

V 産業日本語研究会ワークショップ講演

令和元年 12 月 13 日に開催した産業日本語研究会ワークショップでは、関連研究として、難波英嗣氏（中央大学理工学部教授）より、特許オントロジーを利用した技術動向分析の自動化に向けた「特許オントロジーの自動構築とその応用」に関する講演をいただいた。

また、Japio の清藤弘晃氏（特許情報研究所 調査研究部長）より、「ライティングマニュアル周知活動のご紹介： 産業日本語の考え方と NMT」と題する講演があり、ニューラル機械翻訳を前提とした特許ライティングマニュアル各項目の整理に関する報告があった。

以下に上記 2 件の講演内容を掲載する。

1. 「特許オントロジーの自動構築とその応用」

特許オントロジーの 自動構築とその応用

2019年12月13日

中央大学理工学部

難波英嗣

1

特許オントロジーとその応用

- 特許オントロジーとは？

特許検索や特許を対象にした高度な言語処理を行なうための有用な情報源

- 特許オントロジーを利用した技術動向分析の自動化に向けて

- サーベイ論文作成支援・技術動向分析に関するこれまでの研究
- なぜ特許オントロジーなのか
- 特許オントロジーの生成

2

サーベイ論文作成支援・技術動向分析のポイント

1. 特定分野の論文と特許の自動収集
2. 重要論文・特許の選定
3. 分析で必要となる観点の明確化
4. 情報の提示方法

3

サーベイ論文作成支援(1996)

研究の目的：

ある分野の論文集合からサーベイ論文を自動的に生成する

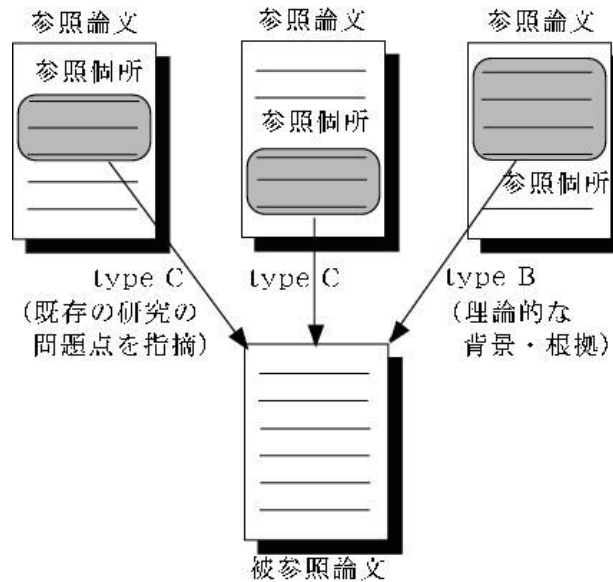
サーベイ論文：複数論文の要約

- 各論文の概要を並べて表示するだけでは不十分
- 論文間の関係(類似点や相違点など)を明らかにする必要がある

4

4

論文間の引用情報



5

引用情報

- (1) 近年、数多くの電子化された論文がWWW上から入手可能になってきている。
- (2) WWW上の論文データを用いた論文データベースとしては、ResearchIndex[1]やCora[2]ある。
- (3) **しかし**、これらは英語で書かれた論文のみを対象にしているため、網羅的なサーベイを行うのに十分であるとは言えない。
- (4) **本研究では**、WWW上にある日英論文データを収集し、より網羅的な論文データベースを自動的に構築する。

6

引用の理由（引用タイプ）の分類

- **Type C**: 他の論文の理論や手法等の問題点を指摘する引用
- **Type B**: ある理論を提案する際、根拠となる引用
- **Type O**: Type B, Type C以外の引用

7

引用情報の抽出

- (1) 近年、数多くの電子化された論文がWWW上から入手可能になってきている。
- (2) WWW上の論文データを用いた論文データベースとしては、ResearchIndex[1]やCora[2]ある。
- (3) **しかし**、これらは英語で書かれた論文のみを対象にしているため、網羅的なサーベイを行うのに十分であるとは言えない。
- (4) **本研究では**、WWW上にある日英論文データを収集し、より網羅的な論文データベースを自動的に構築する。



Type C

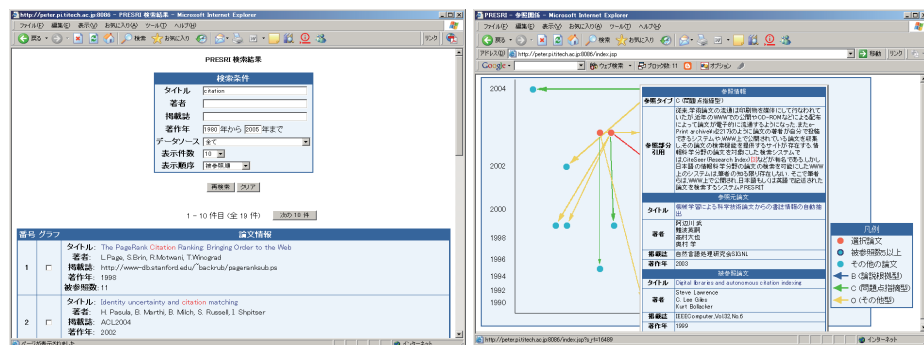
8

引用論文データベースPRESRI(2002-2003)



IPA未踏ソフトウェア創造事業

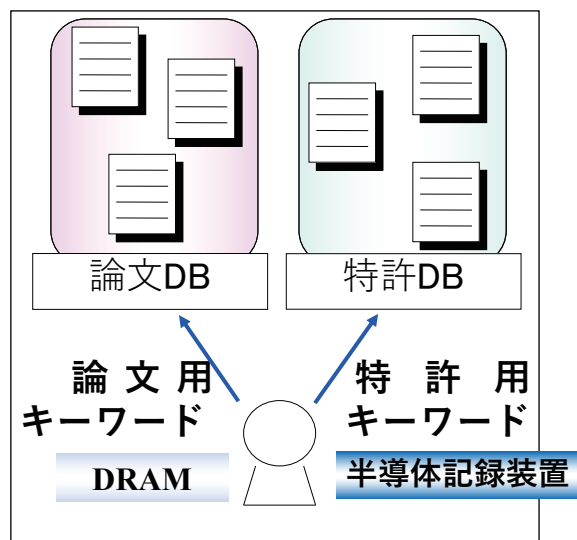
Web上のPostscript及びPDF形式の日英論文データ約
8万件を収集して構築した引用論文データベース
(<http://www.presri.com>)



9

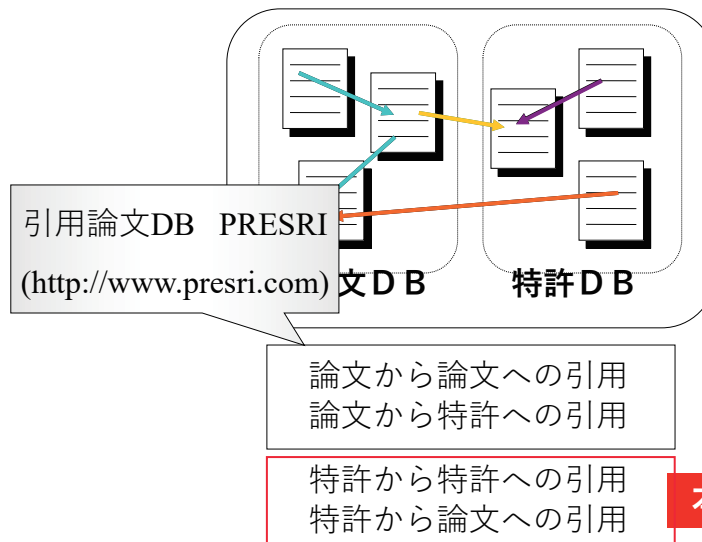
特許と論文データベースの統合的検索システムの構築
および技術動向分析

NEDO



10

引用関係の解析



11

PRESRI 検索結果 - Mozilla Firefox

http://192.168.67.128:8086/index.jsp

Google 検索 PageRank ASX チェック 登録 オプション

PRESRI 検索結果

表示順序 検索照準

両方再検索 論文検索に変更 特許検索に変更 クリア

1 - 10 件目 (全 80 件) 次の 10 件

番号	グラフ	論文情報/特許情報
1		タイトル: フロッピーディスク装置のすべて 著者: 高橋昇司 掲載誌: 著作年: 1989 被参照数: 0
2		特許番号: 特開平5-174371 発明の名称: 磁気ディスク装置及びその構成材料決定方法 発明者: 栗田 典之 田子 一農 小林 金也 出願者: 株式会社日立製作所 小川 勝男 出願日: 1990-01-01
3		特許番号: 特開平5-197934 発明の名称: 磁気ディスク装置のヘッドアーム 発明者: 小崎 雅弘 出願者: 富士通株式会社 井島 藤治 (外1名) 出願日: 1990-01-01

完了

12

12

出力例

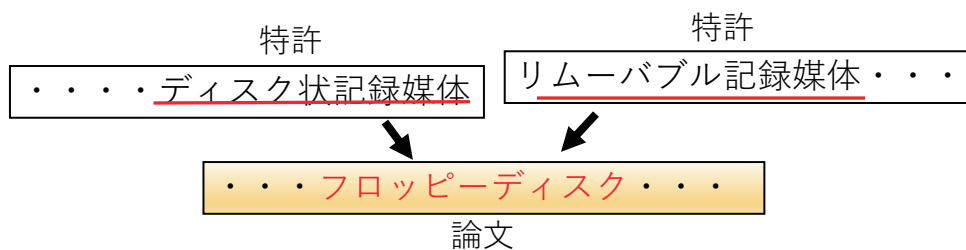
入力用語	ワードプロセッサ	DRAM
出力用語	16 文書編集装置	119 半導体記憶装置
	14 文書作成装置	92 ダイナミックランダムアクセスメモリ
	13 文字処理装置	92 半導体メモリ
	13 文書作成支援装置	91 集積メモリ
	12 共起関係計算装置	90 キャッシュメモリ
	12 定型文書処理装置	90 シリアルアクセスメモリ
	3 ヘルプ機能表示システム	90 ダイナミックメモリ
		90 メモリ

13

13

引用関係を用いた特許用語の収集（1）

フロッピーディスク”の特許用語を収集したい。

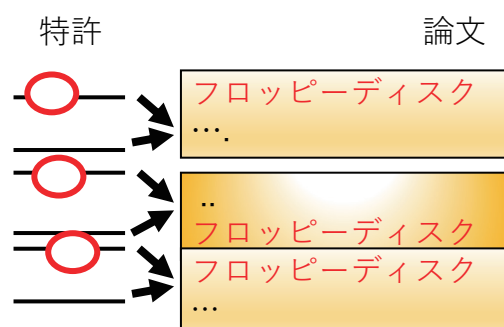


14

14

引用関係を用いた特許用語の収集（２）

“フロッピーディスク”の特許用語を収集したい。

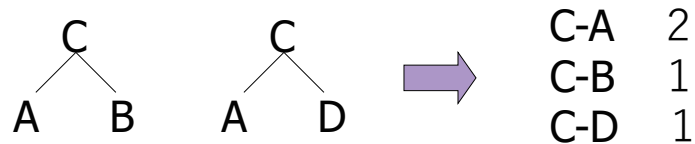


15

15

特許用語シソーラス

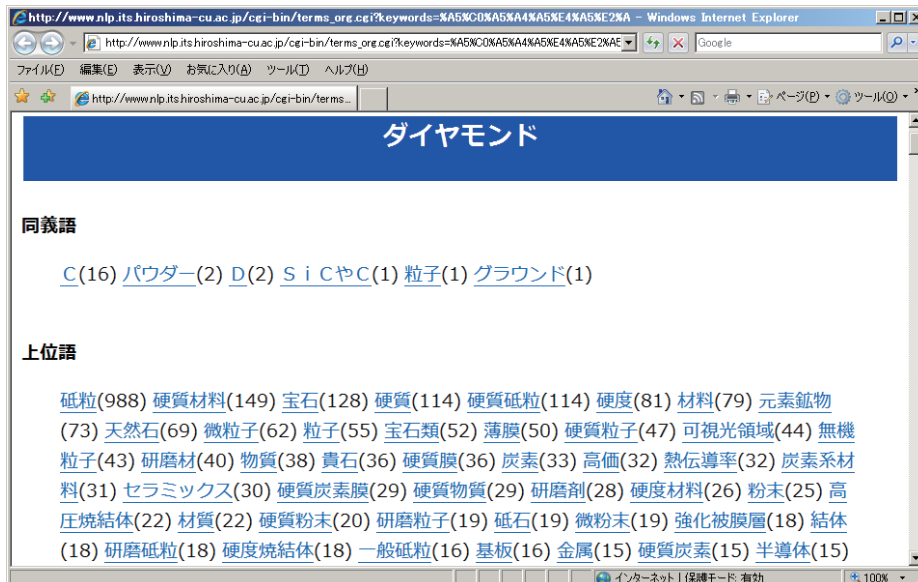
公開公報中の「**A**や**B**等の**C**」、「**A**や**D**等の**C**」
という表記から、用語の上位・下位関係を抽出



16

16

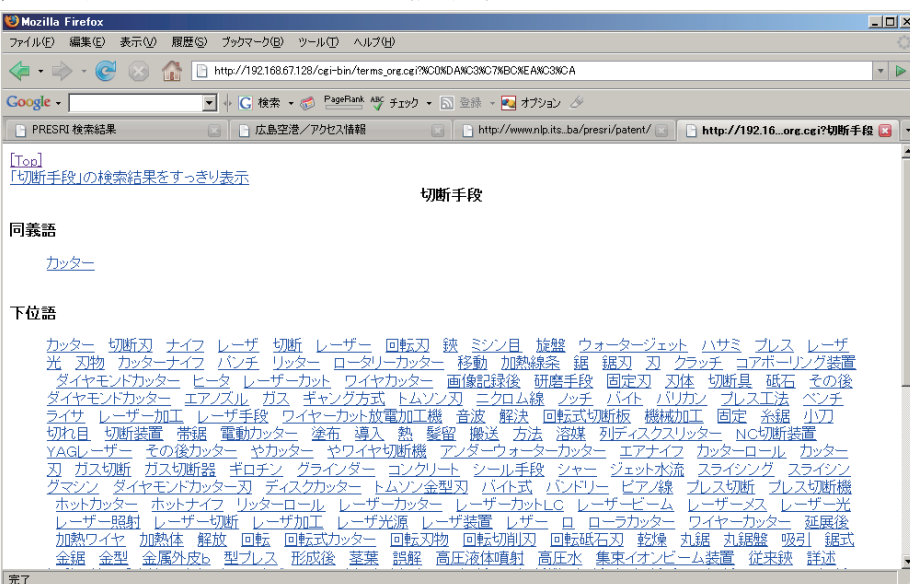
特許シソーラス検索システム



17

17

特許シソーラス検索システム



18

18

NTCIR特許マイニングタスク(2007-2010)

目的

特許と論文を対象にした検索や技術動向分析など様々な
目的に利用可能な言語処理技術の開発

技術動向マップの自動生成

	効果 1	効果2	効果 3
技術 1	[AAA 1993] [US Pat. XX-XXX]		[BBB 2002]
技術 2		[DDD 2001]	
技術3	[CCC 2000]	[US Pat. YY-YYY]	[US Pat. ZZ-ZZZ] [US Pat. WW-W]

19

19

技術動向マップの生成

ある分野の論文と特許から技術動向マップを生成するには、まず、
その分野の論文と特許を収集する必要がある。

- 特許には国際特許分類などの分類コードが付与されている
→論文にも国際特許分類を付与
- 技術動向を要素技術とその効果でまとめる
- タスク1：学術論文の国際特許分類への自動分類
- タスク2：学術論文と特許からの要素技術と効果の抽出

20

論文と特許からの要素技術と効果の抽出

- 論文および特許を対象（一部抜粋）

「PM磁束制御用コイルを設けて閉ループフィードバック制御を適用するため、電気損失を最小化できる。」

↓

PM磁束制御用コイルを設けて<TECHNOLOGY>閉ループフィードバック制御</TECHNOLOGY>を適用するため、
<EFFECT><ATTRIBUTE>電気損失</ATTRIBUTE>を<VALUE>最小化</VALUE></EFFECT>できる。

21

21

特許マイニングタスクタスク1参加者数(NTCIR-7)

	日本	日本以外のアジア	欧州	北米
大学	3	4	0	2
企業	2	0	1	0

22

22

タスク1(論文の国際特許分類への自動分類)参加者(NTCIR-8)

	日本	日本以外のアジア	欧米
大学	1	3	1
企業	1	0	0

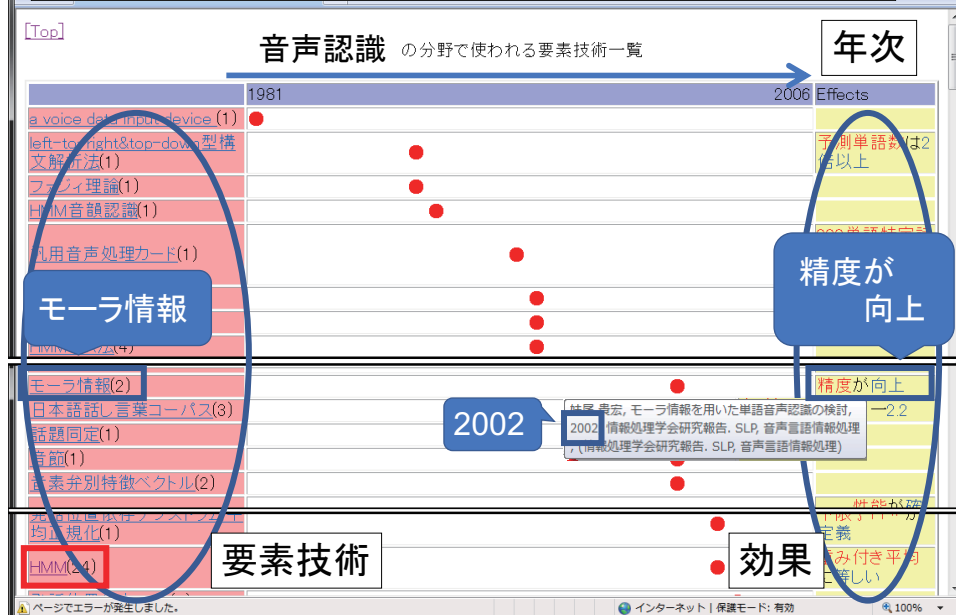
タスク2(論文と特許からの要素技術と効果の抽出)(NTCIR-8)

	日本	日本以外のアジア	欧米
大学	2	3	1
Company	3	0	0

23

23

技術動向分析システムの動作例



24

24

新情報の追加による サーベイ論文の自動更新(2013-2014)

学術情報量の爆発的増加

すべての関連論文に目を通すことができない

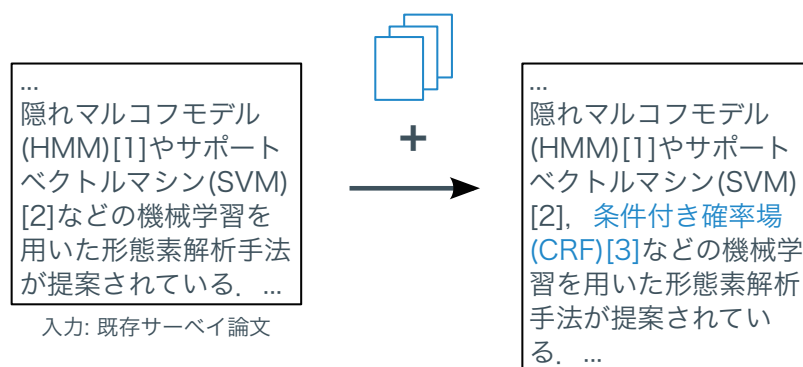
→サーベイ論文や専門書籍の必要性

しかし、ある分野のサーベイ論文や専門書籍が執筆されてかなりの時間が経過していると、それらに最新の研究動向情報が含まれないという問題がある。

サーベイ論文や専門書籍の自動更新

25

サーベイ論文更新のイメージ



26

26

複数テキスト要約研究の中での本研究の位置づけ

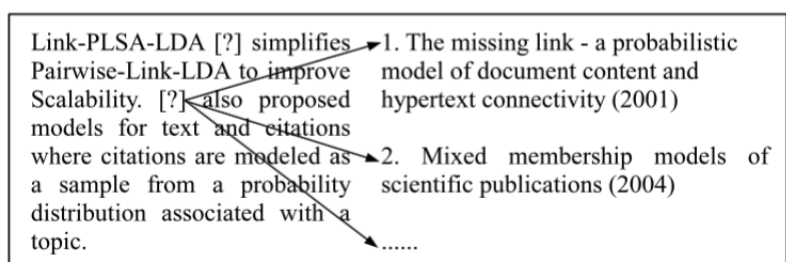
- サーベイ論文や専門書籍のような長いテキストをスクラッチから生成するのは、現在のNLP技術では困難
- 複数の論文から数文程度の要約を出力する研究はある(Radevら)
- Sentence orderingなどの研究もあるが、複数の節、章から構成される文書の出力は実現には程遠い。
- 本研究の枠組みは、Update summarizationの一種

27

27

関連研究

Context-aware citation recommendation [He 10]

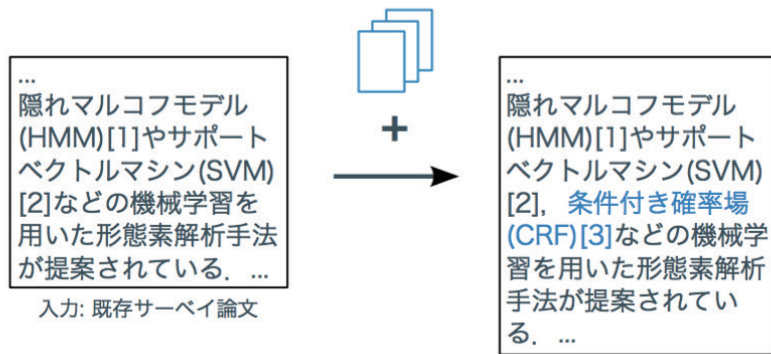


コンテキストに対して引用論文を推薦

28

28

言及されていない、新たに追加すべき論文



29

29

サーベイ更新の手順

1. 追加すべき論文の収集
2. 古いサーベイ論文の更新

30

30

実験に用いるデータ

入力: 既存サーベイ論文 <- 旧版

正解 <- 新版に追加された論文

書籍タイトル	旧版	新版
"information Retrieval"	1 st ed. (1998)	2 nd ed. (2004)
"Modern Information Retrieval"	1 st ed. (1999)	2 nd ed. (2011)
"Speech and Language Processing"	1 st ed. (2000)	2 nd ed. (2009)
"Modern Operating Systems"	2 nd ed. (2001)	3 rd ed. (2007)

章単位での検索, 評価

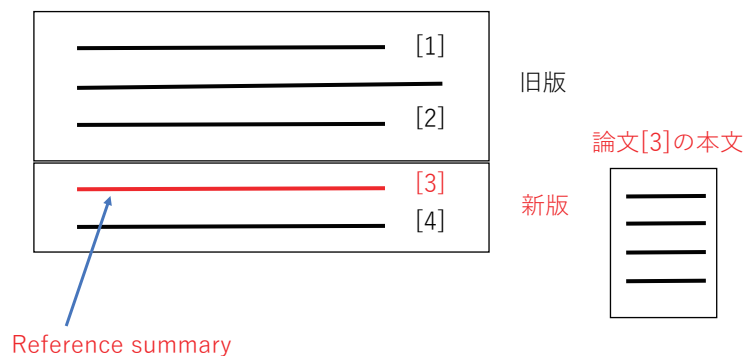
CiteSeer全データ(2012年9月)

31

31

実験条件

- 入力: 新版に追加された論文の本文
- 正解: 新版に追加された論文に関する引用文 (文単位で評価する)



32

32

得られた知見(新情報の追加によるサーベイ論文の自動更新)

- 追加すべき論文の候補(同じ分野の論文)を収集するのはそれほど難しくないが、その中から重要論文を選定するのが非常に難しい。
- 追加すべき論文は発表されてからあまり時間が経っていないため、被引用数による選定が使えない。
- 追加すべき文を抽出するのも容易ではない。

33

得られた知見(新情報の追加によるサーベイ論文の自動更新)

- 追加すべき論文の候補(同じ分野の論文)を収集するのはそれほど難しくないが、その中から重要論文を選定するのが非常に難しい。
- 追加すべき論文は発表されてからあまり時間が経っていないため、被引用数による選定が使えない。
- 追加すべき文を抽出するのも容易ではない。

34

特許や論文の新規性

- 課題としての新しさ
- 解決手段としての新しさ

課題や解決手段が新しいかどうかを計算機で自動的に判断するためには、過去にどんな課題や解決手段が存在したのかをきちんと整理しておく以外にないのでは？



特許オントロジーの自動構築

35

日英特許データベースからのシソーラスの自動構築(2011)

定型表現を用いた上位、下位概念の獲得

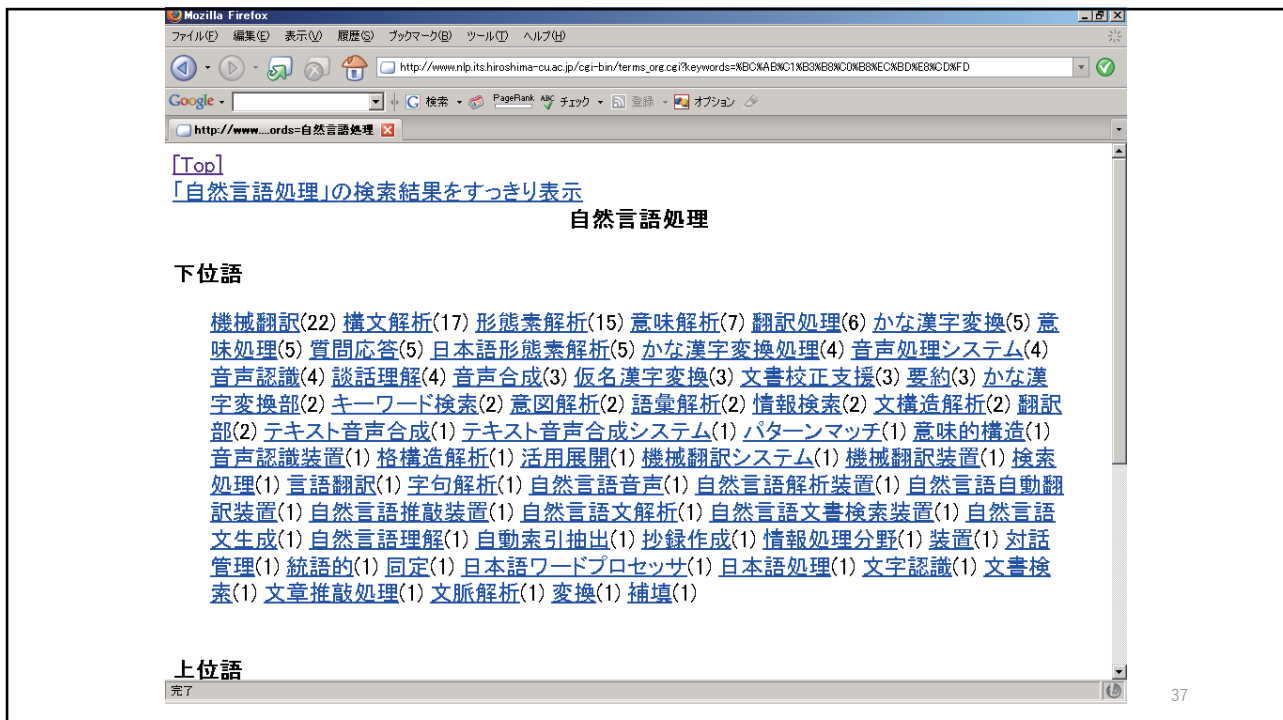
「などの」「等の」の2種類の定型表現に着目

[名詞句]((や|、|,) [名詞句])*等の[名詞句]

アルミニウムや鉄などの金属

公開公報から、これらの定型表現を含む文を収集

36



37

自動抽出された上位、下位概念の誤り例

- 「パソコンなどのOA機器」→「OA機器＞パソコン」○
- 「パソコンなどのキーボード」→「キーボード＞パソコン」×

上記2つ目の事例のような部分－全体関係が誤って上位、下位関係として抽出されることがある。

解決策

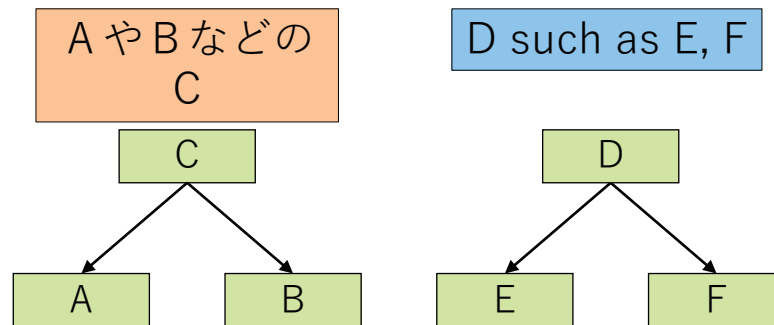
「などの」を「の」に言い換えて成立するかどうか調べる

「パソコンのOA機器」 X → 上位一下位

「パソコンのキーボード」 ○ → 部分－全体

しかし、この方法は意外とうまくいかない。

日英特許DBからの上位、下位 概念の獲得

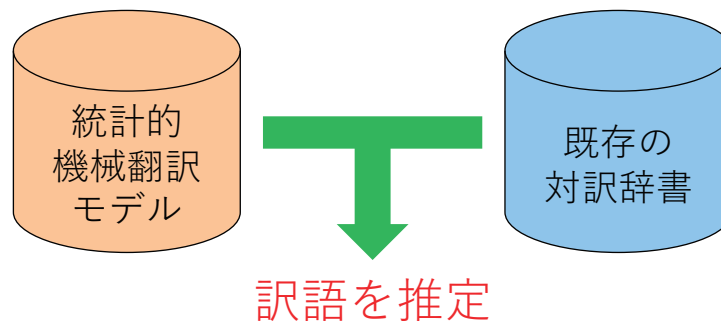


39

39

日英の用語間の対応付け

- 関連研究 [森下他 2010]
- 専門用語の訳語を自動推定する

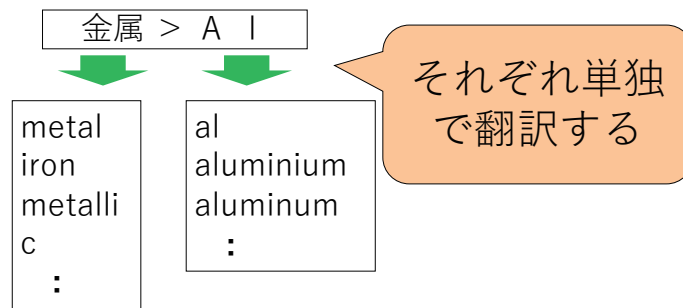


40

40

日英用語対の作成

1. 上位、下位概念の翻訳

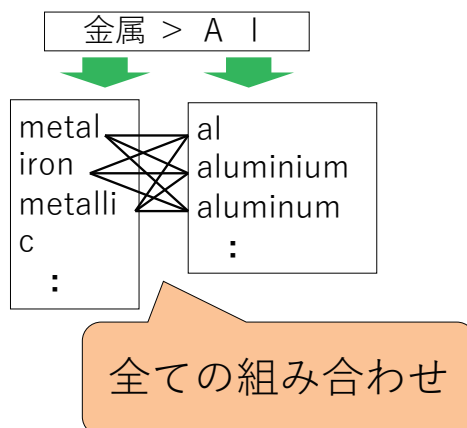


41

41

日英用語対の作成

2. 上位、下位概念の候補を作成



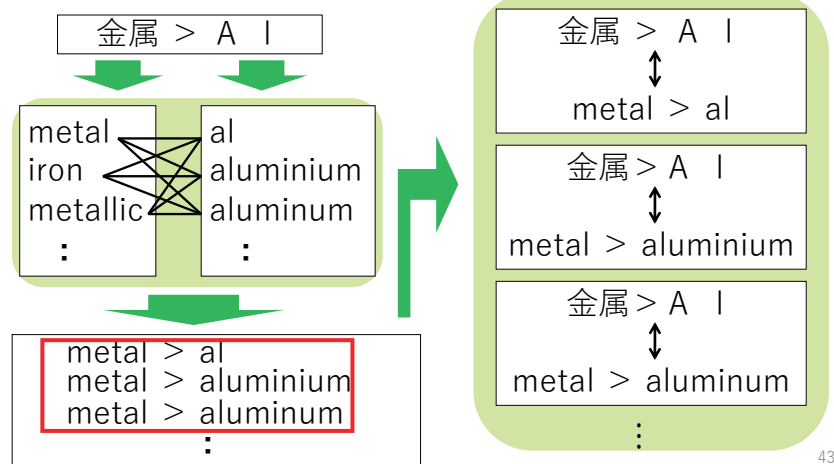
42

42

日英用語対の作成

対応付け

3. 上位、下位概念の対応付け



43

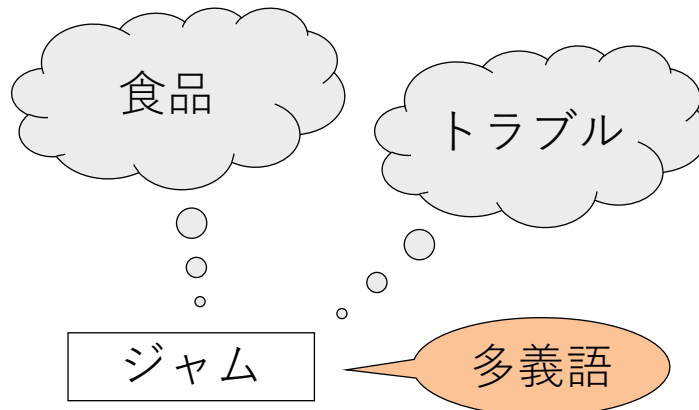
43

得られた知見(日英特許データベースからのシソーラスの自動構築)

- 78%程度の精度で日英の対応付けができる。
- ただし、米国特許DBから抽出される上位・下位関係の数がそれほど多くないため、対応付けられたデータの数そのものが多い。

44

上位・下位関係抽出に関するその他の問題点



45

45

特許検索履歴を用いたシソーラスの自動構築(2011)

[G06F15/167@B+G06F15/16,470+G06F15/16,640@A+G06F15/177,(672@H+678@A)+? 6 4 0 @ A +? 6 7 2 @ H +? 6 7 8 @ A +BB04+JJ(07+09+37+38+45+46+48)]*[サービスプロ/TX+??サービスプロ+SVP/TX+??SVP+監視装置/TX+??監視装置]*[障害/TX+??障害+故障/TX+??故障+異常/TX+??異常+エラ/TX+??エラ] -¥1-¥2327件二次検索クラス+同報+ブロードキャスト+サービス+一斉+通知*[障害+故障+異常+エラ]+セル+サービスプロセッサ+SVP+[ストップ+停止+切離+切り離し+離]*[障害+故障+異常+エラ]

46

クエリログの例

(1) すべて同義語

- ジャガイモ+じゃがいも+バレイショ+馬鈴薯+ばれいしょ

(2) 「茸」「きのこ」「キノコ」は「植物」の下位語

- 植物+茸+きのこ+キノコ

(3) 属性と属性値が混在

- 解像度+デグレード+低下+劣化

論理和で結合されている一連の用語がすべて同義語である可能性の高い個所(上記の場合は例1)を見つけ、その個所から同義語を抽出

47

[\[Top\]](#)
[「自然言語処理」の検索結果をすっきり表示](#)
自然言語処理

下位語

機械翻訳(22) 構文解析(17) 形態素解析(15) 意味解析(7) 翻訳処理(6) かな漢字変換(5) 意味処理(5) 質問応答(5) 日本語形態素解析(5) かな漢字変換処理(4) 音声処理システム(4) 音声認識(4) 談話理解(4) 音声合成(3) 仮名漢字変換(3) 文書校正支援(3) 要約(3) かな漢字変換部(2) キーワード検索(2) 意図解析(2) 語彙解析(2) 情報検索(2) 文構造解析(2) 翻訳部(2) テキスト音声合成(1) テキスト音声合成システム(1) パターンマッチ(1) 意味的構造(1) 音声認識装置(1) 格構造解析(1) 活用展開(1) 機械翻訳システム(1) 機械翻訳装置(1) 検索処理(1) 言語翻訳(1) 字句解析(1) 自然言語音声(1) 自然言語解析装置(1) 自然言語自動翻訳装置(1) 自然言語推敲装置(1) 自然言語文解析(1) 自然言語文書検索装置(1) 自然言語文生成(1) 自然言語理解(1) 自動索引抽出(1) 抄録作成(1) 情報処理分野(1) 装置(1) 対話管理(1) 統語的(1) 同定(1) 日本語ワードプロセッサ(1) 日本語処理(1) 文字認識(1) 文書検索(1) 文章推敲処理(1) 文脈解析(1) 変換(1) 補填(1)

上位語

完了

48

48

(関係2) 効果表現リストの作成 [NTCIR-8]

PM磁束制御用コイルを設けて閉ループフィード
バック制御を施すため、電力損失を最小化できる



PM磁束制御用コイルを設けて<TECHNOLOGY>閉ル
ープフィードバック制御</TECHNOLOGY>を施すため、
<EFFECT><ATTRIBUTE>電力損失</ATTRIBUTE>を
<VALUE>最小化</VALUE></EFFECT>できる

49

49

(関係2) 効果表現リストの作成 [Nanba 2010]

公開公報10年分を解析し、抽出した効果表現を頻度順に出力

22393	信頼性	向上	
21713	信頼性	高い	
17362	構成	簡単	
16870	生産性	向上	
11283	作業性	向上	
10522	操作性	向上	
10376	コスト	低減	計 96,435対
10175	製造コスト	低減	(頻度 属性 属性値)

50

同義関係の抽出

検索履歴

(電力 + 電源) * (停止 + 切断 + オフ + 遮断)

効果表現リスト

電源 - オフ

電源 - 切断

電力 - 停止

51

同義関係の抽出

検索履歴

ブロック1

ブロック2

(電力 + 電源) * (停止 + 切断 + オフ + 遮断)

効果表現リスト

電源 - オフ

電源 - 切断

電力 - 停止

ブロック1と2の間は
属性 - 属性値の関係

52

同義関係の抽出

検索履歴

ブロック1

ブロック2

(電力+電源) * (停止+切断+オフ+遮断)

新たに得られる知識

効果 : 電力一切断、電力一オフ、電力一遮断
 電源一停止、電源一遮断

53

同義関係の抽出

検索履歴

ブロック1

ブロック2

(電力+電源) * (停止+切断+オフ+遮断)

新たに得られる知識

同義関係 : 電力一停止、電力一切断、
 電力一オフ、電力一遮断、電源一停止、
 電源一切断、電源一オフ、電源一遮断

54

実験

データ

- 検索履歴データ：一般財団法人工業所有権協力センター(IPCC)において2001年度から2008年度までの間に知財の専門家が作成した検索報告書に記載された検索式約61万件
- 特許全文データ：公開公報1993-2002年(3,496,252文書、94.5GB)

55

実験結果

(評価1) 属性－属性値の抽出
89.3% (109/122)

(評価2) 同義関係の抽出
81.9% (100/122)

56

特許分類コード体系に基づくオントロジーの構築(2014)

オントロジー構築のためのFタームの利用

Fタームとは、特許を目的、利用分野、材料といった様々な観点から分類することを目的として日本国特許庁が構築した特許の分類体系のひとつ

Fタームの構造そのものがオントロジーに近い体系→Fタームの体系をオントロジーの構築に流用

これに、テキストデータベースから自動構築された各種辞書をマッピングすることで、特許との親和性を保持しながら、学术论文など他のジャンルの文献にも利用可能なオントロジーの構築を目指す

57

テーマ“5B091(機械翻訳)”のFタームコードの例

- | | |
|------|------------------|
| BA11 | ・ 翻訳方式 |
| BA12 | ・ ・ 直接翻訳 |
| BA13 | ・ ・ 間接翻訳 |
| BA14 | ・ ・ ・ トランスファー方式 |
| BA15 | ・ ・ ・ ・ 意味解析 |
| BA16 | ・ ・ ・ ・ ・ 文脈解析 |
| BA17 | ・ ・ ・ ・ ・ ピボット方式 |

ドット表記には、上位・下位や属性などの関係が含まれている。その関係を明らかにし、オントロジー構築のための知識抽出を目指す

58

Fタームから得られる知識

BA11	・ 翻訳方式	属性：方式 定義域：機械翻訳 値域：直接翻訳、間接翻訳
BA12	・ ・ 直接翻訳	
BA13	・ ・ 間接翻訳	
BA14	・ ・ ・ トランスファー方式	
BA15	・ ・ ・ ・ 意味解析	
BA16	・ ・ ・ ・ ・ 文脈解析	
BA17	・ ・ ・ ・ ・ ピボット方式	

テーマ“5B091(機械翻訳)”のFタームコードの例

59

Fタームから得られる知識

BA11	・ 翻訳方式	上位：間接翻訳 下位：トランスファー方式、 ピボット方式
BA12	・ ・ 直接翻訳	
BA13	・ ・ 間接翻訳	
BA14	・ ・ ・ トランスファー方式	
BA15	・ ・ ・ ・ 意味解析	
BA16	・ ・ ・ ・ ・ 文脈解析	
BA17	・ ・ ・ ・ ・ ピボット方式	

テーマ“5B091(機械翻訳)”のFタームコードの例

60

Fタームから得られる知識

BA11	・ 翻訳方式	
BA12	・ ・ 直接翻訳	
BA13	・ ・ 間接翻訳	
BA14	・ ・ ・ トランスファー方式	
BA15	・ ・ ・ ・ 意味解析	
BA16	・ ・ ・ ・ ・ 文脈解析	
BA17	・ ・ ・ ピボット方	属性：利用技術 定義域：トランスファー方式 値域：意味解析

テーマ“5B091(機械翻訳)”のFタームコードの例

61

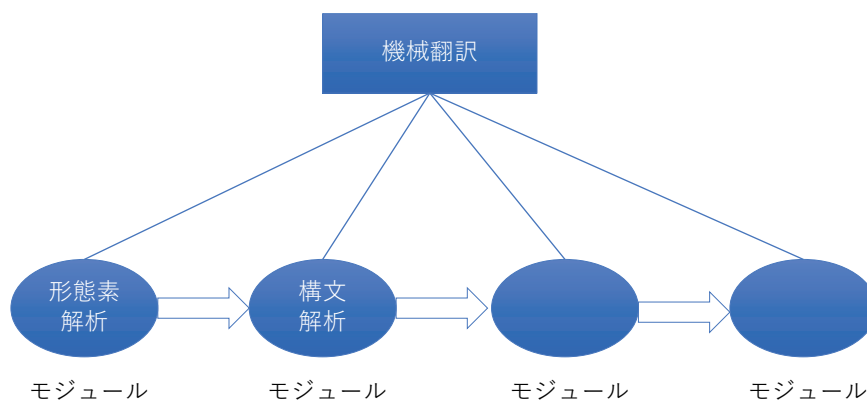
Fタームから得られる知識

BA11	・ 翻訳方式	
BA12	・ ・ 直接翻訳	
BA13	・ ・ 間接翻訳	
BA14	・ ・ ・ トランスファー方式	
BA15	・ ・ ・ ・ 意味解析	
BA16	・ ・ ・ ・ ・ 文脈解析	
BA17	・ ・ ・ ピボット方式	属性：利用技術 定義域：意味解析 値域：文脈解析

テーマ“5B091(機械翻訳)”のFタームコードの例

62

複数手順テキストからの 手順オントロジーの自動構築(2016-現在)



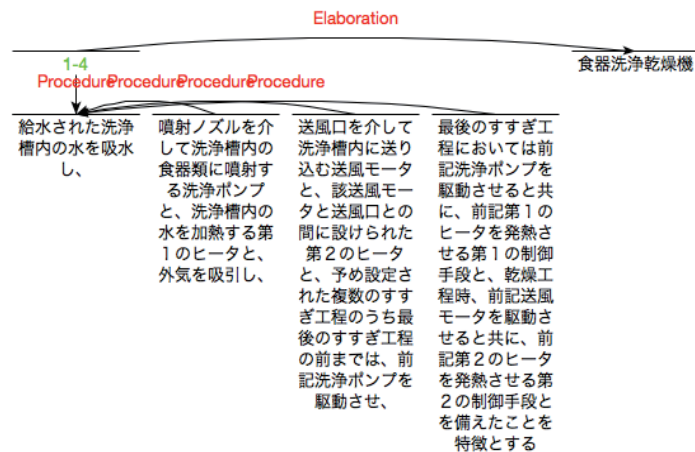
63

手順テキストとは (食器洗浄乾燥機に関する特許の例)

給水された洗浄槽内の水を吸水し⁽¹⁾、噴射ノズルを介して洗浄槽内の食器類に噴射する洗浄ポンプと、洗浄槽内の水を加熱する第1のヒータと、外気を吸引し⁽²⁾、送風口を介して洗浄槽内に送り込む送風モータと、該送風モータと送風口との間に設けられた第2のヒータと、予め設定された複数のすすぎ工程のうち最後のすすぎ工程の前までは、前記洗浄ポンプを駆動させ⁽³⁾、最後のすすぎ工程においては前記洗浄ポンプを駆動させると共に、前記第1のヒータを発熱させる第1の制御手段と、乾燥工程時、前記送風モータを駆動させると共に、前記第2のヒータを発熱させる⁽⁴⁾第2の制御手段とを備えたことを特徴とする食器洗浄乾燥機。

64

「食器洗浄乾燥機」に関する請求項の解析木 [新森 2004]



65

「食器洗浄乾燥機」の検索結果(特許)

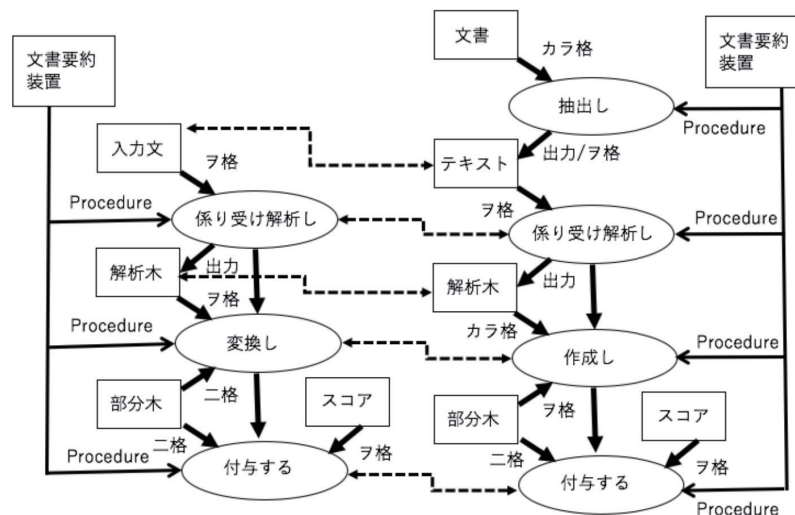
検索結果 1 1 5 件

検索結果 115件

項番	文献番号	発明の名称	筆頭出願人 (登録公報・US和抄は権利者を表示)	発行日	出願番号	出願日	筆頭IPC
1	特開2013-128556	食器洗浄乾燥機	株式会社ハーマン	2013年07月04日	特願2011-278426	2011年12月20日	A47L 15/46
2	特開2013-009706	食器洗浄乾燥機	株式会社ハーマン	2013年01月17日	特願2011-142503	2011年06月28日	A47L 15/46
3	特開2012-110519	食器洗浄乾燥機	株式会社ハーマン	2012年06月14日	特願2010-262462	2010年11月25日	A47L 15/48
4	特開2012-110517	食器洗浄乾燥機	株式会社ハーマン	2012年06月14日	特願2010-262448	2010年11月25日	A47L 15/46
5	特開2012-090838	食器洗浄乾燥機	株式会社ハーマン	2012年05月17日	特願2010-241699	2010年10月28日	A47L 15/42
6	特開2012-045295	食器洗浄乾燥機	株式会社ハーマン	2012年03月08日	特願2010-192337	2010年08月30日	A47L 15/42
7	特開2011-200392	食器洗浄乾燥機	リンナイ株式会社	2011年10月13日	特願2010-069978	2010年03月25日	A47L 15/42
8	特開2011-200391	食器乾燥機	リンナイ株式会社	2011年10月13日	特願2010-069977	2010年03月25日	A47L 19/00
9	特開2010-142322	食器洗浄乾燥機	パナソニック株式会社	2010年07月01日	特願2008-320458	2008年12月17日	A47L 15/42
10	特開2010-142319	食器洗浄乾燥機	パナソニック株式会社	2010年07月01日	特願2008-320455	2008年12月17日	A47L 15/42
11	特開2010-142318	食器洗浄乾燥機	パナソニック株式会社	2010年07月01日	特願2008-320454	2008年12月17日	A47L 15/42
12	特開2008-228889	食器洗浄乾燥機及び	三菱電機株式会社	2008年10月02日	特願2007-070643	2007年03月19日	A47L 15/46

66

手順グラフの対応付け



67

日本語請求項へのタグ付与の例

半導体回路のレイアウトデータに対し、<block id="1" link="7">シミュレーションを施したシミュレーション画像データのパターン線分を膨張する<comp>膨張処理部</comp></block>と、<block id="2" link="7">当該膨張処理が施されたシミュレーション画像データの<comp>パターン内側領域</comp></block>、及び<block id="3" link="7">外側領域の一方にマスクを施す<comp>マスク処理実行部</comp></block>と、<block id="4" link="7">当該マスクが施された<comp>シミュレーション画像データ</comp></block>と、<block id="5" link="7">荷電粒子線装置によって得られた画像を重畳させ、当該マスクが施された領域以外の領域について、輝度信号を検出する<comp>輝度信号抽出部</comp></block>と、<block id="6" link="7">当該輝度変化が所定値を超えた部分の有無の判定、或いは当該所定値を超えた部分の位置情報を抽出する<comp>欠陥抽出部</comp></block>を備えたことを特徴とする<block id="7" link="7">-1</block><head>欠陥検査装置</head></block>。

68

英語請求項へのタグ付与の例

1. A method of `<block id="1" link="-1"><head>optimizing the geometry of a femoral stem of a hip joint prosthesis, the femoral stem</head></block>` comprising

`<block id="2" link="1"><comp>a neck</comp></block>`; and

`<block id="3" link="1"><comp>an anchoring blade</comp>` that is attached to the neck and that tapers towards a distal end with a lateral narrow side comprising a distal straight portion and a proximal arcuate portion corresponding to a curve, a transition between the distal straight portion and said proximal

arcuate portion occurring at an outer lateral point`</block>`; and

`<block id="4" link="1"><comp>said method</comp></block>` comprising `<block id="5 link="4"><proc>optimizing the profile of the curve of said proximal arcuate portion by a process of iterative modeling steps using a series of curves each defined by a path traced by the outer lateral point of the blade on withdrawal of a profile of the stem from a cavity of complementary shape to the stem.</block>`

69

知財工学会設立趣意書

産業の国際競争力強化および経済活性化の観点から、わが国では知的財産の重要性が高まっている。2003年3月には、知的財産の創造、保護および活用に関する施策を作成し、計画的に推進することを目的に、内閣に知的財産戦略本部が設置された。一方、その方法論に関して研究者や技術者が工学的な観点から総合的に議論できる場がこれまでなかった。そこで、知財活動に関する方法論を情報学等の工学的観点から議論できる場を提供し、知的財産の利活用に関する技術を促進することで、知財立国の推進に寄与するため、知財工学会を設立する。

70

特許庁 特許出願技術動向調査を用いた 研究課題の提案

特許庁では、市場創出に関する技術分野、国の政策として推進すべき技術分野を中心に、今後の進展が予想される技術テーマを選定し、特許出願技術動向調査を実施している。

平成30年度

- [三次元計測 \(PDF: 3.272KB\)](#)、[スライド資料 \(PDF: 1.157KB\)](#)
- [電子ゲーム \(PDF: 2.790KB\)](#)、[スライド資料 \(PDF: 2.224KB\)](#)
- [次世代建築技術 \(PDF: 12.449KB\)](#)、[スライド資料 \(PDF: 3.389KB\)](#)

平成29年度

- [超音波診断装置 \(PDF: 2.444KB\)](#)
- [有機EL装置 \(PDF: 1.882KB\)](#)、[正誤表 \(PDF: 145KB\)](#)
- [次世代光ファイバ技術 \(PDF: 1.991KB\)](#)

平成28年度

- [電池の試験及び状態検出 \(PDF: 2.820KB\)](#)
- [移動体用カメラ \(PDF: 7.987KB\)](#)
- [施設園芸農業 \(PDF: 2.172KB\)](#)

71

特許庁 特許出願技術動向調査を用いた 研究課題の提案

しかし、一度作成された調査報告書は、その内容がどんどん古くなる。

↓

調査報告書の自動更新

- 観点ごとに関連論文・特許を自動収集
- 論文・特許集合から観点の候補を自動的に検出
- 文書生成技術を用い、報告書そのものを自動更新

これらを、知財工学会におけるタスクのひとつとして、コンペなどの形式での実施を検討中。

72

まとめ

- サーベイ論文作成支援・技術動向分析の研究事例を紹介
- 特許オントロジーの必要性
- 特許出願技術動向調査を用いた研究課題の提案(知財工学会)

2. 「ライティングマニュアル周知活動のご紹介」

<「ライティングマニュアル周知活動のご紹介」資料>



産業日本語の考え方とNMT (第29回JTF翻訳祭講演Japio担当抜粋版)

令和元年12月

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／知財AI研究センター主幹
清藤 弘晃

「産業日本語」とは

産業日本語

「産業・技術情報を、人に理解しやすく、かつ、機械（コンピュータ）にも処理しやすく表現するための日本語」を意味する造語

- 産業・技術情報を、**客観的かつ正確に伝達**し、機械処理に適した日本語とすることを旨とする
- 「産業日本語」に係る**日本語研究**と、その**研究成果の普及**を推進
- これにより、以下を実現
 - (1) 自然言語処理（翻訳等）における品質と効率の向上
 - (2) 正確かつ円滑な情報発信力の強化
 - (3) 知的生産性の向上

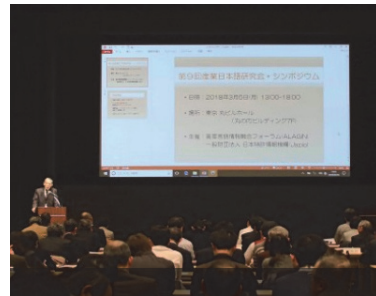


「産業日本語」とは

産業日本語研究会

日本特許情報機構が、高度言語情報融合フォーラム（ALAGIN）と協力して運営・開催している研究会

- 産業分野・科学技術分野における情報発信力や知的生産性・国際競争力の強化に資するような日本語のあり方について総合的な議論を行なうことを目的とする。
- 平成21年度に発足し、1年に1回、産業日本語研究会・シンポジウムを開催。
（今年度は、令和2年3月6日（金）午後第11回のシンポジウムを開催予定。）
- 3つの分科会を開催し、ライティング全般、特許文書、文書作成支援等の研究を継続中。



2

「産業日本語」とは

産業日本語研究会

日本特許情報機構が、高度言語情報融合フォーラム（ALAGIN）と協力して運営・開催している研究会

- 組織体制は以下の通り

顧問：長尾 眞（京都大学名誉教授）
代表：井佐原 均（豊橋技術科学大学）
委員：辻井 潤一（産業技術総合研究所）
橋田 浩一（東京大学）
隅田英一郎（情報通信研究機構）
柏野和佳子（国立国語研究所）
潮田 明（産業技術総合研究所）
清藤 弘晃（日本特許情報機構）



3

特許文書の日本語文

「特許文書」は、**技術文書**であると共に、**法律文書（権利文書）**でもある

- 特許文書は、**高度な正確性**が求められる
- 特許文書の活用場面は海外にも拡大し、**多言語翻訳に耐えられる品質の日本語**を用いることが重要
- しかし、**特許文書に特有の課題**がある
 - ・ 長文で、係り受けが複雑な文が多い
 - ・ 特許独特の表現（言い回し・一般化・抽象化）の多用
 - ・ 一般用語と異なる概念による技術用語の定義 など



4

特許文書の例 (特許請求の範囲 = 特許権の範囲)

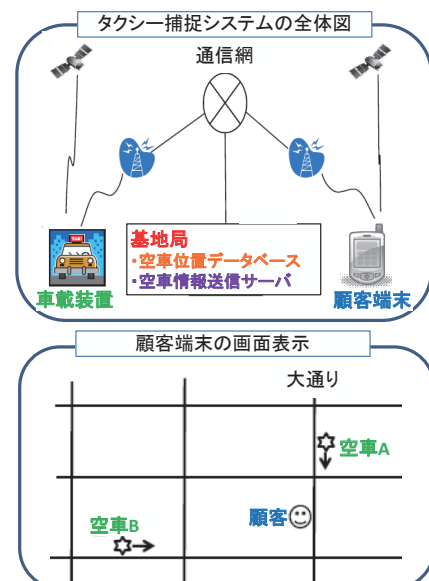
【請求項1】

基地局、**顧客端末**、および**車載装置**からなるタクシー捕捉システムにおいて、

前記**基地局**は、

複数の空車の**車載装置**から受信したそれぞれの空車の位置情報を格納する**空車位置データベース**と、

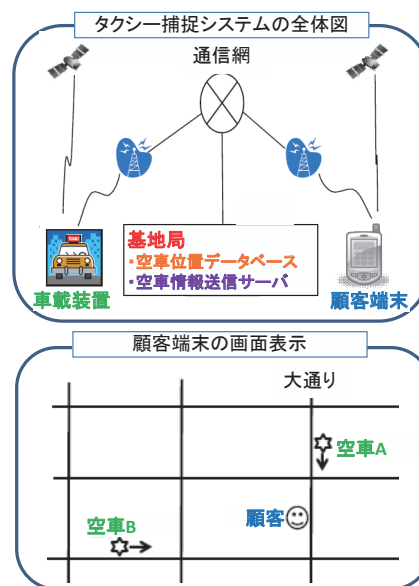
前記**顧客端末**から受信した顧客の位置情報に基づいて、該顧客の周囲に存在する一以上の空車を前記**空車位置データベース**から特定して、地図上に該顧客の位置と該一または複数の空車の位置とが表示される空車状況情報を作成し、該空車状況情報を該顧客端末へ送信する**空車情報送信サーバ**と、を有することを特徴とする、タクシー捕捉システム。



5

特許文書の例 (発明の詳細な説明)

【発明の効果】本発明のタクシー捕捉システムによれば、周囲にいるすべての空車の位置、走行方向を瞬時に把握して、いちばん至近にいる空車を迅速かつ確実に捕まえて乗車することができ、さらに、タクシー運転手とタクシー会社にとってもビジネス機会が増大する。



6

特許文書の例 (発明の詳細な説明)

【発明の効果】本発明のタクシー捕捉システムによれば、周囲にいるすべての空車の位置、走行方向を瞬時に把握して、いちばん至近にいる空車を迅速かつ確実に捕まえて乗車することができ、さらに、タクシー運転手とタクシー会社にとってもビジネス機会が増大する。

and?
or?

主語は？

長文



- ① 並列要素の関係が不明 (and?, or?)
- ② 主語の省略 (誰が乗車する?)
- ③ 一文が長い (115文字)



7

特許文書の例 (特許請求の範囲＝特許権の範囲)

【請求項1】

アントシアニン及びローヤルゼリーを含有することを特徴とする視力回復剤。



8

特許文書の例 (発明の詳細な説明)

【発明の効果】

本発明の実施の形態によれば、目の疲れ・視力低下に対し回復の即効性に優れる視力回復剤を提供することができるし、製造時に調節することで、上記回復効果のみならず、目のかすみなどの目に発症する種々の症状の予防及び治療効果を有する視力回復剤も提供することができる。



9

特許文書の例 (発明の詳細な説明)

【発明の効果】

本発明の実施の形態によれば、目の疲れ・視力低下に対し回復の即効性に優れる視力回復剤を提供することができるし、製造時に調節することで、上記回復効果のみならず、目のかすみなどの目に発症する種々の症状の予防及び治療効果を有する視力回復剤も提供することができる。

長文



- ① 並列要素の関係が不明 (and?、or?)
- ② 目的語の省略 (何を調節する?)
- ③ 一文が長い (127文字)

and?
or?

何を調節?



10

特許ライティングマニュアル



➤ 人による特許明細書作成実務をガイドし、コンピュータによる特許ライティング支援機能を実現する基礎となるもの

➤ 2013年に発行した第1版の内容を見直して、2018年3月に改訂版発行

改訂版発行以来
ダウンロード数は1500以上

➤ 産業日本語研究会のホームページから無料でダウンロード可能
<https://www.tech-jpn.jp/tokkyo-writing-manual/>

11

特許ライティングマニュアル

文 レ ベ ル	1. 短文にする ～短くシンプルな文にする～	・一文を短くして複雑な係り受けをなくすことで、人間の理解が容易になり、誤訳を減少できる。ただし、短文化する際には、意図しない内容とならないように注意する。	
	2. 省略しない ～隠れている要素がないか注意する～	・主語や目的語などの省略された要素を明示して、文意を明瞭にする。	➤ 現状の機械翻訳システム等の状況も踏まえつつ内容を精査
	3. 理解しやすい構成にする ～文の構造に注意する～	・文の構造を工夫することで、理解しやすく翻訳しやすい文にすることができる。	➤ 7つのカテゴリー、27のルールに再構成
	4. 横並びの要素の表現を揃える ～対等に並べ意味に注意する～	・要素を対等に並べることで文章に構造をもたせ、係り受けを明確にし、理解しやすくする。	➤ カテゴリーを、文、節・区、語レベルに分類して体系化
	5. 読点を工夫する ～係り受けや文の構造を明らかにする～	・適切な位置に読点をつけることで、係り受けや文の構造を明らかにする。	➤ 例文の追加・修正を実施
	6. 簡潔にする ～シンプルな表現にする～	・不要、冗長、難解な表現を避け、元の文の意味が変わらない分かりやすい表現を用い、シンプルな表現にする。	
	7. 言い換える ～誰にでも伝わる表現にする～	・多義的な表現やあいまいな表現は、明確・具体的な表現に改める。また、日本語独特の表現は、訳しやすい表現に改める。	

12

特許ライティングマニュアル

ライティングマニュアル 抜粋

1 文レベル

短文にする ～短くシンプルな文にする～

1-1

説明語句が長いときは、短く分ける

説明する語句（修飾語）が長い場合には、以降の文でその説明をするなどして短文に分ける。

例 1

修正前 「本発明は、筐体表面にレジストを塗布する塗布工程を備えることを特徴とする。」

修正後 ➡塗布工程を備えることが特徴の場合：
「本発明は、塗布工程を備えることを特徴とする。この塗布工程では、筐体表面にレジストを塗布する。」
➡具体的な塗布工程が特徴の場合：
「本発明は、後述する塗布工程を備えることを特徴とする。この塗布工程では、筐体表面にレジストを塗布する。」

13

特許ライティングマニュアル

ライティングマニュアル 抜粋

1 文レベル

短文にする ～短くシンプルな文にする～

1-2 複数の主語や述語を含むときは、文を分ける

1文に複数の主語や述語が存在するときは、文を分ける。

例 1

修正前

「仮に、スライドカム取付部材を引っかけてスライドカムを移動させた場合を考えると、この場合には、スライドカム取付部材に応力が局所的に発生し易くなるため、スライドカムの往復移動を繰り返すうちに、スライドカム取付部材にひびが生じることとなり、アダプタの耐久性試験を中断せざるを得なくなるおそれがある。」

修正後

「仮に、スライドカム取付部材を引っかけてスライドカムを移動させた場合を考えると、この場合には、スライドカム取付部材に応力が局所的に発生し易くなる。このため、スライドカムの往復移動を繰り返すうちに、スライドカム取付部材にひびが生じる。その結果、アダプタの耐久性試験を中断せざるを得なくなるおそれがある。」

14

特許ライティングマニュアル

ライティングマニュアル 抜粋

2 節・句レベル

省略しない ～隠れている要素がないか注意する～

2-1 主語を明示する

「何が」「誰が」を明示する。

例 1

修正前

「本実施例によれば、ファスナーの上下部分を保持できる。」

修正後

「本実施例によれば、**作動アーム**はファスナーの上下部分を保持できる。」

例 2

修正前

「この接着液の塗布方法では、被塗布部の下面にのみ塗布される。」

修正後

「この接着液の塗布方法では、被塗布部の下面にのみ**接着液**が塗布される。」

15

特許ライティングマニュアル

ライティングマニュアル
抜粋

2 節・句レベル

省略しない

～隠れている要素がないか注意する～

2-2 目的語を明示する

「何を」「誰を」「何に」「誰に」を明示する。

例 1

修正前 「洗浄ノズルは、高圧の洗浄水を噴射して洗浄する。」

修正後 「洗浄ノズルは、高圧の洗浄水を噴射して**汚染部**を洗浄する。」

例 2

修正前 「自動搬送装置は選択された商品を搬送する。」

修正後 「自動搬送装置は選択された商品を**次のシステムに**搬送する。」

16

特許ライティングマニュアル

ライティングマニュアル
抜粋

3 節・句レベル

理解しやすい構成にする

～文の構造に注意する～

3-1 主語と述部を近づける

「～は、～のとき、～する。」を「～のとき、～は、～する。」のように、
主語を述部に近づけることで、係り受けを明確にする。
ただし、文の構造が短く単純・明瞭であるときは、必ずしも文を修正する必要はない。

例 1

修正前 「**応力の分散による可動部および支持構造部の変形は**、1つの可動部に対し複数のアクチュエータを分散して配置することにより、アクチュエータの発熱による周囲の熱影響が軽減される結果、抑制される。」

修正後 「1つの可動部に対し複数のアクチュエータを分散して配置することにより、アクチュエータの発熱による周囲の熱影響が軽減される結果、**応力の分散による可動部および支持構造部の変形は抑制される。**」

特許ライティングマニュアル

ライティングマニュアル
抜粋

5 節・句レベル

読点を工夫する

～係り受けや文の構造を明らかにする～

5-4

修飾先を明らかにするためにつける

読点の有無や位置によって、文の意味が変化するときには、修飾先を明らかにするように読点をつける。必要に応じて、語順を変えることも検討する。
(参考：ルールカテゴリ3)

例 1

修正前 「安価な装置の部品」

修正後 → 装置が安価：
「安価な装置の、部品」
→ 部品が安価：
「安価な、装置の部品」

例 2

修正前 「高速で走行する車両を検査する」

修正後 → 高速で走行：
「高速で走行する車両を、検査する」
→ 高速で検査：
「高速で、走行する車両を検査する」 /
「走行する車両を、高速で検査する」(ルール3-2(語順変更)も適用)

18

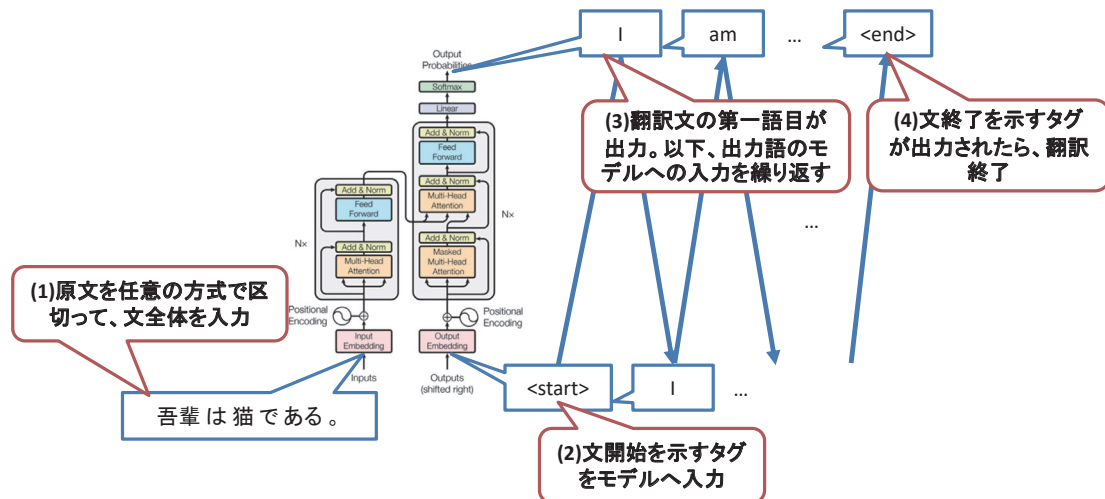
特許ライティングマニュアル

- 特許ライティングマニュアルは、特許文書を人に理解しやすく、かつ、機械にも処理(翻訳)しやすい様に作成するためのもの
- 人にとっての「理解しやすさ」と機械にとっての「翻訳容易」とは同じではない
- AI翻訳(ニューラル機械翻訳)にとっての「翻訳容易」は、これまでの機械翻訳(ルールベース翻訳、統計翻訳)にとっての「翻訳容易」と異なるのだろうか？
- ニューラル機械翻訳が前提となる昨今、特許文書として気をつけるべきポイントは？



19

ニューラル機械翻訳の仕組み (Transformer方式の例)



20

産業日本語から見た ニューラル機械翻訳の特徴

1. コーパスベース機械翻訳

学習データ(原語と訳語の対訳コーパス)から知識を自動獲得し、それを利用して機械翻訳を行う
⇒ルールベース機械翻訳のように(人手で作ったルールに合うように)文法関係などを精緻に整えなくとも、
学習データに一定頻度で対応関係が出現していれば翻訳可能

2. 基本は一文単位での翻訳

基本的に一文ごとに訳文を出力する
⇒他の文との関係や、文書全体の文脈を考慮するのは困難

21

産業日本語から見た ニューラル機械翻訳の特徴

3. ニューラル機械翻訳に関する注意事項

- 長距離の依存関係: 関連のある語同士の距離が離れていくと、
* 統計翻訳も同様 語の関係性を把握する能力が落ちる
- 低頻度語が苦手: コーパスに出現する頻度が低い語
(例: 専門用語)は翻訳不能か、
関連のある別の語に誤訳(関連語に混同)
- 翻訳結果が流暢: 誤訳であっても流暢な日本語となるため、
翻訳結果の確認/ポストエディット時には注意が必要。
そのため、原文の十分な理解がますます重要
* 訳語も統一されていない場合があるためご注意ください。

22

NMTを前提とした特許ライティングマニュアル各項目の整理

1. コーパスベース機械翻訳

たとえ悪文であっても、学習データに一定の頻度で存在していれば、適切に翻訳可能

4 横並びの要素の表現を揃える

～対等に並べ意味に注意する～

要素を対等に並べることによって文章に構造をもたせ、係り受けを明確にし、理解しやすくする。



修正前: 第1領域と拡散領域としての第2領域が...
修正後: 第1領域と拡散領域としての第2領域とが...

⇒文法的に必ずしも正しくない横並び表現

6 簡潔にする

～シンプルな表現にする～

不要、冗長、難解な表現を避け、元の文の意味が変わらない簡潔な表現を用い、シンプルな文にする。



修正前: 約10%程度、低速の速度で動く
修正後: 約10%、低速で動く

⇒不要、冗長、難解な表現

7 言い換える

～指にでも応じる表現にする～

多義的な表現やあいまいな表現は、明確・具体的な表現に改める。日本語特有の表現は、訳しやすい表現に改める。



修正前: <の字型の電極、オーソリ
修正後: Uの字型の電極、オーソリゼーション

⇒日本語特有の表現、省略語

23

NMTを前提とした特許ライティングマニュアル各項目の整理

2. 基本は一文単位での翻訳

基本は一文単位で翻訳語を出力しているため、その文以前の文脈を考慮できない

2 省略しない
～隠れている要素がないか
注意する～

主語や目的語などの省略された要素を明示して、文意を明確にする。



修正前: この作業終了後、乾燥する。
修正後: この作業終了後、**樹脂層**を乾燥する。

⇒ 前の文から省略された要素を補うことができず、適当なものを補ってしまう可能性があるため、省略された要素を明示することが有効

5 読点を工夫する
～係り受けや文の構造を
明らかにする～

適切な位置に読点をつけることで、係り受けや文の構造を明らかにする。



修正前: 安価な装置の部品。
修正後: 安価な**、**装置の部品。安価な装置の**、**部品。

**4 横並びの要素の
表現を揃える**
～対等に並べ意味に注意する～

要素を対等に並べることで文章に構造をもたせ、係り受けを明確にし、理解しやすくする。



修正前: ガラス**間**又はガラスと保護膜の間
修正後: ガラス**と**ガラス**との間**又はガラスと保護膜の間

⇒ 文脈をヒントに係り受け関係を特定することができず、単に尤度の高い翻訳をしてしまうので、係り受け関係の明示や、多義的な表現やあいまいな表現の明確化は有効

7 言い換える
～誰にでも伝わる表現にする～

多義的な表現やあいまいな表現は、明確・具体的な表現に改める。日本語特有の表現は、訳しやすいく表現に改める。



修正前: 制御部は、CPUで構成される。
修正後: 制御部は、CPU**である**。制御部は、CPU**を含む**。

⇒ 文脈をヒントに多義語の意味を特定することができないため、多義的な表現やあいまいな表現の明確化は有効
但し、一文内の語の関係から意味が明らかな場合は、必ずしも必要ではない

24

NMTを前提とした特許ライティングマニュアル各項目の整理

3. ニューラル機械翻訳に関する注意事項(1)

長距離の依存関係: 関連のある語同士の距離が離れていくと、語の関係性を把握する能力が落ちる

**3 理解しやすい
構成にする**
～文の構造に注意する～

文の構造を工夫することで、理解しやすく翻訳しやすい文にすることができる。



修正前: **保持レバーは**、収容ポトルを...阻止するように、折り曲げられる。
修正後: 収容ポトルを...阻止するように、**保持レバーは**、折り曲げられる。

⇒ 関連のある語同士の距離を近づける(主語と述部)ルールなどが有効

1 短文にする
～短くシンプルな文にする～

一文を短くして複雑な係り受けをなくすることで、人間の理解が容易になり、機械翻訳時の誤訳を減少できる。ただし、短文化する際は、意図しない内容とならないように注意する。



修正前: 本発明は、筐体表面にレジストを塗布する塗布工程を備えることを特徴とする。
修正後: 本発明は塗布工程を備えることを特徴とする。**この塗布工程では**、筐体表面にレジストを塗布する。

⇒ 文が短いほど無条件に精度が高くなるわけではないが、短文にすることで、関連のある語同士が近づくことが多いので有効

25

NMTを前提とした特許ライティングマニュアル各項目の整理

3. ニューラル機械翻訳に関する注意事項(2)

低頻度語問題: コーパスに出現する頻度が低い語(例: 専門用語)は、翻訳不能か、関連のある別の語に翻訳

もし、コーパスにおける出現頻度が高くなければ、出現頻度の高い別の表現にすることが有効

6 簡潔にする

～シンプルな表現にする～

不要、冗長、難解な表現を避け、元の文の意味が変わらない簡潔な表現を用い、シンプルな文にする。



修正前: 約10%程度、低速の速度で動く
修正後: 約10%、低速で動く

⇒ 出願頻度の低い不要、冗長、難解な表現

7 言い換える

～誰にでも伝わる表現にする～

多義的な表現やあいまいな表現は、明確・具体的な表現に改める。日本語特有の表現は、訳しやすい表現に改める。



修正前: <の字型の電極、オーソリ、嵌合
修正後: Uの字型の電極、オーソリゼーション

⇒ 出願頻度の低い日本語特有の表現、省略語

正しい翻訳結果が出力されるかは
コーパス内の出現頻度によります

26

産業日本語から見たニューラル機械翻訳に有効と考えられるルール

文 レ ベル	1 短文にする ～短くシンプルな文にする～	一文を短くして複雑な係り受けをなくすることで、人間の理解が容易になり、機械翻訳時の誤訳を減少できる。ただし、短文化する際は、意図しない内容とならないように注意する。	○	文が短ければ短いほど翻訳精度が上がるわけではなく、短文化により関連のある語同士が近づくと有効
	2 省略しない ～隠れている要素がないか 注意する～	主語や目的語などの省略された要素を明示して、文意を明確にする。	○	基本的に、前の文を考慮して翻訳できないので、省略された要素を明示することで、精度の高い翻訳が可能
	3 理解しやすい 構成にする ～文の構造に注意する～	文の構造を工夫することで、理解しやすく翻訳しやすい文にすることができる。	○	関係のある語同士の距離を近づけることで、長距離依存の問題を低減でき、精度の高い翻訳が可能
	4 横並びの要素の 表現を揃える ～対等に並べ意味に注意する～	要素を対等に並べることによって文章に構造をもたせ、係り受けを明確にし、理解しやすくする。	△	文法的に正しくなくても、コーパスに存在する形で、並列要素を列挙すれば正しい翻訳が可能 一方、係り受け関係が明確になる場合は有効
	5 読点を工夫する ～係り受けや文の構造を 明らかにする～	適切な位置に読点をつけることで、係り受けや文の構造を明らかにする。	○	係り受け関係が明確になるので有効
	6 簡潔にする ～シンプルな表現にする～	不要、冗長、難解な表現を避け、元の文の意味が変わらない簡潔な表現を用い、シンプルな文にする。	△	不要、冗長、難解な表現であっても、正しい対訳がコーパスに存在していれば、精度の高い翻訳が可能(低頻度語問題に注意)
	7 言い換える ～誰にでも伝わる表現にする～	多義的な表現やあいまいな表現は、明確・具体的な表現に改める。日本語特有の表現は、訳しやすい表現に改める。	△	多義的な表現やあいまいな表現について明確化することは有効 一方、日本語特有表現も、正しい対訳がコーパスに存在していれば、精度の高い翻訳が可能(低頻度語問題に注意)

27

原文の修正と機械翻訳結果の例

日本語(原文/リエディット後)	ニューラル機械翻訳結果(*下線、マーカ、太字等を追記)	コメント
蒸着工程が繰り返される過程でチャンバ20と噴射ノズル10は、繰り返し熱によって膨張と収縮が繰り返される繰り返し荷重に露出し、このような熱による繰り返し荷重により、噴射ノズル10をチャンバの外壁21に結合させるねじ結合部12の部位に緩みやクラックが発生する恐れがあり、噴射ノズル10とチャンバの外壁21の間には微細な隙間が形成される可能性がある。	In the process of repeating the vapor deposition process, the chamber 20 and the injection nozzle 10 are exposed to a repeated load that repeatedly expands and contracts due to repeated heat, and the injection nozzle 10 is coupled to the outer wall 21 of the chamber by the repeated load caused by such heat. There is a possibility that a looseness or a crack may occur at the site of the screw coupling portion 12, and a fine gap may be formed between the injection nozzle 10 and the outer wall 21 of the chamber.	★後半の文章に出てくる「ねじ結合部12」の説明に関し、本来一文で記載されるべき「噴射ノズル10」、「チャンバの外壁21」との関係が、2箇所に分断して記載されている。
蒸着工程が繰り返される過程でチャンバ20と噴射ノズル10は、繰り返し熱によって膨張と収縮が繰り返される繰り返し荷重に露出する。このような熱による繰り返し荷重により、噴射ノズル10をチャンバの外壁21に結合させるねじ結合部12の部位に緩みやクラックが発生する恐れがあり、噴射ノズル10とチャンバの外壁21の間には微細な隙間が形成される可能性がある。	In the process of repeating the vapor deposition process, the chamber 20 and the injection nozzle 10 are exposed to repeated loads that are repeatedly expanded and contracted by repeated heat. Due to the repeated load caused by such heat, there is a <u>risk of loosening or cracking</u> at the site of the screw coupling portion 12 that couples the injection nozzle 10 to the outer wall 21 of the chamber. <u>Fine gaps</u> may be formed. ()	リエディット例1 「1-2 複数の主語や述語を含むときは、文を分ける」を実施 ★誤訳部分は解消 ※意図せぬ変更が発生(緑色部分):さらなる原文修正には、作成者の意図や明細書の詳細確認が必要 ※最後に訳抜けが発生

28

原文の修正と機械翻訳結果の例

日本語(リエディット後)	ニューラル機械翻訳結果(*下線、マーカ、太字等を追記)	コメント
蒸着工程が繰り返される過程でチャンバ20と噴射ノズル10は、繰り返し熱によって膨張と収縮が繰り返される繰り返し荷重に露出する。このような熱による繰り返し荷重により、噴射ノズル10をチャンバの外壁21に結合させるねじ結合部12の部位に緩みやクラックが発生する恐れがあり、噴射ノズル10とチャンバの外壁21の間には微細な隙間が形成される可能性がある。	In the process of repeating the vapor deposition process, the chamber 20 and the injection nozzle 10 are exposed to repeated loads that are repeatedly expanded and contracted by repeated heat. Due to the repeated load caused by such heat, there is a <u>risk of loosening or cracking</u> at the site of the screw coupling portion 12 that couples the injection nozzle 10 to the outer wall 21 of the chamber. <u>Fine gaps</u> may be formed. ()	リエディット例1 「1-2 複数の主語や述語を含むときは、文を分ける」を実施 ★誤訳部分は解消 ※意図せぬ変更が発生(緑色部分):さらなる原文修正には、作成者の意図や明細書の詳細確認が必要 ※最後に訳抜けが発生
蒸着工程が繰り返される過程でチャンバ20と噴射ノズル10は、繰り返し熱によって膨張と収縮が繰り返される繰り返し荷重に露出する。このような熱による繰り返し荷重により、噴射ノズル10をチャンバの外壁21に結合させるねじ結合部12の部位に緩みやクラックが発生する恐れがある。その結果、噴射ノズル10とチャンバの外壁21の間には微細な隙間が形成される可能性がある。	In the process of repeating the vapor deposition process, the chamber 20 and the injection nozzle 10 are exposed to repeated loads that are repeatedly expanded and contracted by repeated heat. Due to the repeated load caused by such heat, there is a risk of loosening or cracking occurring at the site of the screw coupling portion 12 that couples the injection nozzle 10 to the outer wall 21 of the chamber. As a result, a fine gap may be formed between the injection nozzle 10 and the outer wall 21 of the chamber.	リエディット例2 「1-2 複数の主語や述語を含むときは、文を分ける」を2か所を実施。 ★最後の訳抜けが解消

29

産業日本語の今後の課題

- 産業日本語研究会、各分科会等の場におけるさらなる議論と研究
- 研究成果により、人が理解しやすい特許文書の実現に貢献
- 翻訳精度の向上により、海外への発信力を強化
- 文章の意味の理解が困難と指摘される人工知能などの機械による処理結果の精度を、言語資源側から向上
- 特許ライティングマニュアルを始めとした活動の成果の現場への普及

30

ご清聴ありがとうございました。

＜産業日本語に関するご意見、ご質問はこちらへ＞
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／知財AI研究センター主幹 清藤 弘晃
TEL: 03-3615-5513
E-mail: hiroaki_kiyoto@japio.or.jp

31

— 禁無断転載 —

令和元年度
産業日本語研究会 報告書
「産業日本語」
Technical Japanese

令和2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