張する
  - 着物：着物 (衣服) を揃える ⊃ 着物 (和服) を着付ける
  - 酒：熱燗の酒 (日本酒) ⊂ 世界各国の酒 (アルコール飲料)
  - スピード：スピードが遅くなる ⊃＋ スピードがある
  - 広さ：部屋の広さに圧倒された ＋⊂ グラウンドの広さをはかる

– メタファー（隠喩）　類似性によって意味を拡張する
  - 肩：人の肩 ＝＞ 洋服の肩，…＞ 山の肩，文字の肩
  - 皮：動物の皮 ＝＞ みかんの皮，まんじゅうの皮，…＞化けの皮

– メトニミー（換喩）　密接な関係にあるものや属性などの隣接関係によって意味を拡張する
  - 肩：肩がこる [身体部位] ー＞ [機能] いい肩をしている
  - 鍋：鍋で煮る [容器]ー＞ [中身] 鍋を食べる
  - 茶：茶を摘む [材料] ー＞[加工品]茶を入れる
    茶を入れる[手段] ー＞[行為]お茶にする

（柏野和佳子，本多啓(1998) 多義構造を辞典に書く．『日本語学』明治書院，17(14)，pp.54-63.
ほか，ウルマン1962，レイコフ・ジョンソン1986，国広1982，山梨1995）

10

要件1-3 多義語の意味関係の構造②

提案する記述の具体例1

「のど」

[区分01]　管の形をしている口の奥の部分。
　例：のどに魚の骨がひっかかってしまった。
[区分02]　頭と肩をつなぐ身体部分の前面。
　例：喉に湿布をする。
[区分03]　歌う声。
　例：彼女の 自慢の喉を聞かせてもらいましょう。
[区分04]　本の綴じ目の部分。
　例：本ののどがこわれる。

＜関係＞ 01［物］ ー＞ ［側面］02。　：メトニミー
　　01［手段］ ー＞ ［行為］03。　：メトニミー
　　02 ・・・＞ 04。　：メタファー

11

提案する記述の具体例2

例：「あし【足，脚】」

[区分01]　動物の胴に付属し、歩行や体を支える機能を果たす身体部分。
　例：彼女は足が長い。
[区分02]　「あし01 」の一部で、くるぶしより下の部分。
　例：彼は足が大きい。
[区分03]　歩いたり、走ったりすること。　例：彼は足が速い。
[区分04]　移動の手段となるもの。　例：台風で市民の足が乱れる。
[区分05]　椅子などの下部で、台の部分や上にある物を支える役割を果たす部分。
　例： 机の脚が折れる。

01 ＞＞ 02。　01［身体部位］ ー＞ ［機能］03 ...＞ 04。　01 ...＞ 05。

[区分02]は[区分01]と全体部分関係のメトニミー
[区分03]は[区分01]と［身体部位］と［機能］の関係のメトニミー
[区分04]は，その[区分03]からの機能の類似性によるメタファー
[区分05]は[区分01]からの形態的な類似性によるメタファー

12

多義構造の記述の試行

辞書に収録する多義語の意味的関係を記述 →　全体の約 86% の記述が可能

- 多義語560語の区分間の数　1,364
- そのうち比喩的関係として捉えて記述できた数　1,174(86%)

表4.14　各カテゴリにおける意味的関係の記述数

カテゴリ	メタファー		メトニミー			シネクドキー				(他)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	
サ変	53	18	3		7	34				8
形容詞転成	67				1	10		4	3	
形容動詞	12	1			8	21		2		12
相対	75	42			38	24	1			49
副詞	19	2			5	8		1		3
述語用法2	23	27		1	12	21		15		14
被修飾2	24	13		1	7	52		1		19
動詞転成	5	5			5	8				
植物		22	11	2	6	4				
食べ物	6	4			16	6	2			4
身体	25	56	10	4	47	2				17
動物	1	14	3	2	9	9	2			11
場所	5	9	2	1	5	19	1			3
建物・家具	1	12	1	1	6	3				3
乗り物		2		2	3	1				
洋装品	4	11			4	3				3
仕事・役割	15	9			12	20				22
自然物	3	15	2	1	18	5				12
自然現象	5	8			2	5		1		6
組織・団体	2	2			4	7				4
小計	345	272	32	15	215	262	6	24	3	190
合計	617		262			295				190

合計 1,174

13

多義性をもつコロケーション

表5.1　重複記載の4パターン

a. 異区同素(異なる区分で同じ素性と重複)　3,608例
(例)　[区分01] ACT　(敵軍の)攻撃が始まる
　　[区分02] ACT　(議長への)攻撃が始まる

b. 異区異素(異なる区分で異なる素性と重複)　2,260例
(例)　[区分01] MEA　(大豆の)収穫が多い
　　[区分02] GRA　(会議の)収穫が多い。

c. 同区同素(同じ区分で同じ素性と重複)　821例
(例)　[区分01] CON　自転車(のサドル)ガ高い
　　区分01]　CON　自転車(の価格)ガ高い

d. 同区異素(同じ区分で異なる素性と重複)　934例
(例)　[区分01] CON　(シャワーデ)汗ヲ流す
　　[区分01] LIQ　(額ニ)汗ヲ流す

合計　7,623例

14

コロケーション重複の要因

要因(A)-1　多義である名詞が同じ素性を持つため，素性に特有の同じ述語と結び付く。

例えば，〈LOC〉(場所)であれば「行く，戻る，広い，狭い」，〈ACT〉(行為)であれば，「始まる，終わる，止まる，早い，遅い」といったものが素性に特有の述語である。

要因(A)-2　述語の抽象度が高いため，同じ述語と結び付く。

例えば，「ある」「ない」「する」「なる」といった述語は抽象度が高いため，見出し語が多義であるとき，それぞれに結び付きやすく，その結果，重複する数も多くなる。

要因(B)-1　名詞と述語の組合せであるコロケーション自体が比喩的意味を持つため，同じように結び付く。

区分間にメタファーの関係があると，コロケーションが重複しやすい。

要因(C)-1　多義である述語が同じ名詞と結び付く。

述語の指す意味はまったく異なるのに，それが異なる述語ではなく，同じ1つの述語の多義によるものであるため，結果として重複が生じる。

15

重複記載を区別する違い

表5.2　重複記載を区別する違いの5種類

a. 他の名詞句が立つ格助詞が違う(他格)	2,421例
(例) 02:CON～ガ～ニお茶ヲ入れる	
02:PRO～ガお茶ヲ入れる	
b. 他の名詞句が違う(他名)	1,145例
(例) 02:HUM(控え室)ニ家族ガいる	
02:ROL(彼女)ニ家族ガいる	
c. 見出し語の先行句が違う(先行)	4,888例
(例) 01:ACT～ガ(原野の)開拓ヲ行う	
02:ACT～ガ(技術の／新市場の)開拓ヲ行う	
d. 見出し語の後行句が違う(後行)	91例
(例) 01:POTのど(の調子)ガいい	
03:FORのどガいい(＝歌う声がいい)	
e. 見出し語が立つ格助詞が違う(格助)	1,794例
(例) 01:INT～ガ舞台ヲ／カラ出る	
01:SPA～ガ舞台ニ出る	

16

要件2-1 多義性をもつコロケーションの意味の分離①

コロケーションへの修飾情報を記述し，コロケーションのもつ多義性を解消

本論文での適用対象：「多義性をもつ名詞のコロケーション」　（『IPAL名詞辞書』）

1 コーパスからコロケーション（名詞句＋格助詞＋述語）を抽出する。

2 「～の」の形で名詞句を連体修飾する用例を詳細に記述する。

3 形容詞や「～な」の形で名詞句を連体修飾する用例を詳細に記述する。

4 テイル形に関するアスペクト情報や副詞表現についての情報抽出が可能であれば抽出し，それも詳細に記述する。

例：「骨が折れている」→慣用句（苦労という意味）の用法ではない。

桑畑(柏野)和佳子，橋本三奈子，青山文啓 (1998) IPAL名詞辞書による多義性解消のためのコロケーションの分析．『情報処理学会論文誌』39(6), pp.1925-1934.

17

要件2-1 多義性をもつコロケーションの意味の分離②

提案する記述の具体例

「いろ【色】」

[区分01] GRA(口紅の／…)色が濃い
〈対象〉[CON｜TIM｜…]信号の―，表紙の―，服の―，炎の―，花の―，醤油の―，夕暮れの―，緑の―，葉の―。

[区分03] APPあせりの／…色が濃い
〈部位〉顔の―，目の―。
〈具体化(心情を表す名詞)〉不安の―，警戒の―，反省の―，焦慮の―，憂慮の―，苦悩の―，傷心の―，いらだちの―。

[区分05] GRA敗戦の／…色が濃い
〈具体化(状況を表す名詞)〉秋の―，自己防護の―，敗北の―，崩壊の―。

[区分07] GRAハト派の／…色が濃い
〈具体化(思想を持つ団体やその考え方・行動を表す名詞)〉タカ派の―，保守の―，過激なイスラム原理主義の―，徒弟制度の―，仏教国の―。

18

辞書情報による重複記載の多義性解消の確認

辞書に収録された多義性のある用例を記述　→　全体の約 84.4% の多義性が解消

- 重複するコロケーション(延べ10,415例，異なり4,578例)の重複対　7,623例
- そのうち記述により多義性の解消ができた数　6,436例(84.4%)

表5.2重複記載を区別する違いの5種類

タイプ	重複記載を区別する違い5種類					タイプ	重複記載を区別する違い5種類				
	他格 他名	先行	後行	格助	出現数(%)		他格 他名	先行	後行	格助	出現数(%)
1	−	−	−	−	1,187(15.6)	13	+	+	−	−	594(7.8)
2	−	−	−	+	77(1.0)	14	+	+	−	+	913(12.0)
3	−	−	+	−	20(0.3)	15	+	+	+	−	2(0.0)
4	−	−	+	+	2(0.0)	16	+	+	+	+	3(0.0)
5	−	+	−	−	2,594(34.0)	17	×	−	−	−	524(6.9)
6	−	+	−	+	162(2.1)	18	×	−	−	+	15(0.2)
7	−	+	+	−	13(0.2)	19	×	−	+	−	0(0.0)
8	−	+	+	+	2(0.0)	20	×	−	+	+	1(0.0)
9	+	−	−	−	330(4.3)	21	×	+	−	−	551(7.2)
10	+	−	−	+	533(7.0)	22	×	+	−	+	52(0.7)
11	+	−	+	−	12(0.2)	23	×	+	+	−	2(0.0)
12	+	−	+	+	34〈0.4)	24	×	+	+	+	0(0.0)
計											7,623(100.0)

これ以外の合計 6,436

19

要件2-3 語の使用頻度の経年変化①

頻度情報により経年変化を明示化

本論文での適用対象：「外来語」

1 経年のデータが蓄積している新聞データベース等を時間的使用推移の頻度情報を得て，使用推移のグラフを作成する。

2 1で作成したグラフを基に，使用推移が次のいずれかの型であるかを判断する。

平行型　平ら：出現期間内をならすと平ら
増加型　右上がり：出現期間内をならすと増加
減少型　右下がり：出現期間内をならすと減少
山：3年以上の増加後3年以上の減少
山の後平ら：増加，減少，平らが各3年以上で連続
凹凸型　突出：特定の年のみ増加
上記以外

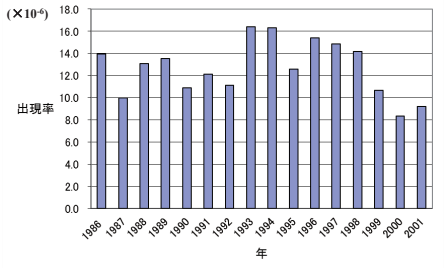
3 2で判断した推移の型を辞書に明記するとともに，その判断の元となった頻度情報を記載する。頻度情報は，人が閲覧する用途では，閲覧用プログラムが必要に応じてグラフとして表示してもよい。

柏野和佳子，山口昌也，桐生りか(2005) 新聞記事データベースを利用した外来語の出現率の推移調査．『自然言語処理』12(4) pp. 97-116

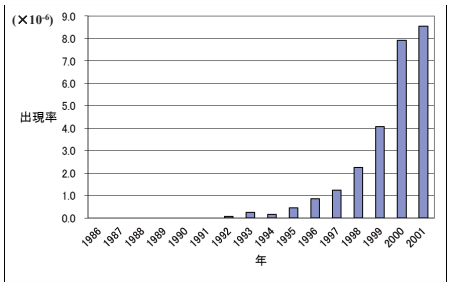
20

要件2-3 語の使用頻度の経年変化②

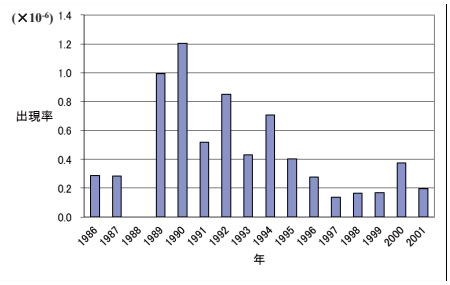
例1：シェア　平行型



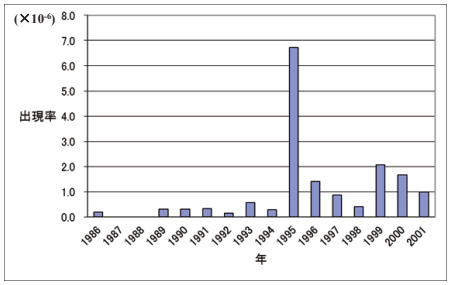
例2：バリアフリー　増加型



例3：フレックスタイム 減少型



例4：ライフライン 凹凸型



21

分類型の安定性の確認①

外来語156語について，新聞二紙間の頻度推移4分類の型の安定性を調査
→　型分類で67.3%が一致，相関あり(相関関数0.5以上)も同じく67.3%

- 読売新聞，毎日新聞の1991年から2004年までの14年分の記事データを使用。
- 156語について，両紙の頻度間の相関を調査。平均0.57。0.5以上が105語(67.3%)。0.5以上を相関あり，とすれば，全体の67.3%に相当する語が相関あり。

表5.20 頻度推移類型の分類結果

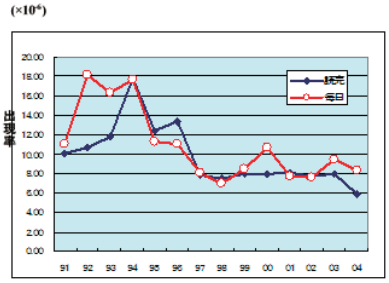
分類	両紙一致の語数	両紙不一致の語数	
	読売・毎日	読売	毎日
平行型	9	1	2
増加型	20	9	7
減少型	56	25	16
凹凸型	20	16	26
合計（語）	105	51	51

67.3%

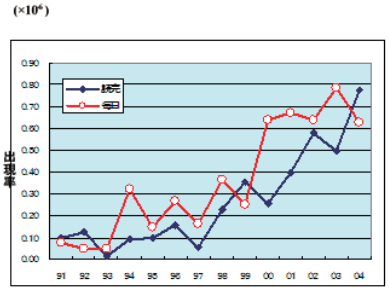
22

分類型の安定性の確認②

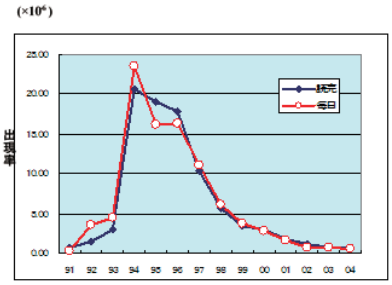
例1：ビジョン　平行型



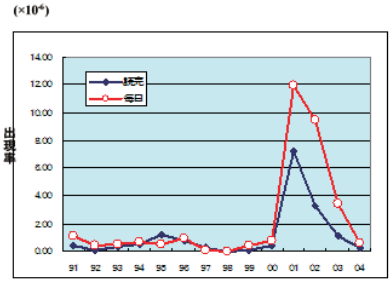
例2：スキル　増加型



例3：マルチメディア　減少型



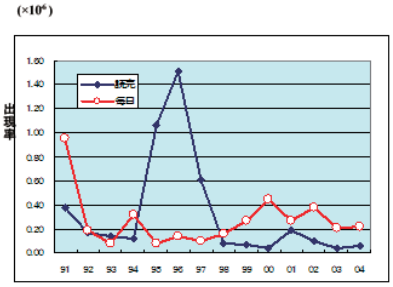
例4：セーフガード　凹凸型



23

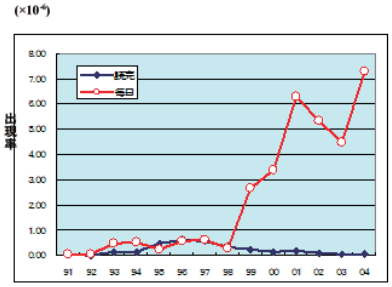
分類型の安定性の確認③

例7：ポートフォリオ　不一致



「ニッセイ投資顧問チーフポートフォリオマネージャー」の連載コラム

例8：インタラクティブ　不一致



毎日新聞の新しいホームページ「毎日インタラクティブ」

24

本論文の到達点

- 辞書記述の骨格における客観化・明示化の提案
  - コーパスの頻度情報に基づく見出し語の選定方法とその有効性を示した。
  - 名詞に着目し，名詞を下位区分する際の客観的・明示的な区分方法とその有効性を示した。
  - 名詞の多義語に着目し，その意味関係の記述方法とその有効性を示した。
- 辞書に記述すべき新規な情報の提案
  - 用例を網羅的に収集した場合に生じるコロケーションの多義性の問題に着目し，それを解決するための記述方法とその有効性を示した。
  - 「古風な語」に着目し，使用域・位相情報の記述方法とその有効性を示した。
  - 「外来語」に着目し，使用頻度の経年変化に基づく辞書記述方法とその有効性を示した。

期待される効果
- 多義性解消の精度向上，語選択・生成処理の精度向上。
- 多義の関係，豊富なコロケーション，文体・位相情報による人の参照性の向上。

25

残された課題

- 辞書記述の精緻化に関する課題
  - 低頻度語の扱い
  - コーパスの規模と見出し語・用例の採集効果の関係の統計的な解析
  - コロケーションの結び付きの強度の明示化
  - 文体・位相情報の体系的な記述方法の提案
  - 意味・用法の自動分類
- 大規模辞書の構築
  - 数万語以上の規模の実用的辞書の構築
    - 本論文では，古風な語，外来語などといった限られた対象について議論した。これを足掛かりとし辞書を構築する。

26

国語辞典・シソーラス

27

技と工夫　その１

- 似たものを使って説明する。

パセリ
畑に作る二年草。葉は**ニンジンの葉のように**細かく切れている。
かおりが強く、洋食のつまにする。オランダゼリ。
（新明解国語辞典）

28

技と工夫　その２

- 意味，用法を似たもの、近いものをまとめて説明する。

VS

- 意味，用法を細く分けて並べて説明する。

とる【取る・採る・執る・捕る】：

『岩波国語辞典』８個
（①～⑧の下に、小項目が３個、８個、３個、２個、３個、３個、１個、１個、計２４個あり）

『明鏡国語辞典』１～５９まで、59個

学習用では・・・
『新レインボー小学国語辞典』１～１４まで、14個

29

技と工夫　その３

- 意味を主観的に（自分の感じ方や考え方で）説明する。

VS

- 意味を客観的に（ものごとの見え方や考え方で）説明する。

こい【恋】
特定の異性に深い愛情を抱き、その存在が身近に感じられるときは、他のすべてを犠牲にしても惜しくないほどの満足感・充足感に酔って心が高揚する一方、破局を恐れての不安と焦躁(シヨウソウ)に駆られる心的状態。「一に落ちる／一は盲目／一は思案の外（ホカ）〔=恋する人の気持や行動は常識や理性では割り切れないものだ〕／老いらくの一」　（新明解国語辞典）

こい【恋】
異性に愛情を寄せること、その心。恋愛。「一は盲目」（恋をすると無分別になる）▽本来は、（異性に限らず）その対象にどうしようもないほど引きつけられ、しかも、満たされず苦しくつらい気持を言う。「一に焼けて死ぬ虫になったって、思いはとげて見せるぞ」のような用例から見て取れるとおり、心の活動（やその内容）を言う「思い」とは区別される。「一は楽し、野辺の花よ」のような言い方は、1910年代ごろからのもの。→こう（恋）　（岩波国語辞典）

30

技と工夫　その４

- 言葉を記述的に（あるがまま）説明する。
- 新語・新用法を積極的にのせる。

VS

- 言葉を規範（きはん）的に（お手本として）説明する。
- 新語・新用法は慎重にのせる。

さわり【触り】
④曲などの出だしの部分
〔参考〕④は最近の用法。　（新選国語辞典）

『三省堂国語辞典』は改版時に新語を中心に約4,000語が増補されている。
六版（2008年）「イケメン，がっつり，逆ギレ，激～，～じゃないですか，どん引き，パニクる，真逆，むずい，やさぐれる，ロン毛」など。
七版（2014年）「いらっと，上から目線，オソロ，空気感，恋バナ，ニーハイ，まるっと，スイッチが入る，パンダ目、ほうれい線，変顔」など。

31

技と工夫　その５

- 類義語だけで説明しないようにする。

つきあう【付き合う】
交際する。

こうさい【交際】
つきあうこと。　（多くの辞典）

つきあう【付き合う】
①よその人となかよく話したり，何かをいっしょにしたり，助け合ったりして，したしい関係をつくる。
②恋人として交際する。
（以下，略）

こうさい【交際】
〔人間どうしの〕つきあい。つきあうこと。まじわり。
（三省堂国語辞典）

32

技と工夫　その6

- 語釈にあう用例を入れる。

岩国《進む》

(1)(ｴ) 仕事がはかどる。	「工事が着着と—（=完成に向かう）」
(3)(ｳ) 勢いが増す。⇔衰える。	「食が—」（食欲が盛んになる） 「病勢が—」（病気が悪化する）

明鏡《進む》

③物事が（計画した通りに）進行する。	「仕事が予定通りに—」 「工事が—」 「筆が—」 「大いに食が—」
④物事の状態・程度がはなはだしくなる。進行する。	「過疎化［相互理解・世代交代・活字離れ］が—」 「病状［近視］が—」

33

技と工夫　その7

- 見出し語を一つにする，あるいは分ける。

岩波国語辞典　第七版新版　岩波書店

こ-える【越える】

①手前から向こうに、その上を過ぎて移る。「山を–」「国境を–」「2メートルのバーを–」

②ある時期、特に年を過ごしてその後になる。「暑い夏を–」「–・えて天明9年の正月」

③ある基準より上になる。「分ぶんを–・えてしゃしゃり出る」

㋐…より大きな値になる。「100人を–」「30度を–暑さ」「予想を–」

㋑ぬきんでて、まさる。「人に–」（人よりまさる）。超越する。「現代を–」「党派を–・えて大局を見る」㋐㋑は「超える」とも書く。

㋒順序を追わず、先に出る。飛び越す。「兄を–・えて弟が家をつぐ」

34

技と工夫　その７

辞林２１　三省堂

こ・える【越える】

(1)障害物や境界線の上を通り過ぎて，その向こう側へ行く。越す。「峠を–・える」「–・え難い壁」(2)ある日時が過ぎる。「–・えて1990年の1月」

こ・える【超える】

(1)ある基準・数値を上まわる。「4万人を–・える大観衆」(2)自分の考え方や立場をすてて先へ進む。超越する。「怨讐を–・えた援助」(3)他よりもすぐれる。まさる。「力量が人に–・える」

※大辞林（三省堂），広辞苑（岩波書店）は，

こ・える【越える・超える】

35

国語辞典の情報には多くの技・工夫がある

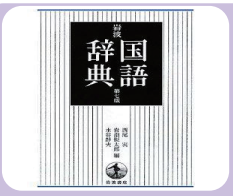
- 見出し語
- 表記
- 品詞
- 意味
- 用法
- 用例
- その他

どれをとっても，辞典間でさまざまな考え方や立場の違いによって書き方に差がある。

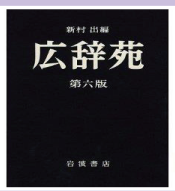
国語辞典それぞれに，それぞれの，技や工夫が多くある！！

36

たくさんの国語辞典



小型　現代代生活に必要な語，使用頻度の高い語を収録。約6～9万語。『岩波国語辞典』（岩波書店），『三省堂国語辞典』（三省堂），『新選国語辞典』（小学館），『新明解国語辞典』（三省堂），『明鏡国語辞典』（大修館書店）など



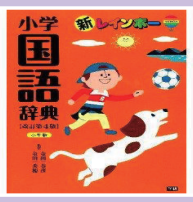
中型　現代語と百科項目を収録。約20～25万語。『広辞苑』（岩波書店），『大辞林』（三省堂），『大辞泉』（小学館）など

大型　国語項目と百科項目を収録。約50万語。『日本国語大辞典』（全13巻，小学館）。

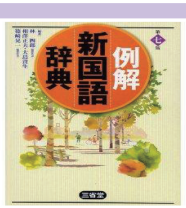
37

学習用国語辞典

ほとんどがひらがな。小学校入学前から小学校低学年向け。440～1,200語。『三省堂こどもこくごじてん』(三省堂)，『こどもにほんごじてん』(講談社)など



漢字全てに読み仮名。小学校低学年から。2.5～3.7万語。『例解学習国語辞典』(小学館)，『新レインボー小学国語辞典』(学研)，『下村式小学国語学習辞書』(偕成社)，『小学新国語辞典』(光村教育図書)『チャレンジ小学国語辞典』(ベネッセ)，『くもんの学習 小学国語辞典』(くもん出版)，『例解小学国語辞典』(三省堂)，『旺文社小学国語新辞書』(旺文社)など



中学生から。4.5～5.7万語。『例解新国語辞典』(三省堂)，『学研現代標準国語辞典』(学研)，『旺文社標準国語辞典』(旺文社)，『ベネッセ新修国語辞典』(ベネッセ)，など

38

シソーラス・類語辞典とは・・・

「シソーラス」
「類語辞典」

語句が五十音順ではなく，意味によって分類，配列され，似た言葉どうしが近くに並べられているもの。

39

どんなふうに利用できるか

表現辞典：
適切な言葉

言語研究：
ことばの分布
作品の用語調査
ことばの結びつき

情報処理：
情報検索
類似度の計算

40

シソーラス・類語辞典の目的と体裁

目的　①適切な言葉探し
②語彙の分布や偏りを見る
③使い分けを知る

シソーラス

類語辞典

体裁　A) 語釈や用例などがあるか
B) 分類体系は品詞別であるか
C) どのような語がどれくらいの数，収録対象か

41

『分類語彙表』とは

意味の世界を分類したもの

『分類語彙表』の「まえがき」より：

ここに分類語彙表というのは，一般に一つの言語体系の中で，その語彙を構成する一つ一つの単語が，それぞれどのような意味で用いられるかを一覧できるように，単語が表わし得る意味の世界を分類して，その分類の各項にそれぞれの単語を配当したものである。

42

『分類語彙表』 の公開

・国立国語研究所資料集6 『分類語彙表』　秀英出版　1964年
　初版刊行 （1997年 31版）

・『分類語彙表 フロッピー版』　秀英出版　1994年 刊行
約3万７千語

・『分類語彙表　増補改訂版』　大日本図書　2004年 報告書刊行，
電子化データベースの公開
約9万6千語

いずれも研究利用に無償で公開されている。

語彙調査データ

http://pj.ninjal.ac.jp/corpus_center/archive.html

43

『分類語彙表』 の担う役割

①言葉や概念を手がかりに，適切な言葉を見つけるもの

②語彙の分布や偏りを見るための「物差し」となるもの

44

『分類語彙表』の紹介記事

オトナの心得

語彙不足につける「薬」

「確か、こういう言葉があったはずだが……」

懸命に脳から引き出そうとしても、まず出てこない。生まれつき、語彙（ごい）が不足しており、人前で話すとき的確な表現ができない。原稿を書くときも、言葉を絞り出すのに、苦労する。

二十年ほど前、人知れず悩んでいるときに出合ったのが国立国語研究所が編集した「分類語彙表」（秀英出版）。三万六千の言葉を意味によって体系的に分類した本で、一九六四年に初版が出た。

手の内を明かすと、このコラムのタイトルを考えるとき

も役立った。「常識」と同じような意味で、もっと親しみのわく言葉はないか。「常識」を索引で調べると、「知識・意見など」の項目に導かれる。心得や知識、良識、……などが並ぶ。そこから「心得」を拾い出したのである。

同研究所が編集した書籍の中で、この本の発行部数は突出している。三十一回も版を重ね、九万部が出た。なのに、九六年には出版社が自主解散。大日本図書が発行を引き継いだものの、九九年絶版に。

改訂版の出版予定はないか、研究所に問い合わせた。初版の編纂（へんさん）に当たった元所長の林大（おおき）さん（88）らが目下、増補版作りの準備中。新語を加え、しめて九万語。来年度中に発行する計画という。

語彙不足を嘆く人が増えているようだ。そんな人は、この種の本が身近にあると安心だ。増補版の「分類語彙表」が安価でかつ早めにお目見えすることを願ってやまない。

（編集委員　足立則夫）

イラスト・南　伸坊

日経新聞 プラス１

（朝刊別刷り）

2001年12月8日(土)

45

主なシソーラス・類語辞典①

（１）『角川類語新辞典』 (1981)

- 収録語数：約60,000語
- 分類の特徴：大分類(10)，中分類(100)，小分類(1,000)，細分類(約3,000)の４階層。品詞の区別なし。大分類は，「自然」「人事」「文化」という枠の中で設けられたもの。
- その他特徴：語釈，用例あり。位相が付与されている。

（2）『日本語語彙体系』 (1997)

- 収録語数：約300,000語
- 分類の特徴：１階層目が品詞。2階層～12階層目に2,876個の意味属性を木構造で配置。
- その他特徴：さらに6000語の用言には日英の対訳文型14,000パターンを付す。機械翻訳を主目的にして開発されたため，固有名詞も多く収録されている。

46

主なシソーラス・類語辞典②

（3）『類語大辞典』(2002)

- 収録語数：約79,000語
- 分類の特徴：カテゴリー(100)，小分類(916)，品詞(11)，小見出しの４階層。カテゴリーは用言に基づいて作成し，人に近いところから遠いところへという原則にしたがって配置。
- その他特徴：語釈，用例あり。位相が付与されている。

（4）『日本語大シソーラス』(2003)

- 収録語数：延べ約320,000語，異なり約200,000語
- 分類の特徴：カテゴリー(1,044)，小語群(14,000)，セミコロン（任意）の３段階。品詞の区別なし。カテゴリー部分が先に２～４階層に分類されているため，あわせると一番深い場合で６階層。語釈なし。文法的カテゴリーなどの一部に用例あり。
- その他特徴：故事成句，人名やオノマトペが多く収録されている。

47

主なシソーラス・類語辞典③

（5）『類語例解辞典』新装版(2003)

- 収録語数：約25,000語
- 分類の特徴：大分類(10)，中分類(200)，グループ(約6,000)の３段階。助詞・助動詞は別枠で78に下位分類。品詞の区別なし。
- その他特徴：語釈，用例のほか，使い分けの解説，対比表などがある。

（6）『三省堂類語新辞典』(2005)

- 収録語数：約50,000語
- 分類の特徴：柱(3)，ジャンル(18)，分野，領域の３段階。品詞の区別なし
- その他特徴：語釈，用例のほか，位相，使い分けの解説，図解などがある。また，類語のニュアンスを解説するコラムが86ある。オノマトペも積極的に収録されている。

48

比較ー分類体系は品詞別であるか

品詞別である	『分類語彙表』　『類語大辞典』 cf.『ロジェのシソーラス』
品詞別ではない	『日本語大シソーラス』 『角川類語新辞典』 『三省堂類語新辞典』

＜品詞別の利点＞

- 同じ品詞どうしで，語の比較がしやすい
- 品詞別にどのような意味分布があるかの見通しが良い

49

比較ー人名はどう分類されるか

『分類語彙表』	＜1.2390 人名＞ 01　佐藤　鈴木　田中　高橋　山田 スミス　ジョンソン 09　紫式部　清少納言　鴨長明　滝沢馬琴 夏目漱石　森鴎外　芥川龍之介 17　浦島太郎　シンデレラ　ピノキオ ロミオ　ジュリエット
『日本語大シソーラス』	「シンデレラ」 ＜架空人物＞ ＜美女＞ ＜なさぬ中＞（継子関係） ＜いじめる＞

「人名は人物典型を表現する際の最初の手掛かりになるものである」（『日本語大シソーラス』序）

50

比較–属性分類か主題分類か①

属性分類　語が本来もつ性質によって分類することを優先するもの

（例）『分類語彙表』　＜1.5　自然現象＞「発病」

＜1.3　人間活動＞「治療」

主題分類　ある主題のもと関連する語を集めて分類することを優先するもの

（例）『角川類語新辞典』 073「発病」＞　C「治療」

51

コーパス

52

コーパスとは何か

目的

言語の実態を分析するための重要なデータとして、言語学、言語教育、辞書編集、言語情報処理などの分野で利用されている。

歴史

Brown Corpus（1964年、米、100万語）が最初。その後、British National Corpus（1994年、英、1億語）など、英語のコーパスを中心に発展してきた。

言語

英語、韓国語、中国語、スペイン語、ドイツ語、フランス語など、さまざまな言語でコーパスが作られている。

種類

書き言葉コーパス、話し言葉コーパス、通時コーパス、学習者コーパス、対訳コーパスなど、目的により多様なコーパスが作られている。

規模

書き言葉では、1億語以上のものが多い。話し言葉では、British National Corpusの一部(約1000万語)や『日本語話し言葉コーパス』の約700万語が最大級。

構築

コーパスの構築には、多額の費用が必要なため、大規模なコーパスのプロジェクトには国費が投入されている。

53

『現代日本語書き言葉均衡コーパス』

- 「コーパス」とは、言葉の実際の使われ方をしらべるために作られた，言語研究のための資料。体系的に言語の資料を収集し，研究用の情報（たとえば，品詞情報）を付加したもの。

- 目的： 現代日本語の書き言葉を代表するコーパス
- 対象： 1976年～2005年に出版された書き言葉
  - 書籍，雑誌，新聞，白書，教科書，Web文書…
  - 全部で1億語　（文庫本1700冊分相当）
- 公開： Web上でオンライン検索『少納言』にて公開中

書き言葉の総体 → コーパス

54

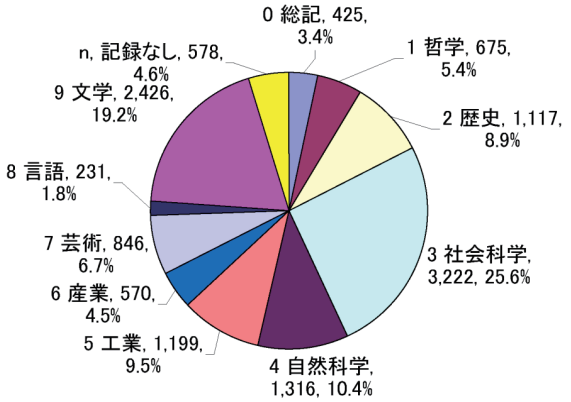
『現代日本語書き言葉均衡コーパス』の構成

現代日本語書き言葉均衡コーパス

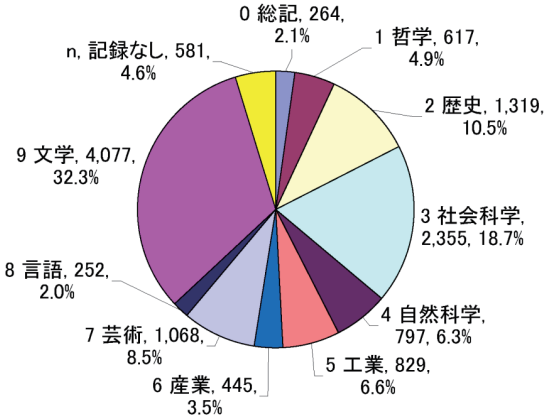
出版サブコーパス	図書館サブコーパス	特定目的サブコーパス
書籍・雑誌・新聞	書籍	白書・教科書・広報紙・ベストセラー・Yahoo!知恵袋・Yahoo!ブログ・韻文・法律・国会会議録
2001-2005年	1986-2005年	
約3,500万語	約3,000万語	約3,500万語

55

書籍　出版・図書館サブコーパス



出版サブコーパスにおける取得サンプル数のNDC別内訳（サンプル数およびその割合（%））



図書館サブコーパスにおける取得サンプル数のNDC別内訳（サンプル数およびその割合（%））

実際と同じ配分になるように設計する

56

現代語の中の「さもありなん」

まさか和久さんが指導員として復帰してるなんて
思わなかったから。でも、さもありなん、という気もする。

君塚良一|脚本；丹後達臣|ノベライズ『踊る大捜査線スペシャル』フジテレビ出版；扶桑社(発売)，1998

美也子たちの離婚を、夫の実家では、さもありなん、といった諦めの心境で受けとめてくれた。

藤堂志津子|著『恋人よ』講談社，2003

「でもさ、霊ってお墓に住んでいるわけじゃないと思うんだけど…」あまりにも「さもありなん」なオチだったので私は呆れて言った。

田口ランディ|著『ほつれとむすぼれ』角川書店，2004

「第十二深宇宙探検隊の隊長の秋篠氏と君は同じ苗字だけど、親戚なの？」教室中に緊張が走った。さもありなん。さっき同じ質問をしたルイスは完璧に打ちのめされたんだ。

榎木洋子|著『Step out』集英社，1996

57

現代語の中の「なきにしもあらず」

また、さざ波のように押しよせる技術革新と対峙してきたブルーカラーにしてみれば、ＯＡ化に浮足だつホワイトカラーの姿は、いまさらの感も無きにしもあらずで、滑稽にすら映るかもしれない。

加藤仁|著『日本の課長』PHP研究所，1988

その声には、質問せずにいられないだけでなく、思いあたる節がなきにしもあらずという気持ちもあらわれていた。

クーパー|作；村山淳彦|訳『開拓者たち下』岩波書店，2002

「そのデータが中尾弥生に渡っていた可能性もなきにしもあらずね」
「そうですね」

志茂田景樹|著『美雪総監と消えたクルーザー』光文社，1993

でないと、“ピアノ離れ”がドーッと出てしまった『子供のピアノ教室』の二の舞になる恐れがなきにしもあらずだわ。

大村典子，大崎妙子|著『大人のピアノ長続きのコツ』ヤマハミュージックメディア，1997

58

現代語の中の「べし」

男は黙って「魂」を聴くべし。

鈴木敦史|著『クラシック批評こてんぱん』洋泉社，2001

もしあなたが深夜に鍾路界隈にいて、空腹を満たしたくなったらこのエリアに直行すべし。

松井聖一郎|著『韓国 長期滞在者のための最新情報55 』三修社，1996

◆賢い消費行動のための基本原則は「資産と負債の総合管理」にあり。
◆資産が増えないのであれば、負債を縮小させるべし。

小学館文庫編集部|編『デフレ時代を生きる知恵』小学館，2001

これはこれでいいクルマ。いやはやイタリア恐るべし

福野礼一郎|著『礼一郎式外車批評』双葉社，2003

そんな尾崎さんの描くＥＭＥが、数量限定でクリアファイルと下敷きになりました。ゲットすべし。

瀧川武司|著『EME blue』富士見書房，2002

59

コーパス：公開①

http://www.kotonoha.gr.jp/shonagon/

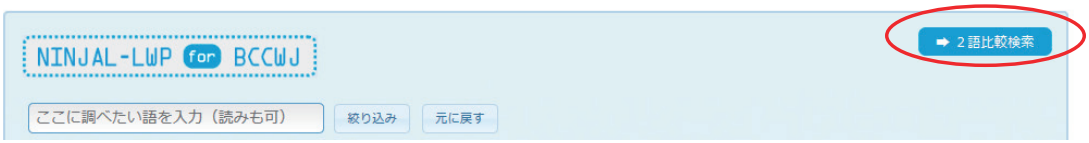
60

コーパス：公開②

http://nlb.ninjal.ac.jp/

61

『LWP』の２語比較


NINJAL-LWP for BCCWJ
２語比較検索
ここに調べたい語を入力（読みも可）
絞り込み
元に戻す

NINJAL-LWP for BCCWJ
問題　課題
絞り込み
元に戻す
２語比較
名詞
動詞
形容詞・連体詞
副詞
オノマトペ
見出し
読み
ローマ字表記
頻度
問題　モンダイ　mondai　68,064
課題　カダイ　kadai　8,035

63

問題を... 749
課題を... 247
問題
課題
コロケーション
頻度
MI
LD
LD差
問題を引き起こす 111 9.6 8.59 8.59
問題を起こす 163 8.32 8.48 8.48
問題をはらむ 49 10.39 7.68 7.68
問題を生じる 57 6.47 6.79 6.79
問題を生み出す 30 7.17 6.56 6.56
問題を回避する 20 8.59 6.35 6.35
問題を討議する 18 10.71 6.31 6.31
問題を先送りする 15 10.83 6.06 6.06
問題を持ち出す 16 7.58 5.94 5.94
問題を重視する 18 6.66 5.89 5.89
問題を生ずる 15 7.24 5.81 5.81
問題を追及する 13 8.38 5.76 5.76
問題を内包する 12 9.79 5.72 5.72
問題を解明する 12 8.16 5.64 5.64
問題を担当する 14 6.59 5.6 5.6
問題を含める 79 7.13 7.36 課題を含める 1 3.6 1.78 5.58
問題を扱う 95 8.23 8.02 課題を扱う 1 4.43 2.5 5.52

助詞+動詞
問題
課題
パターン
頻度
問題が... 7738 課題が... 447
問題に... 7552 課題に... 771
問題を... 7268 課題を... 1063
問題と... 3189 課題と... 1378
問題は... 2338 課題は... 130
問題も... 1379 課題も... 79
問題で... 1018 課題で... 51
問題から... 203 課題から... 11
問題の... 185 課題の... 13
問題まで... 45
問題へ... 43 課題へ... 3
問題より... 25 課題より... 1

問題を... 749
課題を... 247
問題
課題
コロケーション
頻度
MI
LD
LD差
課題を遂行する 13 10.79 7.87 -7.87
課題を採択する 6 9.67 6.76 -6.76
課題を公募する 3 11.46 6.39 -6.39
課題をやり遂げる 3 10.14 6.2 -6.2
課題を達成する 7 8.26 6.07 -6.07
課題を取り払う 2 9.87 5.68 -5.68
課題を掘り起こす 2 9.85 5.67 -5.67
課題を実行する 5 7.4 5.32 -5.32
課題を内含する 1 14.06 4.94 -4.94
課題を担わす 1 12 4.91 -4.91
課題を感得する 1 10.74 4.87 -4.87
課題を羅列する 1 10.54 4.86 -4.86
課題を深化する 1 10.36 4.84 -4.84
課題をレポートする 1 10.05 4.82 -4.82
課題を摘出する 1 9.53 4.77 -4.77
課題を上乗せする 1 9.27 4.74 -4.74

コーパス：公開③

https://chunagon.ninjal.ac.jp/ （申請が必要）

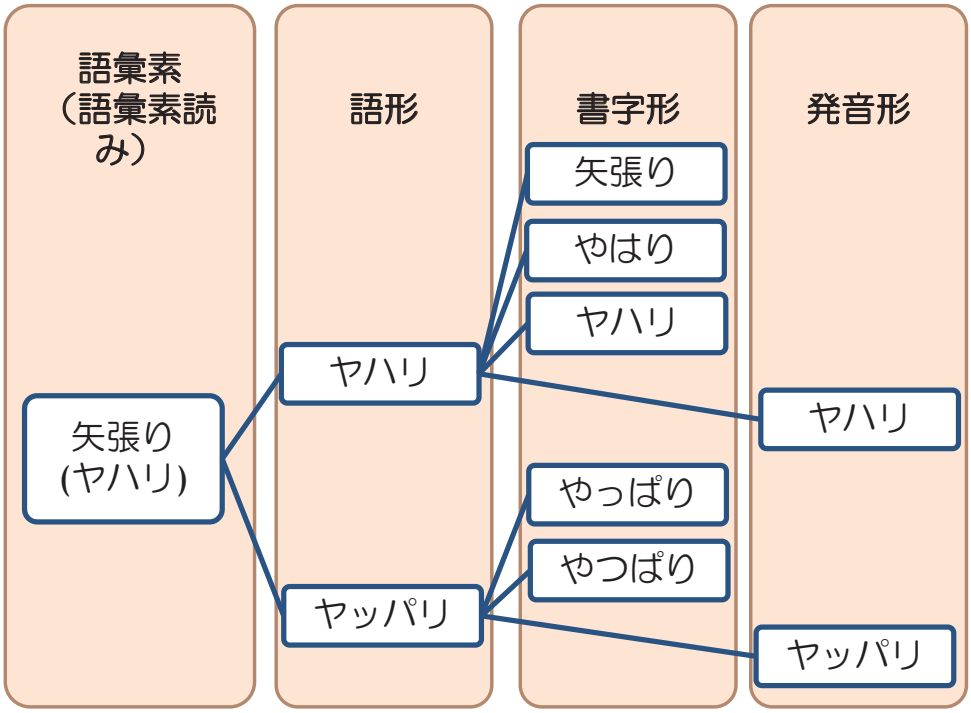
65

中納言の特徴

- 短単位と長単位とで検索できる
- 文字列検索だけではなく、見出し語検索ができる
- 品詞を指定した検索ができる
- 前にくるもの、後ろにくるもの、といった検索ができる

66

「矢張り(ヤハリ)」の階層構造



※「ヤッパ」「ヤッパシ」「ヤパ」「ヤパリ」「ヤーッパリ」もあり。67

68

1,120 件の検索結果が見つかりました。そのうち 500 件を表示しています。
検索対象語数：124,624,010　記号・補助記号・空白を除いた検索対象語数：105,377,883

サンプルID	開始位置	通番	前文脈	キー	後文脈	語彙素読み	語彙素	語彙素細分類	品詞	活用型	活用形	レジスター	執筆者	書名/出典	編著者等	出版者	出版年
OM21_00004	555560	348470	この二月一日の議事録の中で#「共同研究開発という言葉は非常な	いろいろ	(な)(言葉)を含むのでございまして、場所によりましてその定義が区々で	イロイロ	色々		形状詞一般			特定目的・国会会議録		国会会議録			1982
OT02_00019	20330	13070	、外来語を駆使することで表現を豊かにしてきた。#日本語の	微妙	(な)(言葉)の使い分けを知ることで、留学生たちの言葉に対する感覚も磨かれる	ビミョウ	微妙		形状詞一般			特定目的・教科書		国語 1	宮地裕|ほか著	光村図書出版株式会社	2005
OT02_00019	38660	24810	な違い。#慣用句・ことわざ#「耳を傾ける」「寄とったきねつかりの	よう	(な)(言葉)を集める。#日本語の特色#・外国人にとって、日本語のどんな	ヨウ	様		形状詞-助動詞語幹			特定目的・教科書		国語 1	宮地裕|ほか著	光村図書出版株式会社	2005
OC02_00626	1330	730	で喜んでる連中です。#普通スラングと言うのはその仲間内でしか理解できない	特殊	(な)(言葉)ですが2ちゃんねるのは素人でもすぐに判っちゃいますから。#変味ない	トクシュ	特殊		形状詞一般			特定目的・知恵袋		Yahoo!知恵袋		Yahoo!	2005
OY14_13248	19480	13080	！」#「・・・・・オイ。」#「へ？」#「オマエ、いま	重要	(な)(言葉)をさらっと受け流しただろ。」#「へ？」#「だからその・・	ジュウヨウ	重要		形状詞一般			特定目的・ブログ		Yahoo!ブログ		Yahoo!	2008
OV2X_00013	9960	6280	人の耳にあわれにひびいた#寒い地方　#暑い地方#極国をまわって来たその	わずか	(な)(言葉)はその季節季節の風のなかにあわれにしわがれて消えていった#鳴島	ワズカ	僅か		形状詞一般			特定目的・詩文		中野重治詩集	中野重治|著	思潮社	1988
OC14_10116	110	50	ありがとう、をもっと	素敵	(な)(言葉)で言えませんか？#「君の真心、確かに受け止めたよ」#「	ステキ	素敵		形状詞一般			特定目的・知恵袋		Yahoo!知恵袋		Yahoo!	2005

69

276,104 件の検索結果が見つかりました。そのうち 500 件を表示しています。
検索対象語数：124,624,010　記号・補助記号・空白を除いた検索対象語数：105,377,883

サンプルID	開始位置	通番	前文脈	キー	後文脈	語彙素読み	語彙素	語彙素細分類	品詞	活用型	活用形	レジスター	執筆者	書名/出典	編著者等	出版者	出版年
LBf7_00035	124780	79500	カウンターに立つことになる。#もちろん、金山は技術面の習得だけにとどまらず、	アメリカ的	なサービス業の取り方をも身につけ、オフリミットの外へ出て	アメリカ	アメリカ	America	名詞-固有名詞-地名-国			図書館・書籍	東谷　譲(著)	進駐軍クラブから歌謡曲へ	東谷譲|著	みすず書房	2005
LBl9_00175	55930	35340	いった。#「二足すこは絶対に五にはならない」彼以上に	アルテレス的	確信を込めようとしながら、私はまず手始めにそういった。#「二足す			未知語				図書館・書籍	マイク・ティッドウェル(著)/吉嶺英美(訳)	アマゾンの白い酋長	マイク・ティッドウェル|著/吉嶺英美|訳	翔泳社	1997
LBe2_00047	85230	59490	ものばかりでなく、一見そうしたものと関係のない施設や事業までも。	イデオロギー的	意味合いが色濃く付与されていた。#この点は、先の県	イデオロギー	イデオロギー	Ideologie	名詞-普通名詞一般			図書館・書籍	中島　三千男(著)	天皇の代替りと国民	中島三千男|著	青木書店	1990
OC03_00410	490	240	あるクレジットカードは何が思い浮かび上がすか？#アメリカンエキスプレスかな。#中身は知らんけど	イメージ的	には。	イメージ	イメージ	image	名詞-普通名詞-サ変可能			特定目的・知恵袋		Yahoo!知恵袋		Yahoo!	2005
PB49_00275	41340	25680	べき規範を命令する。#それらはみな何らかの意味において、あるいはいは	カント的	な「意識一般」の意味において、あるいはいイデッガー的な「ひと」即ち	カント	カント	Kant	名詞-固有名詞-人名一般			出版・書籍	三木　清(著)	日本近代文学評論選	千葉俊二, 坪内祐三|編	岩波書店	2004

70

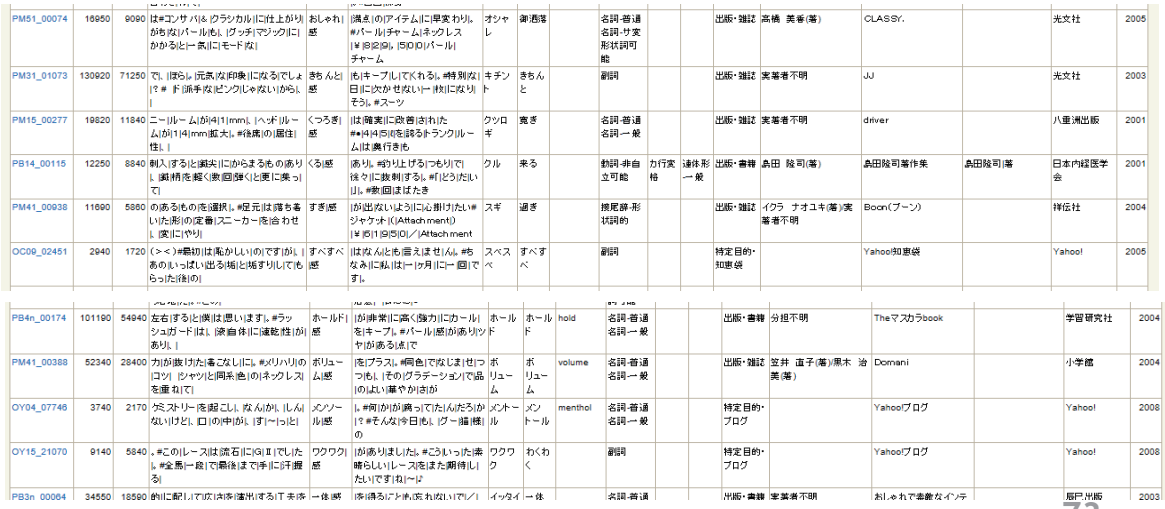
LBk3_00141	9080	5420	、起こる可能性がある場合、その情報を国際社会とメディアに伝える。#●	国際的	な平和維持要員や警察力が配置されている国の場合は、安全	コクサイ	国際		名詞-普通名詞一般			図書館・書籍	Jeff Crisp(著)/実著者不明	世界難民白書	国連難民高等弁務官事務所(UNHCR)編著;UNHCR駐日事務所訳・編	読売新聞社	1996
PB32_00034	23570	16010	を剥離した後、力を加える部分である打面の調整を入念に行い、	最終的	に細石刃を剥離している。#剥離は石や鹿角のハンマーで叩くの	サイシュウ	最終		名詞-普通名詞一般			出版・書籍	加藤　真二(著)/岡田　康博(著)	発掘された日本列島	文化庁編	朝日新聞社	2003
PB43_00615	13840	9000	とか「君は理系へ進んだ方がいいと思うよ」などという、	具体的	なアドバイスはとうとうできませんでした。#二人が帰ったあと、私は	グタイ	具体		名詞-普通名詞一般			出版・書籍	宮地　正卓(著)	教育改革への哲学的視点	宮地正卓|著	日本図書センター	2004
LBl3_00070	6420	4150	みることができる。#新聞の投稿欄は六〇年以降、選考基準を変え、	政治的	な問題にふれる投書を避け、ただ、そのひたすら受動的な感情の記録を	セイジ	政治		名詞-普通名詞一般			図書館・書籍	天野　正子(著)	フェミニズムのイズムを超えて	天野正子|著	岩波書店	1997
LBp2_00025	17520	11980	農業界の支配体制が確立し、戦後も踏襲されていく。# なお、	伝統的	な農本主義に基づく皇国農村の建設が政策目標として提唱された	デントウ	伝統		名詞-普通名詞一般			図書館・書籍	山路　健(著)	明治・大正・昭和の世相史	山路健|著	明治書院	2001
LBs4_00034	61540	38580	、重力的な影響は大きい。#しかし、このような遭遇が起こるかどうか、	確率的	に比べると、衛星銀河の合体が起こる確率はかなり高い。#伴銀河の存在	カクリツ	確率		名詞-普通名詞一般			図書館・書籍	谷口　義明(著)	クェーサーの謎	谷口義明|著	講談社	2004
PB44_00091	8890	5350	対応できなかった従業員を、やる気がないとか、覇気がないとか、	人格的	欠陥人間としうとまれ切り捨てられていく例もあった。# しかし、	ジンカク	人格		名詞-普通名詞一般			出版・書籍	斎藤　茂太(著)	「ストレス知らず」の処方箋	斎藤茂太|著	麗澤大学出版会;廣池学園事業部(発売)	2004
PB20_00093	75970	50920	申し渡された。#片や検察の最高位をきわめ、敗者は法務大臣どころか、	社会的	にも葬り去られたのである。#明暗を分けるというが、この懸隔を	シャカイ	社会		名詞-普通名詞一般			出版・書籍	本田　靖春(著)	本田靖春集	本田靖春|著	旬報社	2002

71

609 件の検索結果が見つかりました。そのうち 500 件を表示しています。
検索対象語数：1,419,729　記号・補助記号・空白を除いた検索対象語数：1,131,891

名大会話コーパス

会話ID	開始位置	連番	前文脈	キー	後文脈	語彙素読み	語彙素	語彙素細分類	品詞	活用型	活用形	収録日	収録場所	参加者の関係	発話者ID	性別	年齢	出生地1	出生地2	居住地	居住歴
data100	149510	91880	って、Iさんとは全然違うかったもん。レベルが。#うん、そうそう。#でも	積極的	にやらんから。#うーん。#それもあるんだろうね。#ほんとはもっとできてる	セッキョク	積極		名詞-普通名詞一般			2002年1月17日	院生室	大学院の同級生	F140	女性	20代前半	香川県	高松市	愛知県名古屋市	大学卒業まで高松市在住。
data003	80200	52240	行ってさ、2週目行かなかったじゃん。#うん。#忘れるね。#だから全然	感覚的	にさあ、まだ身についてないときに休んでるから、あれどこだった	カンカク	感覚		名詞-普通名詞一般	名詞-サ変可能		2001年10月23日	車中(某大から所属大学への帰り道。運転者F033)	大学院の同級生	F056	女性	20代前半	愛知県	岡崎市	愛知県岡崎市	
data076	119770	74590	さん肉とかよく食べるじゃん。#食べない。#お父さん、肉嫌いだもん。#	本質的	には、#お母さんがやたら肉買ってきて。#うそお#料理するから	ホンシツ	本質		名詞-普通名詞一般			2002年1月1日	F001の実家	M001はF079の娘婿であり、F001の父親。	M001	男性	50代前半	山梨県	北巨摩郡	山梨県北巨摩郡	
data081	140440	86630	があって、その、まあ#うんうんうんうん#1分野、2分野みたいなの？#	全体的	なこと。#で、あとで、次に、生物、物理、化学、地理。	ゼンタイ	全体		名詞-普通名詞一般			2001年12月21日	職場	同僚	F050	女性	20代後半	岐阜県	岐阜市	岐阜市	
data118	156130	99110	できるから。#あ、そうそう、その指導さんの最後のところは、その、ほんとに	日常的	なことでもいいから、なんか話を聞いたとき、あ、そうな	ニチジョウ	日常		名詞-普通名詞一般	名詞-副詞可能		2002年7月14日	F157宅	F158とM037は夫婦。F157はM037の母親。×は二人の娘で、データ収集者。	F158	女性	50代前半	長野県	塩尻市	名古屋市	22歳から現在まで名古屋市に居住。
data124	109840	71650	って。#でも、ほんとわかってんのかな。#どんどんやってくれ。#すーげえー。#	本人的	にはへれすデビューしたいらしいからさ。#D2先生にもこう、頑張っ	ホンニン	本人		名詞-普通名詞一般			2002年7月13日	北海道大学喫煙室	友人	M007	男性	20代後半	北海道	滝川市	札幌市	
data051	3760	2500	にさー、車のデザインさー、別に。#なんか、どこかの国がすごく	先駆的	だとか、そういうことはないでしょ？#いまどきさ。#うーん、と思う。	センク	先駆		名詞-普通名詞一般	名詞-サ変可能		2001年12月13日	院生室	F019はF004の大学院の後輩。	F019	女性	20代後半	東京都	八王子市	愛知県長久手町	0歳から23年間八王子市に居住。
data050	132370	83950	で受け入れられるぐらいの。#一般的うんうん。#だって、河合隼雄が、その、何、	一般的	に結構出てきて、知られてっていうのって、十何年か	イッパン	一般		名詞-普通名詞一般			2001年12月12日	喫茶店	大学の同級生	F050	女性	20代後半	岐阜県	岐阜市	岐阜市	
data050	69830	44930	。#うん#そうそうそう。#ふーん#なんかね。#で、たぶんそれはちょっと精神に。	精神的	にちょっと異常がある人が書いたんがなっていう気がする。#うん	セイシン	精神		名詞-普通名詞一般			2001年12月12日	喫茶店	大学の同級生	F039	女性	30代前半	兵庫県	川西市	宗都市	0歳から約23年間川西市に居住。
data119	125850	80350	特別になりそうなあれかもしれんけど。#ま、お前は結構なんか、	基本的	には仲よくしとるつもりだでさ。#うん#べつにそんなに、ねえ。#ζ	キホン	基本		名詞-普通名詞一般			2002年7月22日	M036宅	高校の同級生	M036	男性	10代後半	愛知県	名古屋市	名古屋市	
data089	95130	59840	フランス式はねー、サージュファムのね、力がすごく強い。#そう、お産の	具体的	なしかたはまあ、知らんけど。#うん#とり、とりあえず、お医者さんなんかほとんど	グタイ	具体		名詞-普通名詞一般			2002年1月17日	レストラン	大学の同僚	F011	女性	40代後半	大阪府	堺市	大阪府堺市	
data012	104950	65370	ああ、あたしめにはって。#うん#やっぱり使っちゃいます？#あります、あります。#	あたしめ	には、桃井かおりとさあ、あのー、ベリーベリーのコマーシャルしてる人だれだ	ワタシ	私		代名詞			2001年10月29日	大学研究室	F017はF102の大学の先輩。	F017	女性	20代前半	愛知県	岡崎市	名古屋市	出身地での居住期間が長く、名古屋市での在住期間は1年未満。
data104	119250	74790	けどさー、ねえ、別に目がぱっちりしてるとかなんか、そのいかにも	モテル的	顔。#うん#まあねー#でも、モテルさんってさー、何、こういうあっさりし	モテル	モテル	model	名詞-普通名詞一般			2002年1月16日	F106宅	中学校の同級生。	F106	女性	10代前半	オーストラリア		名古屋市	3歳から名古屋市に居住。
data116	15210	9880	あんまりよくない。#まじ#それいいの？#うそー。#うん#C家の方がいいわけ？#	顔的	に。#まあ必ずたりよったりだけどね。#目当レベルが。#なんかなが月	カオ	顔		名詞-普通名詞一般			2002年5月4日	M011宅	大学の同級生	M013	男性	20代後半	福岡県	北九州市	名古屋市	
data024	7520	4610	ですね。#柿の種みたいな、ああいうあれがいいかなっ	つまみ的	なものですとチーズの盛り合わせとかナッツとかそう	ツマミ	摘み		名詞-普通名詞一般			2001年11月24	ホテルのレストラン・バー	科研費による研究の共同研究者	×						



73

コーパス：公開④

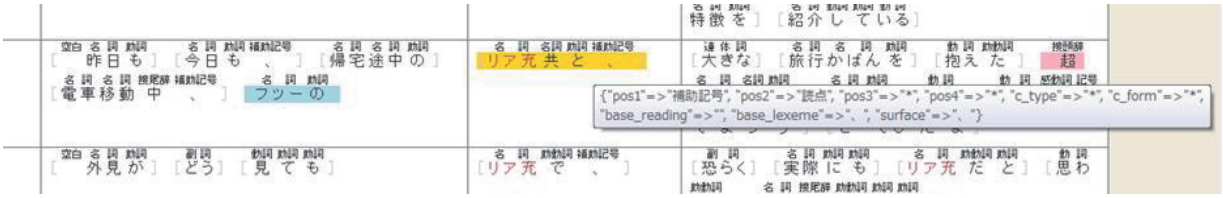
『国語研日本語ウェブコーパス（NWJC）』

258億語の日本語コーパスをウェブで公開～国立国語研究所 -INTERNET Watch
http://internet.watch.impress.co.jp/docs/news/1047913.html

- NWJCは、2014年10～12月にインターネット上で8399万2556のURLから収集した258億3694万7421のコーパスを収録。文数では14億6314万2939、のべ文数は38億8588万9575に上るデータが格納されている。
- 梵天の一般公開版では、この258億のコーパスに対する文字列検索が行える。

「梵天」の検索結果

http://pj.ninjal.ac.jp/corpus_center/nwjc/bonten-results.html



74

『梵天』

「梵天」の検索機能：文字列検索

http://pj.ninjal.ac.jp/corpus_center/nwjc/

「文字列検索」は最も手軽にNWJCのデータを調べることのできる検索方法です。

検索の手順

1. 文字列検索では、はじめに「文字列検索」タブをクリックする。

2. 「検索文字列」の入力欄に調べたい文字列を入力する。

3. 「URLドメイン」を指定したい場合は、末尾２パート（第２レベルドメイン）を入力する。
4. 「検索」ボタンを押す。
5. 検索結果が「検索」ボタンの下部に表示される。
件数が50件以上ある場合の検索結果は、1～50件のみが表示されている。

http://bonten.ninjal.ac.jp/

75

３．「ソフトウェア要求仕様書の品質
～ IEEE Std. 830-1998 より～」

ソフトウェア要求仕様書の品質
～IEEE Std. 830-1998より～

東京女子大学
白銀純子

講演者

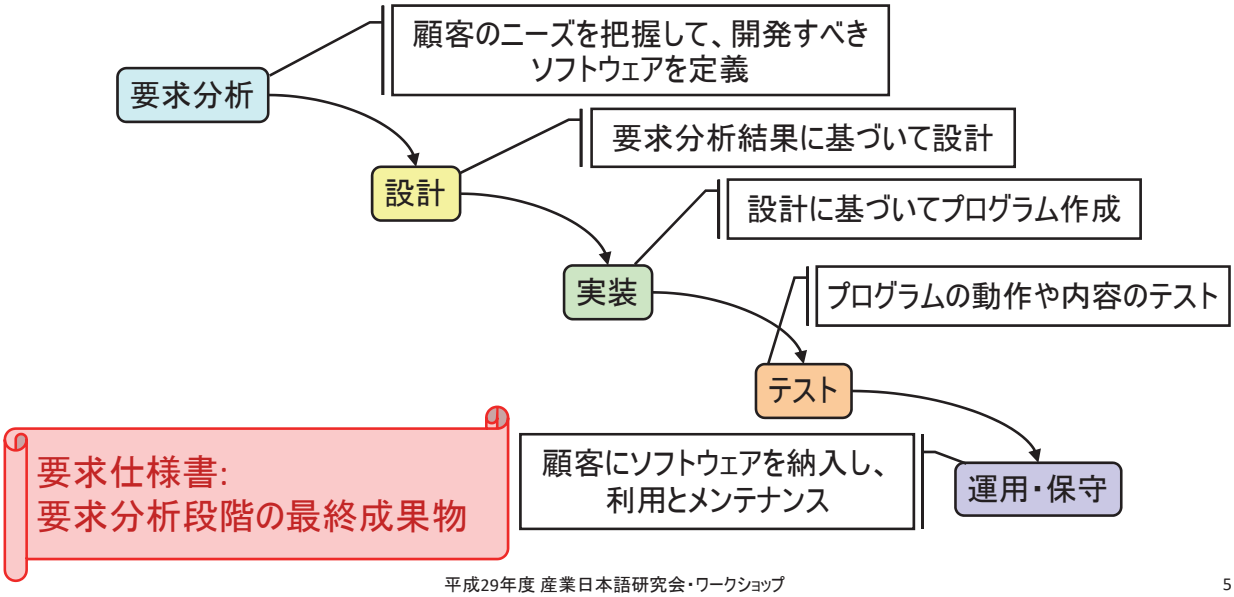
- 白銀純子
  - 現職: 東京女子大学准教授
  - 専門: ソフトウェア工学
  - 主な研究テーマ: 要求工学およびユーザインタフェース(ソフトウェアの画面構成)の開発支援

本日の内容

- 要求仕様書とは
- IEEE Std. 830-1998～ソフトウェア要求仕様書～
- IEEE Std. 830-1998～品質特性～
- 要求仕様書とソフトウェアの品質の相関

要求仕様書とは

ソフトウェア開発の流れ

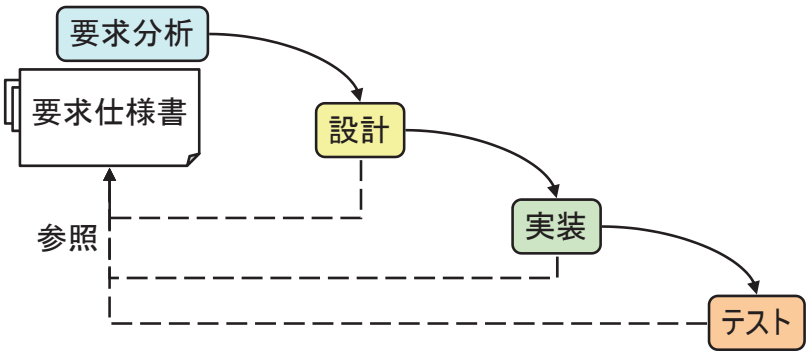


要求仕様書とは?

- どのようなソフトウェアを開発すべきかを文書化したもの
  - ソフトウェアに対して求めること: 要求
    - ソフトウェアが持つべき機能
    - ソフトウェアが持つべき特性
    - 開発時に考慮すべき条件, etc.
- 要求分析段階での最終成果物
  - 顧客からの聞き取りや業務の分析、資料の調査など、様々な作業の後、結果をまとめて作成
- 開発するソフトウェアに対する、顧客側と開発側との合意形成に利用
  - 顧客と開発者との一種の契約書のようなもの

要求仕様書の役割

- 設計、実装、テストといった後の開発の段階で常に参照される
  - どんな機能を実現すれば良いか?
  - どんな制約があるか?
  - ソフトウェアは要求通りに作成されているか?, etc.



要求仕様書の重要性

- 要求仕様書に不備があると…
  - 必要な機能や入出力内容
  - 必要な制約
  - ソフトウェアにとって重要な属性 (セキュリティ, 保守性, etc.)

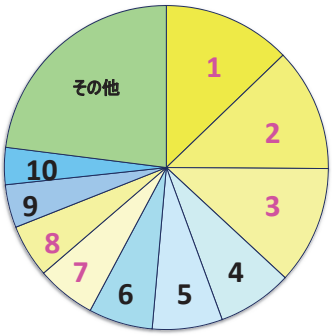
  } 抜けていたり、不明確だと..?

- 顧客に確認 → 手戻りの発生による時間のロス
- 開発側で判断して開発続行
  → 顧客のニーズとは合わないソフトウェアができあがる可能性
  → 顧客に受け入れてもらえず、時間やコストのロスや顧客の不満

顧客のニーズを漏れなく正確に要求仕様書に記述することが重要

参考: ソフトウェア開発プロジェクトが失敗する原因[1]

1. ユーザの参加不足
2. 要求および仕様が不完全
3. 要求および仕様がよく変更になる
4. 経営者側からの支援不足
5. 技術的能力の不足
6. リソースの不足
7. 現実的でない期待
8. 不明確な目的
9. 現実的でない納期
10. 全く実績のない新規技術



失敗する原因として、要求分析に関係するものは多い

IEEE Std. 830-1998～ソフトウェア要求仕様書～

IEEE Std. 830-1998[2]とは

- IEEEという国際的な学会が定めた、ソフトウェアの要求仕様に対して推奨される標準・規格
- 要求仕様書のテンプレートと品質特性を規定
  - この構成の内容を含むことで、少なくとも大きなミスや抜けがなく、ある程度の品質が確保された要求仕様書を作成可能
    - 要求仕様書の記述前に、記述内容を十分に収集することも極めて重要
  - 記述された要求仕様書を、品質特性の観点でチェックすることで、完成度の高いものに洗練

規定されている要求仕様書の構成

- 1. はじめに(Introduction)
  - 1.1 目的(Purpose)
  - 1.2 適用範囲(Scope)
  - 1.3 用語・略称の定義(Definitions、acronyms、and abbreviations)
  - 1.4 参考文献(References)
  - 1.5 概要(Overview)
- 2. 概要(Overall description)
  - 2.1 ソフトウェアの全体像(Product perspective)
  - 2.2 ソフトウェアの機能(Product functions)
  - 2.3 ユーザ特性(User characteristics)
  - 2.4 制約条件(Constraints)
  - 2.5 前提と依存関係(Assumptions and dependences)
  - 2.6 要求の割り当て(Apportioning of requirements)
- 3. 要求仕様の詳細(Specific requirements)
  - 3.1 外部インタフェース(External interfaces)
  - 3.2 機能(Functions)
  - 3.3 パフォーマンス要求(Performance requirements)
  - 3.4 論理データベース要求(Logical database requirements)
  - 3.5 設計上の制約(Design constraints)
  - 3.6 ソフトウェアシステムの属性(Software system attributes)
  - 3.7 個別の要求の構成(Organizing the specific requirements)
  - 3.8 追記(Additional comments)
- 4. 補足情報(Supporting information)
  - 4.1 目次と索引
  - 4.2 付録

IEEE Std. 830-1998～品質特性～

品質特性とは?

- 要求仕様書が「良い」ものであるために必要な要素
- 要求仕様書が品質特性を満たすことで
  - 開発側と顧客側の、開発するソフトウェアに対するイメージに違いがなくなる
  - 開発されたソフトウェアについての検証が容易になる
  - 要求の変更もしやすくなる
  - etc.

要求仕様書の第3章について、品質特性を満たすよう記述することが重要

要求仕様書の品質特性[2][6]

- 妥当であること (Correct)
- あいまいでないこと (Unambiguous)
- 完全であること (Complete)
- 矛盾がないこと (Consistent)
- 重要度と安定度のランク付けがされていること (Ranked for importance and/or stability)
- 検証可能であること (Verifiable)
- 変更が容易であること (Modifiable)
- 追跡可能であること (Traceable)

妥当であること

- 要求仕様書に記述されている要求が、ソフトウェアが満たすべき事柄と一致していること
  - 要求仕様書に記述されている要求が正確に記述されていることを示す特性
    - 顧客のニーズを反映しているか
    - 関連する他のシステムとの整合性や、適用すべき標準・規格を満たしているか, etc.
  - この品質特性を満たしていない = 顧客のニーズが正確に記述されていない
    - 顧客のニーズとは合わないソフトウェアが完成する

➢ 顧客による確認
➢ 関連する文書等との比較
により、できるだけ正確に記述することを目指す

あいまいでないこと

- 要求仕様書に記述されている要求が、ただ一通りに解釈できること
  - この品質特性が満たされていない要求が存在する
    = その要求に対する解釈が複数通り存在する
    - そのまま開発が進行すると、顧客のニーズとは合わないソフトウェアが完成する

平成29年度 産業日本語研究会・ワークショップ　17

あいまいでないこと～自然言語の特性～

- 自然言語はあいまいなところを完全には排除できない
  - 要求仕様書を読む個人によって、記述された要求の解釈が異なることがある
- 顧客側と開発側で、業務において利用している用語が異なることがある
  - 記述された要求に対する誤解が生じることがある
  - 開発側と顧客側で必ずしも同じバックグラウンドを持っていないため、開発側にとって理解しやすい要求仕様書は、顧客側にとっては理解しにくいこともある

⬇

- 様々な人によるレビュー
- 用語集の作成

により、できるだけあいまいでなくすることを目指す

平成29年度 産業日本語研究会・ワークショップ　18

完全であること

- 顧客のニーズが漏れなく要求仕様書に記述されているかを示す特性
  - この品質特性が満たされていない = 要求仕様書に記述されていない要求が存在する
    - 完成したソフトウェアで
      - 必要な機能が実装されていない
      - 実装された機能が必要な振る舞いをしない
      - 必要な性能を満たしていない

      ということにつながる

➢ 機能や性能、設計上の制限、属性、他のシステムとのやりとり等の重要な要求を全て含んでいるか

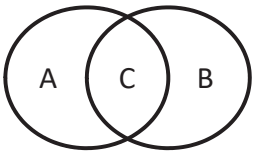
➢ 起こりうる全ての状況における入力に対する応答の定義を含んでいるか

✓ 特に有効・無効な入力値の両方に対する応答が特定されているか

➢ 要求仕様書中の全ての図表へのラベルと参照や、全ての用語の定義、単位の定義をしているかをチェックすることにより、できるだけ漏れをなくすることを目指す

「妥当であること」と「完全であること」の関係[3]

- A: 開発するソフトウェアが満たさなければならない要求
- B: 要求仕様書に記述されている要求
- C: 開発するソフトウェアが満たさなければならない要求で、かつ要求仕様書に記述されているもの(A∩B)



➢ A-C: 満たさなければならない要求で、要求仕様書に記述されていないもの(要求の漏れ)

✓ 品質特性「完全であること」が満たされていない

➢ B-C: 必ずしも満たさなくて良い要求で、要求仕様書に記述されてているもの

✓ 品質特性「妥当であること」が満たされていない

➢ A = B = Cとなったとき、要求仕様書は妥当かつ完全

矛盾がないこと

- 要求仕様書内部で矛盾がないこと
  - 矛盾: ある要求を実現すると、別の要求が実現できない、という状況
    - 記述箇所により、利用するものやその状況が違う
    - 記述箇所により、ソフトウェアの動作が違う
    - 記述箇所により、使う用語が違う, etc.
  - この品質特性を満たしていない = 最終的に達成できる要求と達成できない要求が存在する
    - 設計や実装の段階で、矛盾する要求をどのように扱うか、どちらを実現させるか、混乱が生じる

重要度と安定度のランク付けがされていること(1)

- 各要求について、重要度あるいは安定度を示す指標を明確につけておくこと
  - 要求仕様書に記述された要求は、全てが実現すべき必要な要求とは限らない
    - 必ず実現すべき重要な要求
    - 必ずとは言わないまでもできるだけ実現してほしい要求
    - 余裕があれば実現してほしい要求、etc
  - コストや納期などの面で要求仕様書に記述された要求全てを実現することが難しいこともある
    - ランク付けにより、優先すべき要求が明確になる

重要度と安定度のランク付けがされていること(2)

- 各要求について、重要度あるいは安定度を示す指標を明確につけておくこと
  - この品質特性が満たされていない = 要求仕様書を読む個人によって、各要求の重要度の考え方が異なる
    a. 顧客側と開発側でのコミュニケーションに支障
    b. 設計・実装段階での混乱
    c. 顧客側にとって重要な要求の実現がなし
  - この品質特性が満たされている = 優先的に実現すべき要求が明確になる
    - a. ～ c. の問題の回避
    - コストや納期等の見積もりにも有効

重要度と安定度のランク付けがされていること(3)

- 要求の重要度のランク付け方法(例)
  - 必要なもの (Essential)
    - これらの要求が満たされなければ、ソフトウェアが受け入れられない
  - 条件つき (Conditional)
    - ソフトウェアの価値を高める要求であるが、満たされていなくても受け入れられないという要求ではない
  - オプション (Optional)
    - なくても大きな問題はないが、要求仕様書を超える何かを提供する機会を開発側に与える

※この他、「重要度1～5」のような詳細な順位づけの方法もありうる

重要度と安定度のランク付けがされていること(4)

- 安定度
  - 経験や知識に基づいた要求への将来の変更予測
    - 記述された要求が変更される可能性の有無
    - 予測される変更の回数

    などを明確にしておく

検証可能であること

- ソフトウェアが要求を満たしているかどうかを機械的に確認可能であること
  - 妥当なコスト内、妥当な時間内で、人またはコンピュータがチェックできる手法が存在すること
  - 一般的に、あいまいな要求は検証可能でない

変更が容易であること(1)

- 要求へのあらゆる変更が容易に、完全に、一貫して行えるような構成になっていること
  - ソフトウェア開発の進行中であっても、要求は随時変更
    - ビジネス環境の変化やソフトウェアを稼動させる環境の変化のため
    - 要求の見落としや要求の記述ミスのため, etc.
- 要求の変更にあたっての重要事項
  - 変更対象の要求がどこに書かれているかを容易に発見できること
  - 変更対象の要求に依存する要求(併せて変更が必要となる要求)ができるだけ少ないこと

変更が容易であること(2)

- この品質特性を満たすための重要事項
  - 目次や索引、明確なクロス・リファレンスが整備され、使いやすい構造であること
  - 冗長でないこと
    - 例えば、同じ要求が要求仕様書内で複数箇所に記述されていないこと
  - 他の要求と混ざらず、各要求を独立・分離して表現していること
    - 要求が互いに依存しないこと
    - 互いに依存する要求が多くなると
      - 設計や実装、テスト部分の変更も多くなる
      - 見落としやミスの原因に
      - コスト超過や納期遅延などの原因の原因に

変更が容易であること(3)

- 冗長性
  - 冗長であること自体はエラーにはならないが、エラーに容易に結びつく可能性
  - 冗長であることは要求仕様書をより読みやすくすることも多いが、文書がアップデートされたときに問題が起こる可能性
    - 例えば、複数箇所に記述されている事柄のうちの1つだけを変更するなど(矛盾が存在することに)
  - 冗長性が必要な場合には、明確なクロス・リファレンスを含むこと

平成29年度 産業日本語研究会・ワークショップ　29

追跡可能であること(1)

- 個々の要求の起源が明確で、将来の開発で改良された文書等との対応付けがとれること
  - ソフトウェア開発は、要求仕様書をもとにして設計・実装・テストへと進行
    - 開発の段階ごとに様々な文書が作成
    - 設計・実装・テストの各段階で作成される文書も、必要に応じて随時変更
- この品質特性が満たされていると
  - 要求仕様書に記述された要求から各種文書の内容へたどることができる
  - 各種文書内の記述から要求仕様書に記述された要求にさかのぼることができる

以下を容易に確認することができる
- 要求の変更が各種文書のどの部分に影響するか
- 各種文書の変更によってどの要求が影響を受けるか

平成29年度 産業日本語研究会・ワークショップ　30

追跡可能であること(2)

- 2種類の追跡可能性が存在
  - 後方追跡可能性 (Backward traceability)
    - 開発が始まる以前の段階への参照
      (要求仕様書→要求のソースとなる文書への参照)
    - 各要求が、以前に作成された文書においてそのソースを明確に参照できること
  - 前方追跡可能性 (Forward traceability)
    - この要求仕様書をもとにして作成された全ての文書への参照
      (要求仕様書→各種文書への参照)
    - 要求仕様書内の各要求が、一意の名前や参照番号を持つこと

参考: IEEE Std 830-1998以外の品質特性[4]

- 達成可能であること
  - 要求仕様書に記述された全ての要求を満たすソフトウェアを構築できること
    - 技術的に不可能な要求が含まれていない
    - 矛盾する要求が存在しない, etc.
- 注釈がつけられていること
  - 他の要求との関連を含めて要求の特色がわかりやすく記述されていること
    - 要求の起源や、誰がその要求の実現に責任を持つか、その要求の主要な顧客は誰か、など、要求に関係する様々な情報

要求仕様書とソフトウェアの品質の相関

ソフトウェアの品質

- ソフトウェア開発: ソフトウェアの品質を高める必要
  - 開発のコストやスケジュールを考えると、あらゆる品質を高めるのは難しい
  - 顧客の特性や利用状況、ソフトウェアの特性などにより、重要な品質は異なる

ソフトウェアの品質特性(ISO/IEC 9216 (JIS X 0129-1994))[5]*
- 機能性
- 信頼性
- 使用性
- 効率性
- 保守性
- 移植性

要求仕様書の品質特性
- 妥当であること
- あいまいでないこと
- 完全であること
- 矛盾がないこと
- 重要度と安定度のランク付けがされていること
- 検証可能であること
- 変更が容易であること
- 追跡可能であること

品質特性同士の関連は?
- 要求仕様書のどの品質特性を満たせば、ソフトウェアのどの品質特性を満たせるかが予測できれば、要求仕様書の記述時に、特定の品質特性に注力することができる[3]

*現在は、新しくISO/IEC 25010という規格に置き換わっている

相関関係の検討

- 情報処理学会ソフトウェア工学研究会要求工学ワーキンググループにより検討
  - ワーキンググループのメンバー全員が、相関表に◎(強い相関あり)、○(相関あり)、△(弱い相関あり)、なし(相関なし)を記入して持ち寄り
    - 相関表: ソフトウェアの品質特性を列、と要求仕様書の品質特性を行とする表
  - ◎・○・△・なしの理由を議論し、修正を加え、得点化して平均
  - 平均をもとに、◎・○・△・なしに再度変換

相関表の利用

- 重視するソフトウェアの品質特性を決定し、この相関表をカスタマイズ
  - 開発対象のソフトウェアにより相関表は異なる

	機能性				
	合目的性	正確性	相互運用性	標準適合性	セキュリティ
妥当であること	○	◎	△	△	○
あいまいでないこと	○	◎	△	△	○
完全であること	○	◎	△	△	○
矛盾がないこと	△	△	△	△	○
ランク付けがされていること	◎	なし	なし	なし	なし
検証可能であること	△	○	△	△	△
変更が容易であること	計測しやすい	計測しやすい	計測しやすい	計測しやすい	計測しやすい
追跡可能であること	計測しやすい	計測しやすい	計測しやすい	計測しやすい	計測しやすい

[3]

※条件によって相関が変わることもあるので、注釈がついている項目も多い

参考文献

- [1] Standish Group. CHAOS Report, 1994.
- [2] IEEE, Recommended Practice for Software Requirements Specifications, Std. 830-1998, 1998.
- [3] 要求工学ワーキンググループ: 要求定義で困ってませんか？ 要求仕様書の品質に関する研究成果報告, http://www.selab.is.ritsumei.ac.jp/~ohnishi/RE/rewg-tr1v2.pdf, 2007.
- [4] Davis, A. M., Just Enough Requirements Management: Where Software Development Meets Marketing, Dorset House, 2005.
- [5] International Standard ISO/IEC 9126-1. Software engineering - Product quality - Part 1: Quality model, 2001.
- [6] 妻木, 白銀: 要求工学概論, 近代科学社, 2009

－禁無断転載－

平成 29 年度

産業日本語研究会 報告書

「産業日本語」
Technical Japanese

平成 30 年 3 月
一般財団法人 日本特許情報機構 特許情報研究所
東京都江東区東陽 4 丁目 1 番 7 号
TEL 03-3615-5511