

第6回産業日本語研究会・シンポジウム 予稿集

平成27年2月24日

於 東京大学 情報学環・福武ホール ラーニングシアター

高度言語情報融合フォーラム

言語処理学会

一般財団法人
日本特許情報機構



第6回産業日本語研究会・シンポジウム開催のご案内

平成27年2月

日本語を見つめ直し、今後の産業日本語を考える

市場のグローバル化を背景に、日本企業の海外展開が進み、それに伴い、グローバルな活用場面を想定した産業技術文書（特許明細書、論文、説明書、報告書など）は、正確かつ円滑な情報発信力と知的生産性の向上による国際競争力強化のために、これまでも増して強く求められているところです。また昨今、新興国ビジネスの拡大や訪日外国人観光客の急増、「東京2020オリンピック・パラリンピック」の開催準備など、我が国産業界において、言語処理技術が重要な役割を担うものと期待されています。

「産業日本語研究会」では、以上のような背景のもと、「産業日本語」に関する各種の研究や提言などに取り組み、我が国産業界全体の国際競争力の強化に資する日本語の枠組みのあり方について継続的に議論しています。ここで「産業日本語」とは、情報を正確かつ円滑に伝達できるような情報発信力の強化、そして、コンピュータ処理されやすいような知的生産性の向上に資する産業や科学技術の記述に用いられるべき日本語の枠組みと定義しています。

また、平成22年より毎年開催している「産業日本語研究会・シンポジウム」では、特許、法令工学、翻訳、テクニカルコミュニケーション、システム開発など多分野にわたって、日本語を分かりやすく用いるための取組などが紹介され、その取組を発展させることの重要性も議論されてきました。

本年の「第6回産業日本語研究会・シンポジウム」では、日本語に関する取組や研究について一層の発展を目指すべく、様々な角度・視点から日本語を見つめ直しつつ、今後の産業日本語のあり方について議論する場をご提供したいと考えております。今回のプログラムでは、言語処理分野の若手研究者から産業への応用研究をご紹介いただくほか、有識者によるパネル討論を加えました。

産業日本語に関わる研究の新たなステップにつなげ、我が国産業に大いに貢献できる機会になると期待しております。多くの皆さまのご参加を、よろしくお願いいたします。

産業日本語研究会世話人会

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）

辻井 潤一（マイクロソフトリサーチアジア研究所）

橋田 浩一（東京大学）

隅田英一郎（情報通信研究機構）

横井 俊夫（日本特許情報機構特許情報研究所）

潮田 明（元奈良先端科学技術大学院大学）

河合 弘明（日本特許情報機構）

- 主催：高度言語情報融合フォーラム（ALAGIN）、言語処理学会、日本特許情報機構（Japio）
■後援：総務省、経済産業省、特許庁、国立国語研究所、情報通信研究機構、
工業所有権情報・研修館、情報処理学会、人工知能学会、
アジア太平洋機械翻訳協会
■日時：2015年2月24日（火） 13:00-17:55
■場所：東京大学 情報学環・福武ホール ラーニングシアター
（東京大学 本郷キャンパス）
<http://fukutake.iii.u-tokyo.ac.jp/access/index.html>
■テーマ：日本語を見つめ直し、今後の産業日本語を考える
■参加費：無料（事前登録制）

■プログラム：

- (1) 開会挨拶 (13:00-13:05)

長尾 眞
京都大学 名誉教授

【第一部】招待講演 (13:05-14:05)

- (2) 暗号研究者から見た言語・論理とデジタルフォレンジック

辻井 重男
中央大学研究開発機構 機構教授
／一般財団法人マルチメディア振興センター 理事長
／一般財団法人放送セキュリティセンター 理事長

- (3) 特許審査情報の世界への発信と産業日本語

ーグローバル知財戦略支援の取組ー

大町 真義
特許庁審査第二部福祉機器 審査長

（休憩 10分）

【第二部】活動報告 (14:15-14:55)

- (4) 特許版・産業日本語の活動報告

河合 弘明
一般財団法人日本特許情報機構 調査研究部長
／特許版・産業日本語委員会 委員

- (5) 日本人のための日本語マニュアル

日本語マニュアルの会（順不同）
横井 俊夫
東京工科大学 名誉教授
／一般財団法人日本特許情報機構 特許情報研究所 顧問
石崎 俊
慶応大学 名誉教授
佐野 洋
東京外国語大学 教授

石黒 圭
一橋大学 教授
猪野 真理枝
翻訳家／語学教材作家
烏 日哲
一橋大学国際教育センター 非常勤講師

(休憩 15 分)

【第三部】言語処理分野の最前線 (15:10－16:10)

(6) 産業日本語としての医療カルテ文章

荒牧 英治
京都大学デザイン学ユニット 特定准教授
／独立行政法人科学技術振興機構 さきがけ研究員

(7) 日本語を対象とした統計的機械翻訳の進展

Graham Neubig
奈良先端科学技術大学院大学 助教

(8) 特許オントロジーの構築とその特許検索への応用

難波 英嗣 広島市立大学大学院情報科学研究科 准教授

(休憩 15 分)

【第四部】パネル討論 (16:25－17:55)

座長	井佐原 均	豊橋技術科学大学 教授
パネリスト	乾 健太郎	東北大学 教授
	影浦 峡	東京大学 教授
	外山 勝彦	名古屋大学 教授
	勝田 豊彦	一般財団法人テクニカルコミュニケーター協会 ／ハイテクノロジー・コミュニケーションズ株式会社

【第五部】懇親会・意見交換会 (18:00－19:00)

※第五部 懇親会・意見交換会は、福武ホール ラーニングスタジオにて行ないます。ささやかですが、飲み物など準備しておりますので、こちらも是非ご参加ください。

プログラム詳細

(1) 開会挨拶

(13:00－13:05)

長尾 眞
京都大学 名誉教授

【第一部】招待講演

(13:05－14:05)

(2) 暗号研究者から見た言語・論理とデジタルフォレンジック

辻井 重男

中央大学研究開発機構 機構教授

／一般財団法人マルチメディア振興センター 理事長

／一般財団法人放送セキュリティセンター 理事長

古典暗号時代に、暗号解読とその誤訳などが、第1次・第2次大戦の勃発に及ぼした深刻な影響について触れた後、現代暗号研究者から見た組織通信における言語と論理の役割を多角的に考察する。特に、現代暗号の安全性証明における論理的証明、クラウド環境下における秘匿検索・暗号化状態処理のための自然言語処理と論理学の活用、及び、デジタルフォレンジックや公共情報コモンズ（災害情報伝達システム、Lアラート）における言語翻訳課題の緊急性について考察する。

(3) 特許審査情報の世界への発信と産業日本語

－グローバル知財戦略支援の取組－

大町 真義

特許庁審査第二部福祉機器 審査長

企業等のグローバル知財戦略を支援するうえで望ましい状況は、特許出願人が日本で特許を取得すれば、その審査結果が海外の審査でも参考にされ、海外でも権利を速やかに取得できる状況である。このことを考慮し、特許庁での審査関連書類を海外の庁でも参照可能とするよう、日英機械翻訳を活用した情報発信システムが構築されてきた。本講演では、審査において作成する書類を、日本語としてわかりやすいとともに、同システムにより機械翻訳されても理解されやすい文章により記載することを目指した、特許庁における取組を紹介する。

(休憩 10分)

【第二部】活動報告

(14:15—14:55)

(4) 特許版・産業日本語の活動報告

河合 弘明

一般財団法人日本特許情報機構 調査研究部長

／特許版・産業日本語委員会委員

人に理解しやすく、機械翻訳など言語処理技術を活用するコンピュータにも処理しやすい日本語研究のうち、特許情報への応用を対象としたものを「特許版・産業日本語委員会」で研究・分析してきた。以下のテーマを中心に報告する。

- ・明晰な特許文章を書き、論理的に明確な特許明細書を作成するための「特許ライティングマニュアル（初版）」 改版に向けた作業進捗
- ・法的観点での適正な特許書類を目指す「36 条ルール化検討」 特許法第 36 条（記載要件）の違反類型等に基づく拒絶理由通知書の分析等

(5) 日本人のための日本語マニュアル

日本語マニュアルの会（順不同）

横井 俊夫

東京工科大学 名誉教授

／一般財団法人日本特許情報機構特許情報研究所 顧問

石崎 俊

慶応大学 名誉教授

佐野 洋

東京外国語大学 教授

石黒 圭

一橋大学 教授

猪野 真理枝

翻訳家／語学教材作家

烏 日哲

一橋大学国際教育センター 非常勤講師

日本人は、日本語を使えるが日本語の仕組を十分に知っているわけではない。日本語文章の分析能力を養成するための「日本人のための日本語マニュアル」が必要である。本年 9 月上梓を目標に「日本語マニュアルの会」がマニュアル制作を進めている。マニュアルの書名と章立ては、以下のような予定である。

『日本人のための日本語マニュアル

- 言葉の仕組を学び、外国語との対照を通じて日本語スキルを磨く - 』

はじめに マニュアルの目的、マニュアルの構成と利用の仕方

1 章 文書・文章ライティングのモデルプロセスを学ぶ

2 章 情報を表し情報を伝える言葉の仕組みを学ぶ・日本語と外国語とを照らし合わせる

3 章 「表す日本語」へと言い換え、「伝える日本語」へと言い換える

4 章 「訳せる日本語」へと言い換える

5 章 コンピュータの支援機構を活用する - 文章校正ソフトと機械翻訳ソフト -

おわりに さらなるスキルアップのために - 参考資料 -

産業日本語を定着させるためには、それぞれの機関・組織・業界・分野がそれぞれの文書目的に沿ったライティングマニュアルを整備しなければならない。本マニュアルが各所での整備作業の一助となることを願うと共に、そうなるための効果的な仕組を用意する予定である。

(休憩 15 分)

【第三部】言語処理分野の最前線

(15:10—16:10)

(6) 産業日本語としての医療カルテ文章

荒牧 英治

京都大学デザイン学ユニット 特定准教授

／独立行政法人科学技術振興機構 さきがけ研究員

医療カルテは、施設毎、患者毎に作成されるため大量に生産される。また、臨床的な意義だけでなく、研究に二次利用されるという学問的な価値や訴訟の際には法的根拠にもなりうる法的重要性の高い文章である。しかし、多くの個人情報が含まれるため、一般の目に触れることは滅多にない。本発表では、医療カルテの可能性及び、これを扱う研究の普及活動について議論する。

(7) 日本語を対象とした統計的機械翻訳の進展

Graham Neubig

奈良先端科学技術大学院大学 助教

人手を介さずに翻訳を行う機械翻訳は、言葉の壁を破り情報の流通を促進する可能性があるとして、長年注目されている。しかし、比較的類似しているヨーロッパ言語間の機械翻訳が既に高い精度を実現している一方、他の言語と特徴が大きく異なる日本語では翻訳精度が比較的低い。このため、日本の翻訳産業における機械翻訳の利用実績はヨーロッパに比較して遥かに少ない。しかし、近年の研究では、かつて遠い言語同士の翻訳を苦手としていた統計的機械翻訳の進展により、異なる言語構造を持つ言語でも翻訳性能が着実に向上している。本発表では、日本語を対象とした翻訳の性能向上につながっている「正確な並べ替えモデル」「原文に存在しない単語の補完」「ニューラルネットによる翻訳モデル」など、最新の研究トピックについて説明する。

(8) 特許オントロジーの構築とその特許検索への応用

難波 英嗣

広島市立大学大学院情報科学研究科 准教授

特許データベースからオントロジーを構築する手法について述べる。オントロジーは、文献を検索したり、高度な言語処理を行ったりするための有用な情報源として活用されている。しかし、オントロジーを人手で構築し、更新することは非常にコストがかかる。一方で、テキストデータベースからシソーラスやオントロジーを自動構築する様々な手法が提案されているものの、人手による構築作業に取って代わるレベルまでには至っていない。そこで、本講演では最小減の労力で効率的にオントロジーを構築する枠組みについて述べる。また、オントロジーを用いた特許検索システムを紹介する。

(休憩 15 分)

【第四部】パネル討論

(16:25—17:55)

座長 井佐原 均
パネリスト 乾 健太郎
 影浦 峡
 外山 勝彦
 勝田 豊彦

豊橋技術科学大学 教授
東北大学 教授
東京大学 教授
名古屋大学 教授
一般財団法人テクニカルコミュニケーター協会
／ハイテクノロジー・コミュニケーションズ株式会社

【第五部】懇親会・意見交換会

(18:00—19:00)

目次

第一部 招待講演

1. 暗号研究者から見た言語・論理とデジタルフォレンジック ----- 1
辻井 重男（中央大学研究開発機構 機構教授
／一般財団法人マルチメディア振興センター 理事長
／一般財団法人放送セキュリティセンター 理事長）
2. 特許審査情報の世界への発信と産業日本語 ----- 37
大町 真義（特許庁審査第二部福祉機器 審査長）

第二部 活動報告

3. 特許版・産業日本語の活動報告 ----- 53
河合 弘明（一般財団法人日本特許情報機構 調査研究部長
／特許版・産業日本語委員会 委員）
4. 日本人のための日本語マニュアル ----- 61
横井 俊夫（東京工科大学 名誉教授
／一般財団法人日本特許情報機構特許情報研究所 顧問）

第三部 言語処理分野の最前線

5. 産業日本語としての医療カルテ文章 ----- 75
荒牧 英治（京都大学デザイン学ユニット 特定准教授
／独立行政法人科学技術振興機構 さきがけ研究員）
6. 日本語を対象とした統計的機械翻訳の進展 ----- 97
Graham Neubig（奈良先端科学技術大学院大学 助教）
7. 特許オントロジーの構築とその特許検索への応用 ----- 113
難波 英嗣（広島市立大学大学院情報科学研究科 准教授）

第四部 パネル討論

8. 計算機によるデータに基づく言い換え ----- 127
乾 健太郎（東北大学 教授）
9. 制限言語と言い換え：多言語展開の改善へ向けて ----- 135
影浦 峯 （東京大学 教授）
10. 日本法令の英訳支援 ----- 141
外山 勝彦 （名古屋大学 教授）
11. 「伝える側の視点」から「伝えられる側の視点」へ ----- 155
勝田 豊彦（一般財団法人テクニカルコミュニケーター協会
／ハイテクノロジー・コミュニケーションズ株式会社）

第一部

招待講演 1

暗号研究者から見た言語・論理と デジタルフォレンジック

辻井 重男

中央大学研究開発機構 機構教授

／一般財団法人マルチメディア振興センター 理事長

／一般財団法人放送セキュリティセンター 理事長

第6回産業日本語研究会・シンポジウム

辻井重男

暗号研究者から見た言語・論理とデジタルフォレンジック 概要

古典暗号時代に、暗号解読とその誤訳などが第1次・第2次大戦の勃発に及ぼした深刻な影響について触れた後、現代暗号研究者から見た組織通信における言語と論理の役割を多角的に考察する。

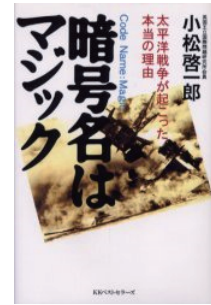
特に、現代暗号の安全性証明における論理的証明、クラウド環境下における秘匿検索・暗号化状態処理のための自然言語処理と論理学の活用、デジタルフォレンジックや公共情報コモンズ(災害情報伝達システム、Lアラート)における言語翻訳課題の緊急性について考察する。

2014年7月辻井企画講演会の動機

法令工学 北陸先端科学技術大学院大学(JAIST) 21世紀COE

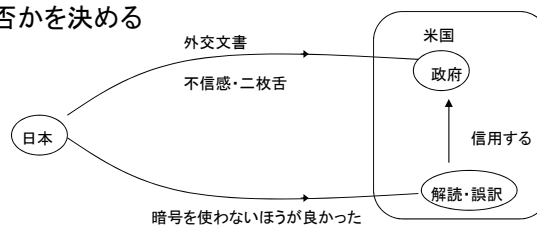
- 論理学と日本語の両面から
- 法律と論理学・言語処理 学際的総合研究
- デジタルフォレンジック 武田薬品 80億円 殆ど翻訳料
- 公共情報コモンズ(マルチメディア振興センター)
- 避難情報のリアルタイム多言語化
- イギリス 17世紀の英語改革 外山滋比古
- 個人的研究動機
- 論理学暗号方式の創始
- 放送通信の4類型と組織通信—情報セキュリティ概念の高度化

誤訳
御前会議 Meeting in the Morning
天聴ニ達セラレアリ(難解な表現、文語調)
甚々恐懼ニ堪エサルモ(天皇に対し出すぎたようで恐縮である)
誤訳
↓
重大な懸念を抱いている(政治家自身が)



最新＝最後 → 最終通告
(電文簡単化)
日米戦争回避条件の一つ
米独開戦時 日独伊三国同盟

↓
日本 自主的に参戦か否かを決める
↓
Automatically と誤訳



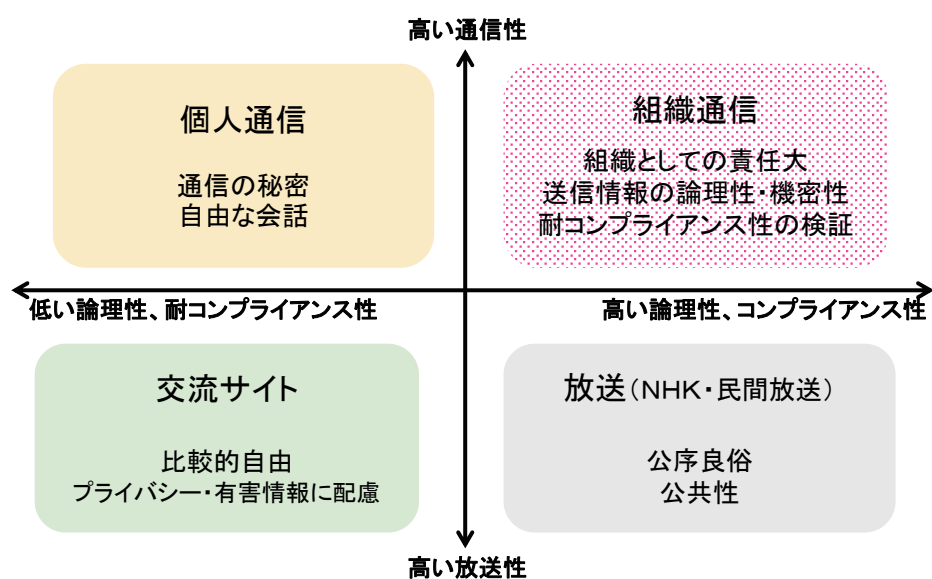
議論の流れ

- 1. 情報伝送・暗号理論研究の立場から
- 2. デジタルフォレンジックと日本語の課題
- 3. 日本語の論理性に対する私見
- 4. 言語と論理に関する課題
- 付録 1. 放送・通信の4類型と情報セキュリティ概念の高度化—第2報
—組織通信と公共情報コモンズ(Lアラート)
- 2. 尸位素餐 (しいそさん) とは？

1. 情報伝送・暗号理論研究の立場から

1. 研究の背景と目的

放送・通信の4類型と組織通信



組織間通信における機密通信の必要性

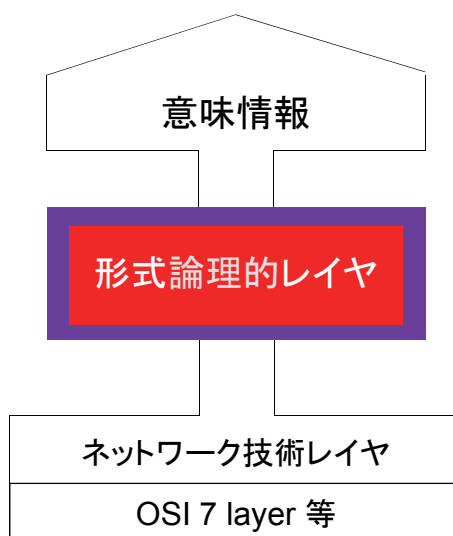
- 情報量の爆発的増大、組織の複雑化、
- 関与者 (Multi-Stake Holder) の拡大を背景に
- 企業、自治体、医療・介護ネットワークなどの
- 組織間における高度な機密通信の必要性増大

例: My Number、税と社会保障、収入・資産

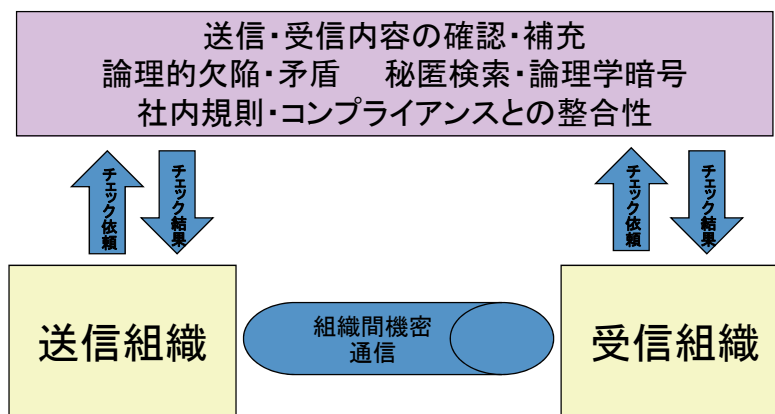
全自治体(居住区域外)に跨る個人情報把握の必要性。
プライバシー保護のため、必要な個人情報のみ把握。

7

情報通信の新階層



Open data, Cloud, Big data, My number環境下での
情報通信・セキュリティ概念の高度化への対応



9

組織間機密通信

- 従来は自治体や病院などの単独組織内で、情報を共有・処理することが多かった。



- ネットワーク化、OCBM環境化、関与者の増大

例えば、医療・介護組織間の連携・業務指示

- 事業所単位での活動から連携へ
- 作業指示書などが書類でなく情報通信で

- ① 機密・プライバシーを保護し、
- ② 情報の論理性・耐コンプライアンス性を高めつつ、複数組織間や民間・行政間での情報交換の増大

10

組織通信に要求される機能

正確性、信頼性、緊急性、迅速性、証拠性(デジタルフォレンジック)

機密性(組織暗号、構造化言語による秘匿検索)

情報限定性、論理的無矛盾性、

法的整合性(耐コンプライアンス性)

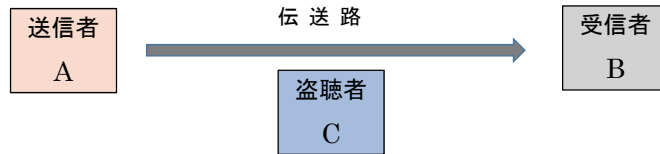
多言語性 (赤字は、本委託研究で実施中)

法令工学

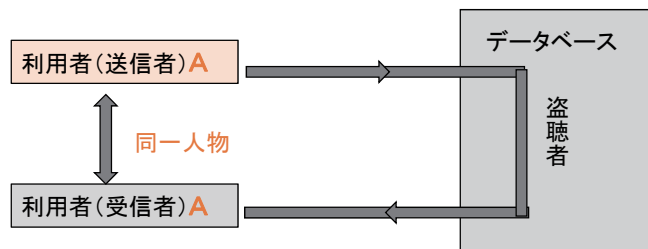
JAIST 21世紀COE 法令工学の提案 片山 卓也

- 「法令工学とは、法令が
 - 1) 制定目的にそって適切に作られ、
 - 2) 論理的矛盾や文書的問題がなく、
 - 3) 関連法令との整合性がとれていることを検査・検証し、
 - 4) 法令の改訂に対しては、矛盾なく変更や追加削除が行われること
- を情報科学的手法によって支援することを目的とする学問分野である」
- 法令工学という法令とは、法律のみでなく、都道府県の条令や、企業の社内規則なども対象にしている。
- 論理学と自然言語処理の相性を両側から高める研究が重要

秘密通信と秘匿検索



(a) 秘密通信の場合



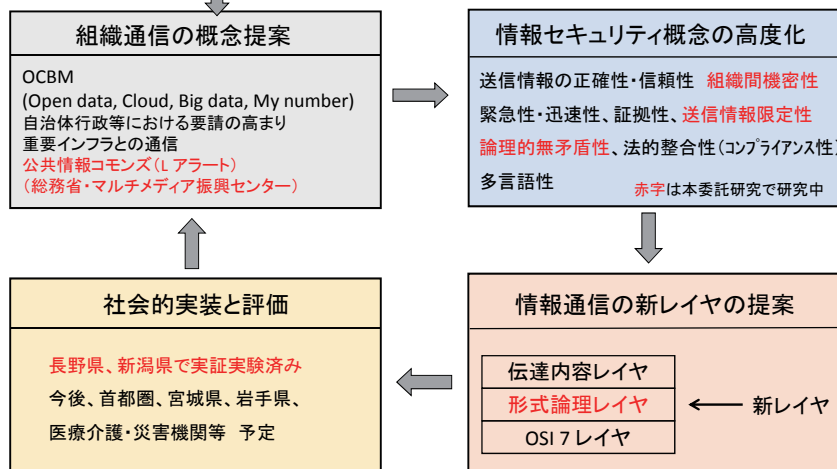
(b) 秘匿検索の場合

自然言語から論理式への変換(暗号化)については伝達には必要はない

13

組織通信概念の提案とその実現に向けて

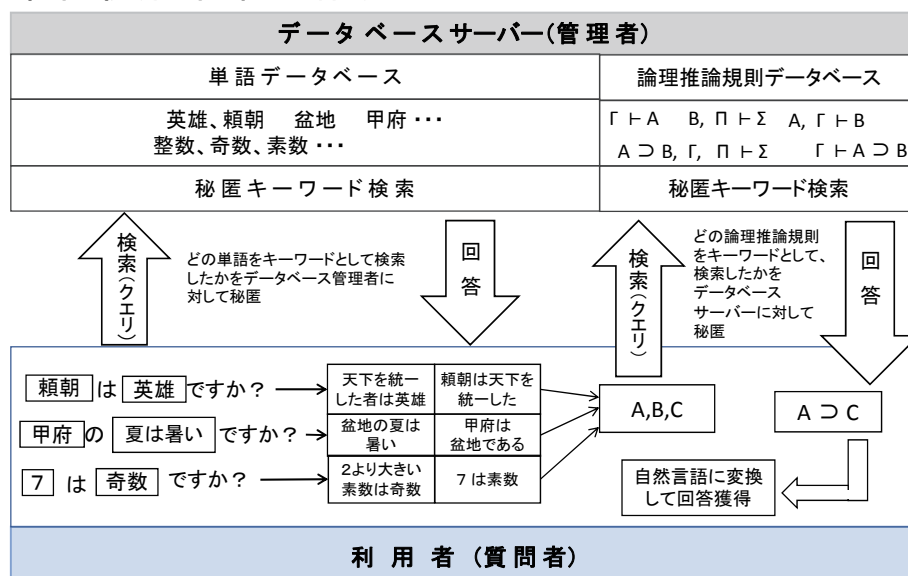
IEEE等国際会議・総務省シンポで賛同的評価



通信の意義 高速大容量→信を通わす

- OCBM時代 Open,Cloud,Big data,My number
- ビジネス、法令、科学技術文書などの論理性
- 送信内容の論理的欠陥・矛盾の検出 CIA概念の拡大
- 質問・回答 証拠の論理的保全・開示
- 暗号化 平文を論理式に直接暗号化
- 現代論理学を利用 古典論理を超えて
- 6月2日 日経朝刊 米国での訴訟和解 深刻な課題
- デジタル・フォレンジック 証拠の保全・開示
- 武田薬品工業 80億 殆どが 翻訳料

3. 研究開発内容 3.2 情報通信・セキュリティ概念の高度化への対応 現代論理学暗号に基づく 秘匿検索・回答文作成システム

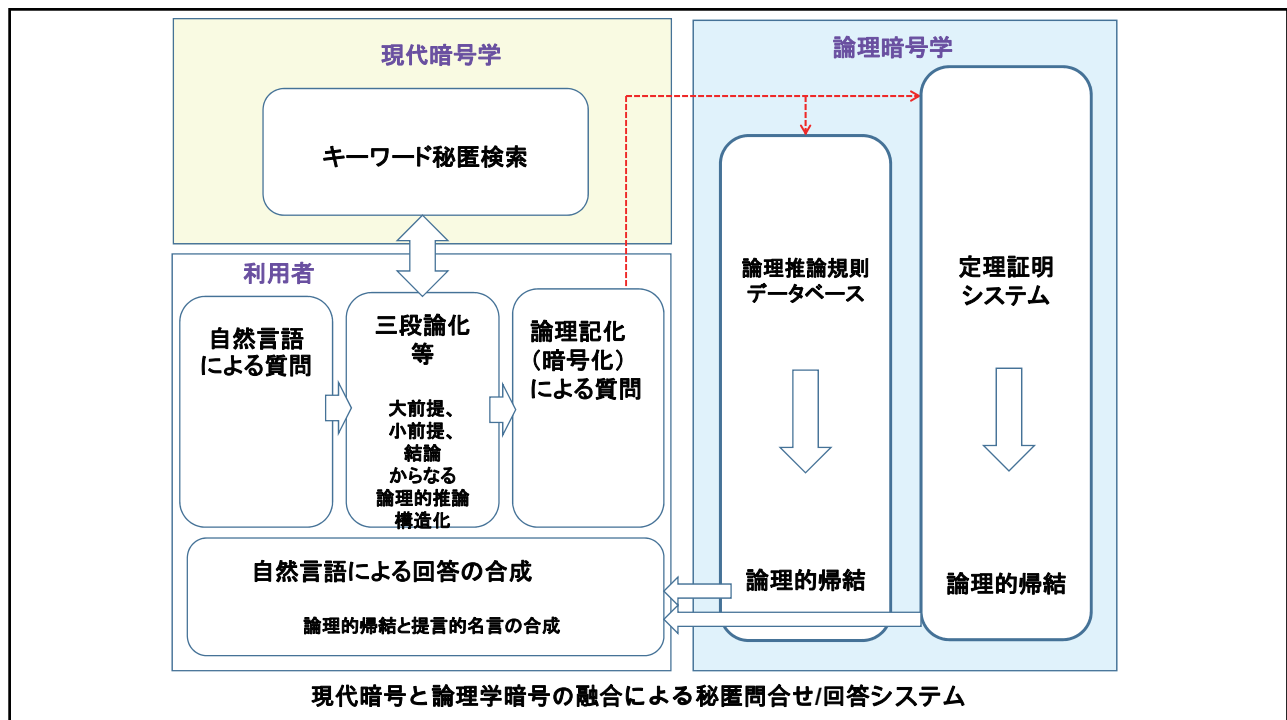


アリストテレスの三段論法のモデルの例

- 言い換えれば、モデルが開示されなければ、何を推論しているのかわからない。
日常言語と導出規則 $2(A \supset B, B \supset C \vdash A \supset C)$ は、多から1への写像である。
- 多から1への写像を説明する「導出規則 $2(A \supset B, B \supset C \vdash A \supset C)$ に対応する日常言語を使った別のモデルによる三段論法」には次のようなものが考えられる。
 甲府は(A), 盆地である(B)
 盆地は(B), 夏暑い(C)
 ゆえに、甲府は(A), 夏暑い(C)

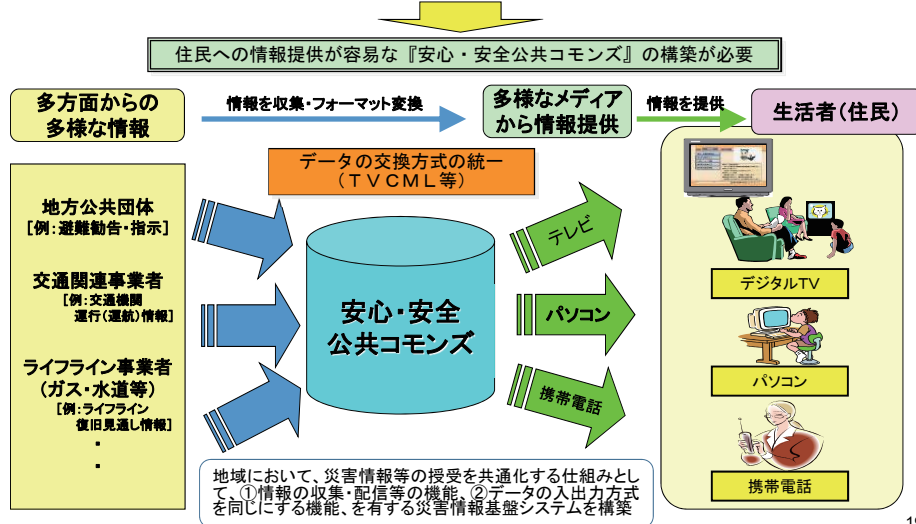
 2より大きい素数は(A),
 2より大きい奇数である(B)
 2より大きい奇数は(B), 7である(C)
 ゆえに、2より大きい素数は(A), 7である(C)
- このように、導出規則 $2(A \supset B, B \supset C \vdash A \supset C)$ に意味を与えるモデルと導出規則自体を分離して考えることが、情報セキュリティにおける「他者からの隠ぺい概念」に結びつく。

17



総務省「地域の安心・安全情報基盤に関する研究会」報告書（平成20年7月2日）より抜粋
『安心・安全公共コモンズ』のコンセプト（FMMC吉田部長）

多様なメディアを活用して、高齢者をはじめ誰もが、いつでもどこでも、分かりやすい形式で地域の安心・安全に係るきめの細かい情報を迅速に入手できるようにするための具体的な仕組みを、早急に構築することが喫緊の課題



19

2. デジタルフォレンジック —国際訴訟と翻訳の課題—

参考

- 1) 守本正宏著「カルテル・PL訴訟・特許訴訟」ディスカバリ
• (企業家大学出版2012年3月)
- 2) 改訂版 デジタルフォレンジック事典
• デジタルフォレンジック研究会編(日科技連)
• 辻井; 日本語の論理性向上のための論理学・自然言語処理
• ・機械翻訳(第10章 10. 1)

米国における 民事訴訟—Pretrial

- 6月2日 日経朝刊 米国での訴訟和解 深刻な課題
- デジタル・フォレンジック 証拠の保全・開示
- 武田薬品工業 80億 殆どが 翻訳料
- -----
- 日本企業 不必要な資料まで翻訳 情報漏洩

e ディスカバリの流れ

- 1. 特定 (identification) データなどの具体的確認
- 2. データの保全 (Preservation)
- 3. データの収集 (Collection)
- 4. データプロセス (Processing)
- 5. 分析 (Analysis) 翻訳の絞込み
- 6. 証拠閲覧 (Review)
- 7. 提出ドキュメントの作成 (Production)

「ディスカバリ」 守本 正宏著

10周年記念 表彰普及・啓発賞

- 日米民事訴訟比較 米国文化 当事者主義；
- プリトライアルの段階で、被告と原告双方が、徹底的に「証拠開示(ディスカバリ)」する。
- デジタル署名・公開鍵暗号の役割
- 余談；池上 彬番組 エニグマ機 出演 しかし・・・
- 相手側が提供した情報の中から証拠を見つけ出す。

- 日本本社の在米国子会社が訴訟にあった場合の
- 翻訳の費用は莫大。
- 機械翻訳・自然言語処理、現代論理学の導入

3. 日本語の論理性に対する私見

- 尸位素餐（しいそさん）とは？
- —ビッグデータ時代の日本語の論理性と国際性
- 付録参照

日本語と日本人のDNA

- 長いトンネルを抜けると雪国であった 川端康成
- The train came out of the long tunnel into the snow country
サイデンスティック
その列車が、その長いトンネルを抜けると、その雪国に入った
再日本語訳

主体—客体？ 純粹経験 主語・定冠詞 不要？
自他—体の境地
フランスの2歳の子供でも（写真）

西田幾多郎著 「日本文化の問題」(1940年、昭和15年)

- 「私は日本文化の特色と云うのは・・・どこまでも自己自身を否定して
- 物となる、物となって見、物となつて行くと云うにあるのではないかと思う。己を空うして物を見る、自己が物の中に没する、・・・
- 日本精神の真髓は、物に於て、事に於て一になると云うことでなければならない。」
- 自己を否定し、私を排し、無心になって、対象と自己を一体化する。
- (佐伯啓思著 西田幾多郎 無私の思想と日本人)

日本語、途方もなく自由だった？

日本語が持つ「途方もない融通無碍な自由さ」だ。「非論理的なものも『てにをは』がつなげてしまうなど意味を超えて感情を喚起する、ある種の分泌性がある」。そして日本語を操る我々にも「つじつまが合わないものを受け入れ、そこに美や叙情を感じる性質がある」というのだ。

(赤田 泰和)

朝日新聞 2013年4月30日 夕刊掲載記事より
結構ですね。

日本語の課題

- 揺れ動く感情を連綿と綴る
湿度100%の演歌？（広田）
- 外山 滋比古
- 150年振りに改革必要
- イギリス 17世紀 大改革
- 情緒・感覚→明晰な論理表現へ

イギリス・ロイヤルアカデミーの英語改革

『今の英語』

- ◎詩、文学など
情緒・感情を中心とした
曖昧模糊たる表現が多い
- ◎これではイギリス人の新しい
考え方を十分に表現できない

英語改革

『改革した英語』

- ◎科学文章など、科学的、論理的、
思考的な表現に重点を置く
- ◎これなら、イギリス人の納得いく
内容を明晰に表現できる



イギリス・ロイヤルアカデミー

トーマス・スプラット
(1635-1713)

The History of the Royal-Society of London
for the Improvement of Natural Knowledge



4. 言語と論理に関する課題

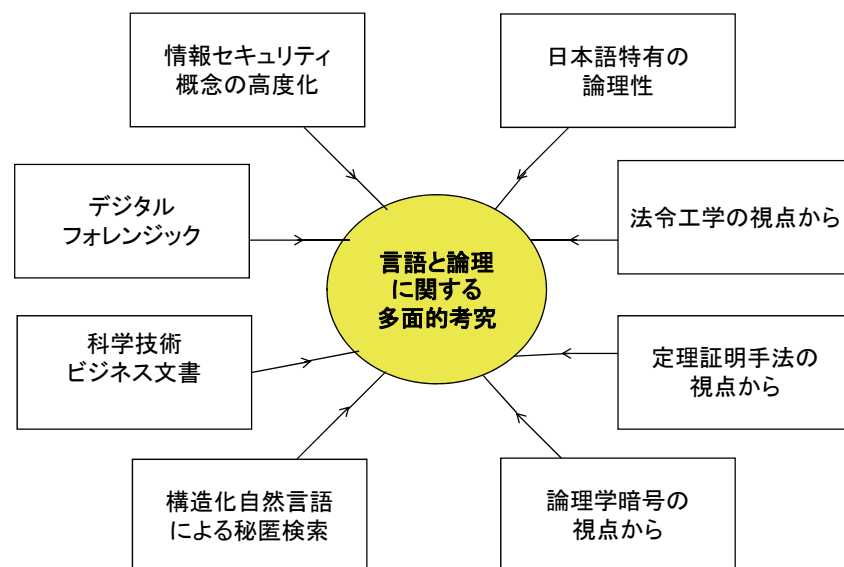


図 言語と論理に関する多面的考究の必要性

自然言語から人工言語へ

与謝野晶子は与謝野鉄幹を愛している。

与謝野鉄幹は与謝野晶子に愛されている。 OK

皆は誰かを愛している。→ 誰かは皆に愛されている。 OK?

「いかなる行為にも、その究極的目的がある」→

「ひとつの究極的目的を目指して、全ての行為がなされる」 OK?

アリストテレス ニコマコス倫理学

「感覚的な事物の存在は、全て何らかの精神に依存している」→

「ある精神が、全ての感覚的な事物の存在に不可欠である」 バークレー OK?

自然言語では隠されている構造→論理形式が記号表現の表面に現れている言語

命題論理から述語論理へ

- 命題論理(文を単位とする推論)やアリストテレスの三段論法では、主語・述語の関連が消えてしまって、推論の正当化が出来ない。
- アリストテレスの方法
- 「全てのAはBである。」項に一定の記号を与える。
- 「全ての」とか「ある」の部分はそのまま温存。
- 正当化には 文を構成する要素部分の取り扱いと普遍量化子・存在量化子の導入が不可欠
- → 述語論理

述語論理 項からモノ(変数)へ

- ・ パンダは笹を食べ、コアラはユーカリを食べる。
- ・ パンダもコアラも動物である。
- ・ 笹もユーカリも植物である。
- ・ -----
- ・ 質問;それならパンダもコアラも草食動物ですか。

記号表現

$F(x)$; x はパンダである。 $G(x)$; x はコアラである。
 $H(z)$; z は笹である。 $F_1(x)$; x はユーカリである。
 $G_1(x,z)$; x は z を食べる。

- ・ $\forall x (F(x) \supset \exists z (G_1(x,z) \wedge H(z)))$
- ・ $\wedge \forall y (G(y) \supset \exists z_1 (G_1(y,z_1)$
- ・ $\wedge F_1(z_1)))$

自然言語と論理学の相性は悪いか？

- 富美子は物理の得意な女性だ。
- 富美子は女性だが、物理が得意だ。
- 富美子は女性のくせに、物理が得意だ。
- 論理学は ニュアンス・感情表現は苦手？
- 科学技術、法律、ビジネス等に限れば
- かなり相性を良く出来る。

今後の展開 多様な論理学の導入が必要

1階述語論理・古典論理から多様な論理へ

神の視点(古典論理) ; 有罪か無罪かは決まっている？
判決の前に死亡するかも？

人間の視点 ; 法律や医療の相談では、**直観主義論理**、**様相論理**

など多様な論理学を導入する必要がある。logics

数学においてすら、構成主義、直観主義の立場がある($n=3, 14, \dots$)

暗号の安全性証明 IND-CCA等では専ら古典論理(背理法)だが。

親が叱らない と 子供は勉強しない→子供が勉強する と 親は叱る？

古典論理・命題論理から非古典論理・述語論理へ

- 2重否定＝肯定 を認めない
- × 排中率、背理、対偶、含意、ド・モルガンの一部
-
- Q 華子さん 好きですか？ A 好きでないことはありません
- Q それでは好きなのですね A ？
- 親が叱らない→子供は勉強しない 子供が勉強する→親が叱る？

多値論理、直観主義論理、様相論理・・・

JAIST 21世紀COE 法令工学 Davidsonian style

フレーゲ(1848－1925)の命題観

- 「伊藤 邦武著；物語 哲学の歴史」より
- 命題とは関係を表現する関数が、その定項や変項として個別的な対象を含むものであり、それに加えて、それらの対象についての量的表記を行う「量化子」を加えたものである。
- 例；師弟 という関数が ソクラテス、プラトンという値をとる。
- 新しい論理学は主語－述語の命題形式を廃棄することで、
- 命題同志の演繹的な関係の推論の力を飛躍的に増大。
- → 形而上学の大きな変革へ(意識の哲学から言語哲学へ)

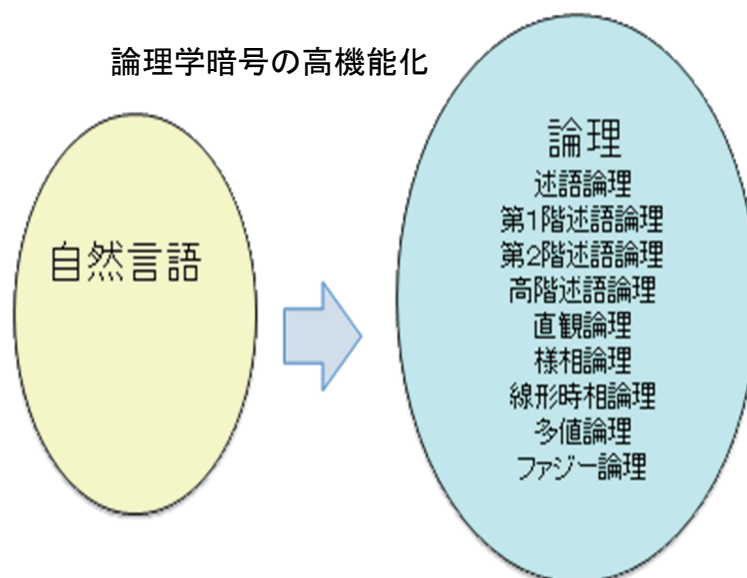
古典論理から非古典論理へ

- 古典論理＝神の論理 nearly＝数学の世界
- 排中律 2重否定＝肯定
- 真か偽か 善か悪か 好きか嫌いか 有罪か無罪か
- 人間世界 好きでないことはない＝好き？

非古典論理

- 直観主義論理 × 排中律、2重否定＝肯定
- 田中角栄 有罪？
- 様相論理 知識、信念、自覚 例；医師 薬100mg communication
-

論理学暗号の高機能化



長期的重要課題

- ー日本語文の論理性向上のための論理学・自然言語処理・機械翻訳ー
- (辻井;IDF事典原稿ー将来の課題より抜粋)
- (1) 日本語の論理性を高めること
- (2) 現代論理学と日本語を含む
- 自然言語処理の親和性を高めること

放送・通信の4類型と情報セキュリティ概念の高度化—第2報 —組織通信と公共情報コモンズ（Lアラート） Advanced Concept of Information Security in 4 Categories of Communication and Broadcast — 2 nd Report — Public Disaster Proof System (L-Alert)

辻井 重男* 吉田 正彦† 柴崎 哲也‡ 小林 正幸‡
Shigeo Tsujii Masahiko Yoshida Tetsuya Shibasaki Masayuki Kobayashi
川喜多 孝之‡
Takayuki Kawakita

あらまし Open data, Cloud, Big data の普及や、2016 年1月から予定されている My number の導入、即ち、OCBM の進展に伴い、従来の紙の文書に代わり、大量の電子化された文書が、組織間で通信される状況に変わろうとしている。本文では、(個人)通信、放送、交流サイトに、組織通信を加えて放送・通信分野を4分野に類型化し、その中で、組織通信のあり方を考察する。組織間の通信では、個人間の自由な会話と異なり、伝達情報の正確性、信頼性、迅速性、緊急性、機密性、送信情報限定性、論理的無矛盾性、各組織が公表している個人情報保護規定との整合性などのコンプライアンス性、或いは、多言語性などが厳しく求められる場合が多い。現在、C・I・A、即ち、機密性（Confidentiality）、完全性（Integrity）、利用可能性（Availability）が、情報セキュリティの3要素とされているが、本文では、組織通信における情報セキュリティ概念を上記のように高度化して考察することとする。組織通信の例として、現在、総務省と一般財団法人 マルチメディア振興センターが協力して、全国的展開を推進している公共情報コモンズ（Lアラート）について考察する。

キーワード 放送、通信、交流サイト、組織通信、情報セキュリティ、公共情報コモンズ、Lアラート、組織暗号

1 放送通信の4類型—公共放送、個人放送、公共通信、個人通信

20 世紀末までは、情報伝達は、放送と通信に大別され、放送は公序良俗、通信は憲法で保証された個人間通信の秘密を守るといった別々の価値観の下に共存していた。

公共放送については、NHK も民間放送も、有害情報や個人情報保護に配慮するなど、法的・倫理的な面や論理的整合性など様々な面からチェックがなされている。放送業界は、自ら放送倫理・番組向上機構を設け、外部評価を受け入れている。

他方、今世紀に入り、フェイスブックやツイッターなど様々な SNS が生まれている。SNS は、誰でも自由に、

いわば個人放送局を開設して表現の自由を楽しむことが出来るようになっている。だが、比較的自由な世界であるだけに、名誉毀損、プライバシー侵害、あるいは児童ポルノなどの問題を生じている。児童ポルノの問題では、表現の自由と倫理との相克が深刻な課題となっており、緊急避難法理により、実在児童が被害を受けないよう規制されている。

1900 年頃、哲学者ヘーゲルは、「歴史とは、自由拡大のプロセスであり、自由の拡大に伴って矛盾も拡大する」と述べているが、最近の IT、特にビッグデータの浸透などを見ると、ヘーゲルの歴史の法則の妥当性を実感させられる。

このような複雑な様相を帯びた SNS は、個人放送局の集合体と見ることも出来るので、放送を、公共放送と個人放送という対立概念で捉えることも出来る。

放送の世界と対称的に、通信の世界では、通信傍受法

* 中央大学研究開発機構、〒112-8551 東京都文京区春日 1-13-27, Research and Development Initiative, Chuo University, Kasuga 1-13-27, Bunkyo-ku, Tokyo, 112-8551 Japan

† 総務省

‡ (一財) マルチメディア振興センター

に抵触するような犯罪に関わる会話は別として、これまで、殆ど制約の無い世界であった。しかし、今後は、Open data, Cloud, Big data の普及や、2016 年 1 月から予定されている My number の導入、即ち、OCBM の進展に伴い、従来の紙の文書に代わり、大量の電子化された文書が、組織間で通信される状況に変わろうとしている。即ち、個人通信とは異なる通信の世界として、今後、組織通信が広がろうとしている。

個人通信の場合、通信の秘密が憲法によって保証される中で、通信相手以外の他者に対しては、責任を強く意識することなく、比較的自由に会話出来る状況にあるが、今後、自治体、企業、医療・介護ネットワークなどの間の組織通信においては、組織体としての、正確性、信頼性、確実性、迅速性、送信情報の限定性、耐コンプライアンス性、論理的無矛盾性や、多言語性などが求められるようになる。即ち、組織通信は、いわば公共通信という性格を帯びてくると言える。

以上を要約すると、やや割り切った分類にはなるが、21 世紀に入って以降、図 1 に示すように、放送の分野では、公共放送に対応して、個人放送が生まれ、通信分野では、個人通信に対して、公共通信が拓けようとしていると言える。

組織通信（公共通信）においては、従来、1 つの病院内で処理していた患者の情報も、ケアマネージャ、介護士、看護師やヘルパーなど、様々な方々からなる医療・介護ネットワークに流されるようになりつつあり、各職種の人に見せる患者情報を制限や個人情報保護法への配慮が必要なケースも生じている。

現在、中央大学研究開発機構では、独立行政法人 情報通信研究機構（NICT）からの委託研究を受け、組織通信という概念を提唱すると共に、その一つの特性として組織暗号と名付ける方式を開発し、長野県や新潟県において、実証実験を行っている。組織暗号とは、組織通信において、暗号文を平文に戻す機会を最小限に抑える機密通信方式である。今、県庁や市町村、或いは病院など異なる組織間で、送信側組織の担当者が、ある文書がある組織に送る場合、受信側担当者が不明な場合、受信側組織の代表者（庶務課長、事務局長等）に、暗号化された文書と、その平文のラベルを送り、受信側組織の代表者は、原則として、暗号文を復号することなく、ラベルにより、適切な担当者に再送信する方式である。暗号方式としては、楕円暗号や多変数公開鍵暗号が考えられるが、現在、楕円エルガマル暗号により鍵管理が安全に行える、

図 2 に示すような方式を開発し、自治体を対象に実証実験を行い、関係者から好評を得ている。

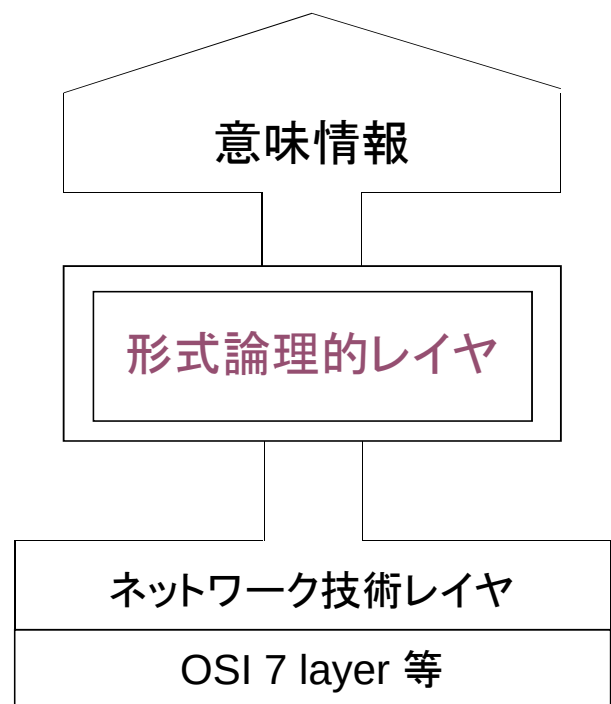


図 3: 情報通信の新階層

2 公共情報コモンズ（Lアラート）

（一財）マルチメディア振興センター（理事長 辻井重男）では、公益目的事業として、総務省と協力して、公共情報コモンズというプロジェクトを全国的に展開している。このプロジェクトは、災害発生時などに各自治体や消防庁などから、ばらばらに発信されていた情報を、データ形式などを統一し、放送メディアなどに伝達する社会システムを構築しようというプロジェクトであり、平成 19 年 7 月の新潟県中越沖地震を発端として、検討が始められたものであるが、東日本大震災以降、日本各地の自治体との連携が急速に進められている。この際、発信者側では、情報の正確性をよくチェックした上で、公共情報コモンズを通して放送メディアなどに向けて情報発信することが必須である。公共情報コモンズは、送信側組織としての各地域の自治体、国土交通省の組織、消防庁、気象庁、ライフライン系組織（電力、ガス、水道、交通機関など）と、受信側組織としての NHK・民間放送、ヤフーなどのメディアとの間の組織通信である。

2014 年 11 月現在、運用を開始した都道府県数は 22、準備中のものを含めると 35 に達しており、順調に利用・運用が進んでいる。2014 年度に公共情報コモンズへ発信された避難勧告の件数は、2014 年 10 月末までの時点で約 2600 件と、2013 年度の 4 倍を超えている。

以下、図 5、図 6、図 7、図 8、図 9、図 10 に公共情報コモンズ概念、システム構成、運用状況などを示し

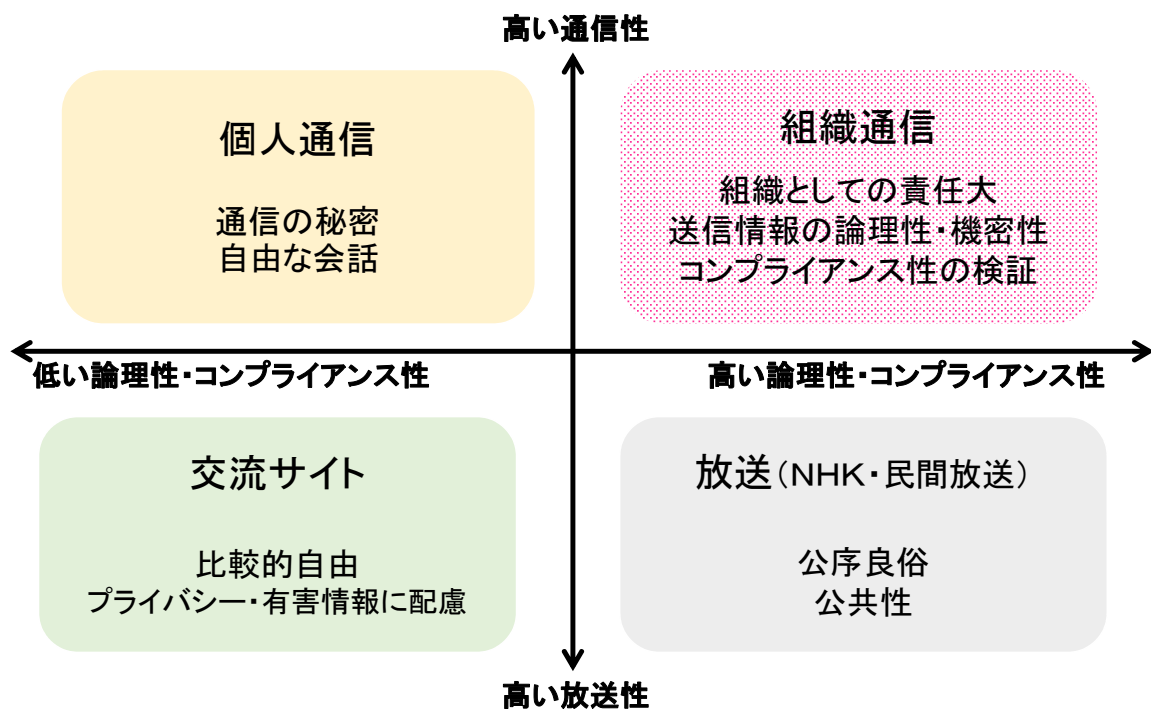


図 1: 放送・通信の 4 類型と組織通信

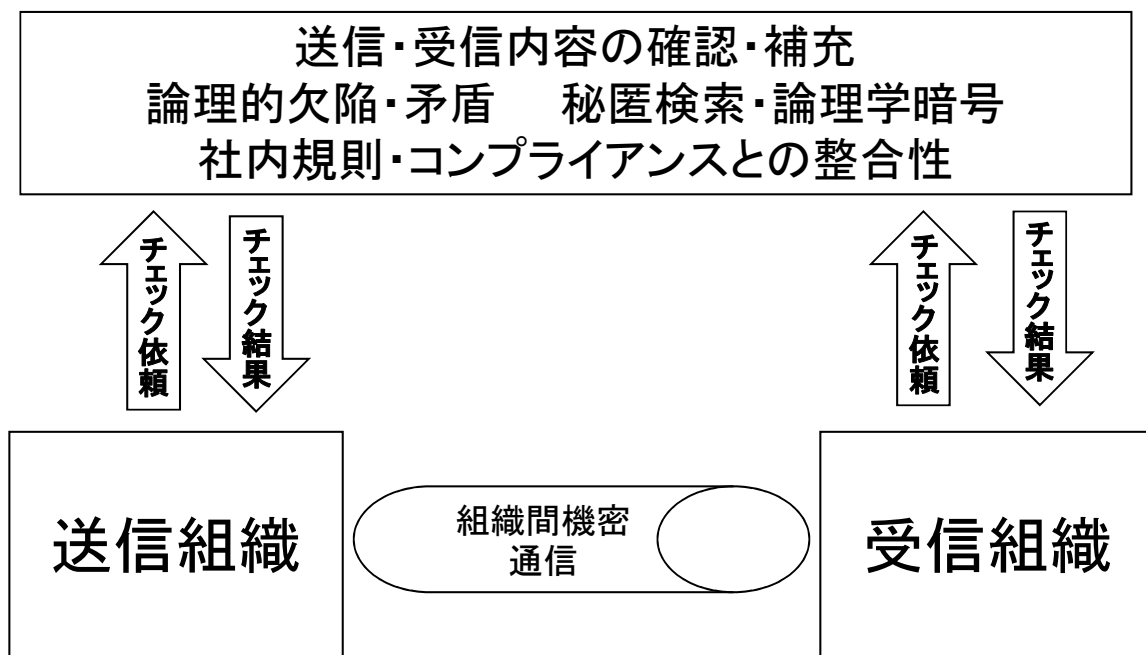


図 2: OCBM (Open data, Cloud, Big data, My number) 環境下での情報通信・セキュリティ概念の高度化への対応

ておく。

謝辞

本研究は、独立行政法人 情報通信研究機構 (NICT) の委託研究「組織間機密通信のための公開鍵システムの

研究開発—クラウド環境における機密情報・パーソナルデータの保護と利用の両立に向けて」の一環として行ったものである。公共情報コモンズ (L アラート) の推進に尽力しておられる総務省、並びに (一財) マルチメディア振興センターの関係各位に謝意を表する。また、本論

IEEE等国際会議・総務省シンポで賛同的評価

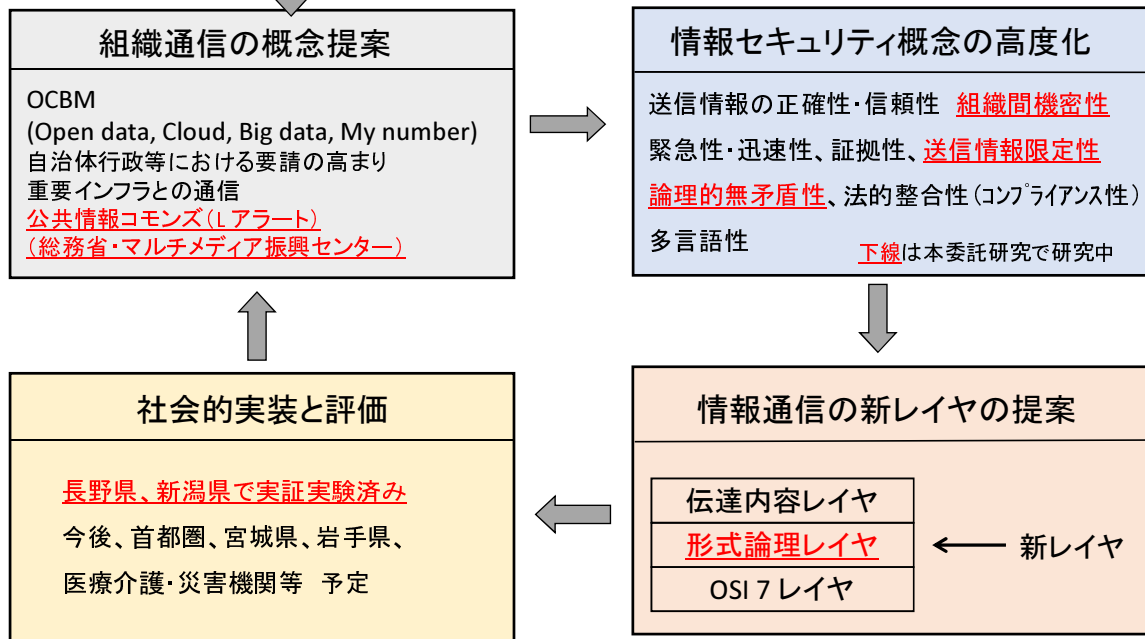


図 4: 組織通信概念の提案とその実現に向けて

多様なメディアを活用して、高齢者をはじめ誰もが、いつでもどこでも、分かりやすい形式で地域の安心・安全に係るきめの細かい情報を迅速に入手できるようにするための具体的な仕組みを、早急に構築することが喫緊の課題

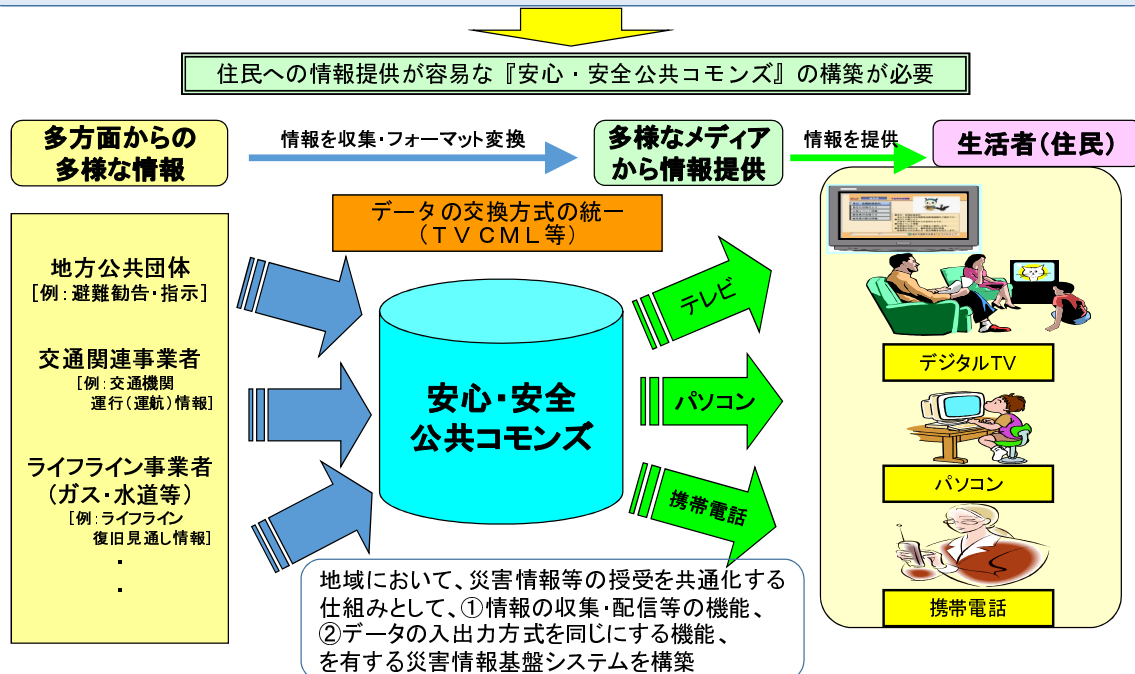


図 5: 『安心・安全公共コモンズ』の構築(総務省「地域の安心・安全情報基盤に関する研究会」報告(2008年7月))

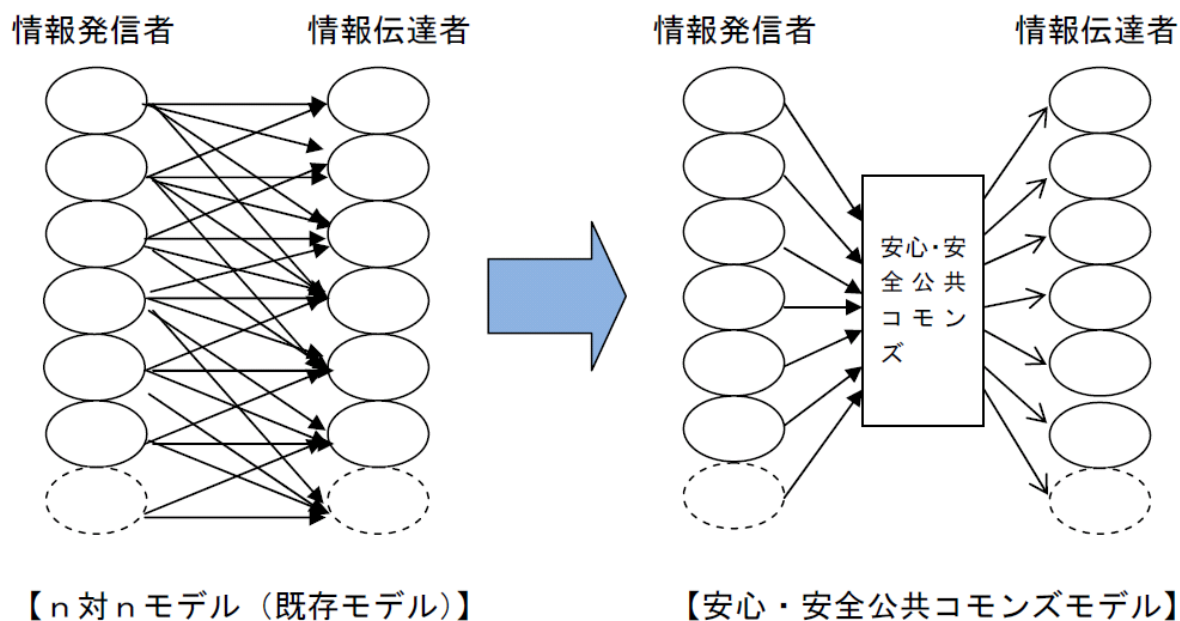


図 6: 「N 対 N モデル」と「安心・安全公共コモンズモデル」（文献 [8] P7 より）

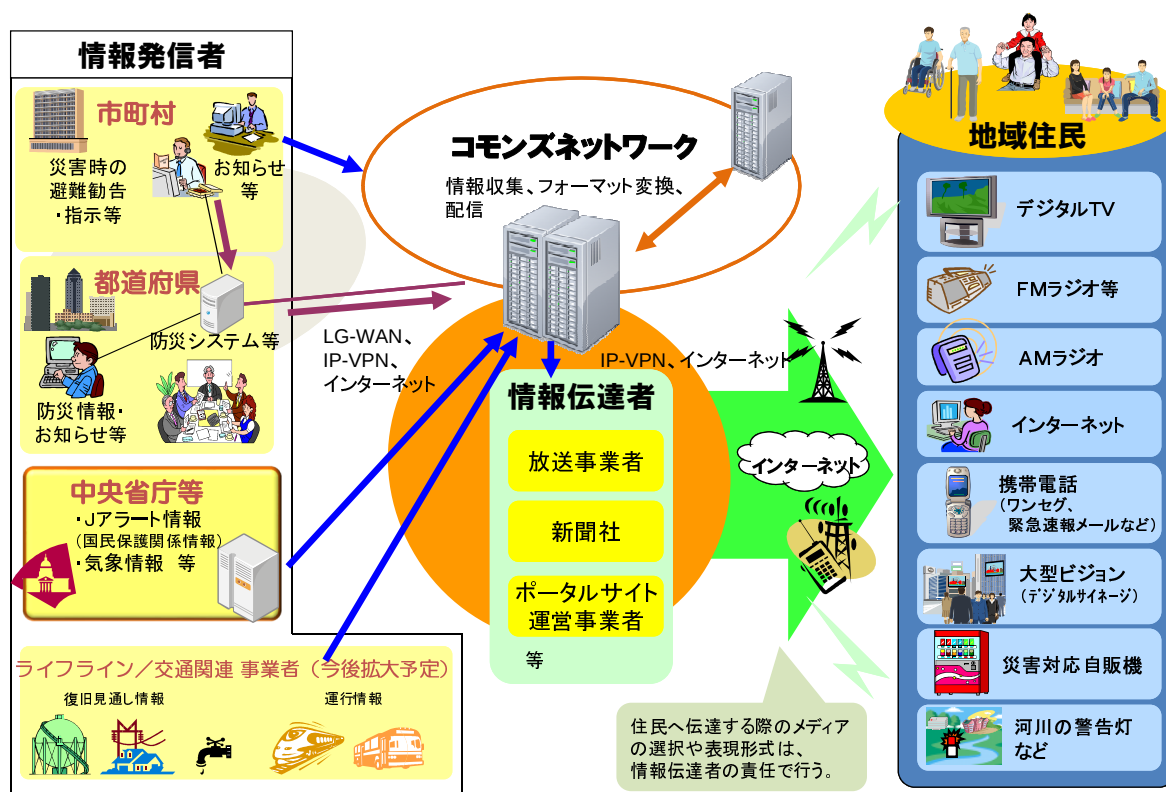


図 7: 公共情報コモンズの利用イメージ

文の編集に協力頂いた藤田 亮 中央大学研究開発機構専任研究員・助教に感謝する。

参考文献

- [1] 辻井重男, “放送・通信の4類型と情報セキュリティ概念の高度化 — 組織通信・組織暗号の普及に向

情報種別	サービス開始時点	平成25年6月3日～	平成26年6月16日～
気象警報・注意報※	○	○	○
気象特別警報・警報・注意報※		※	○
指定河川洪水予報	○	○	○
土砂災害警戒警報	○	○	○
記録的短時間大雨情報		○	○
竜巻注意報		○	○
震度速報		○	○
震源に関する情報		○	○
顕著な地震の震源要素更新のお知らせ		○	○
地震回数に関する情報		○	○
地震の活動状況等に関する情報		○	○
震源・深度に関する情報		○	○
津波情報		○	○
津波警報・注意報・予報		○	○
沖合の津波観測に関する情報		○	○
噴火警報・予報			○

※ 特別警報については、既存の気象警報・注意報においても、見出し文のようなところに、“【特別警報(大雨、暴風、波浪、高潮)】”などと追記

図 8: 防災気象情報（気象情報・注意報など）の対応状況

- ・ 発令情報と報告情報をマージした形でコモンズに情報発信する必要がある

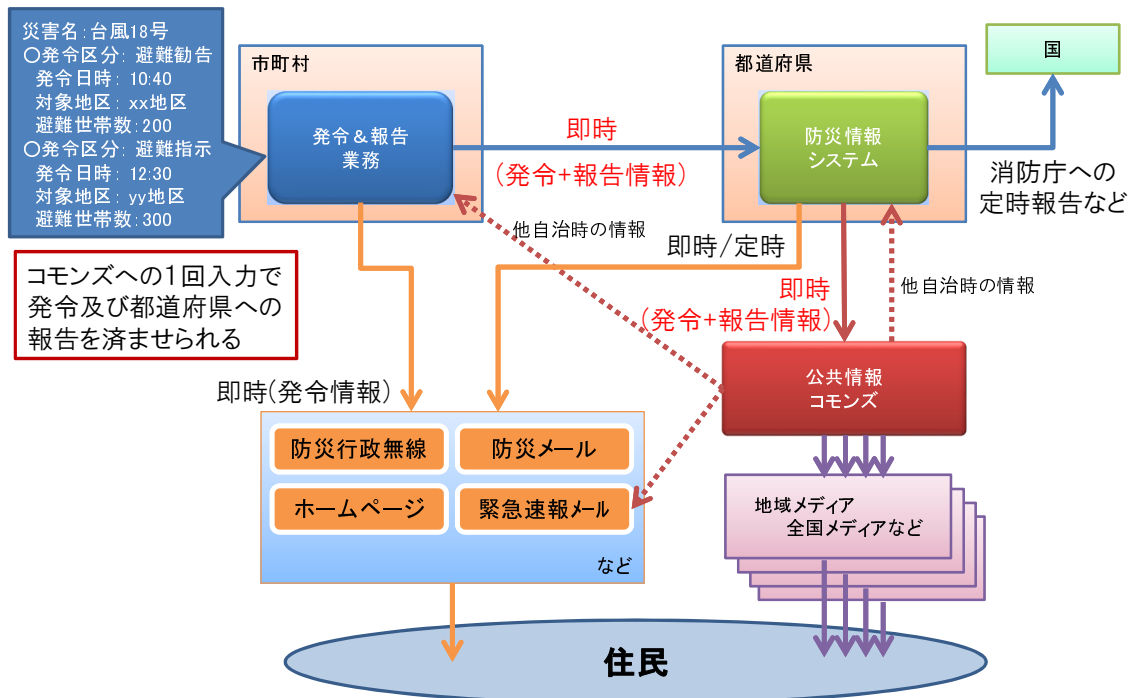


図 9: コモンズ連携後の情報の流れ

○ コモンズ連携システムによる情報収集

利用者が開発したコモンズ連携システムを接続して情報の収集を行う。

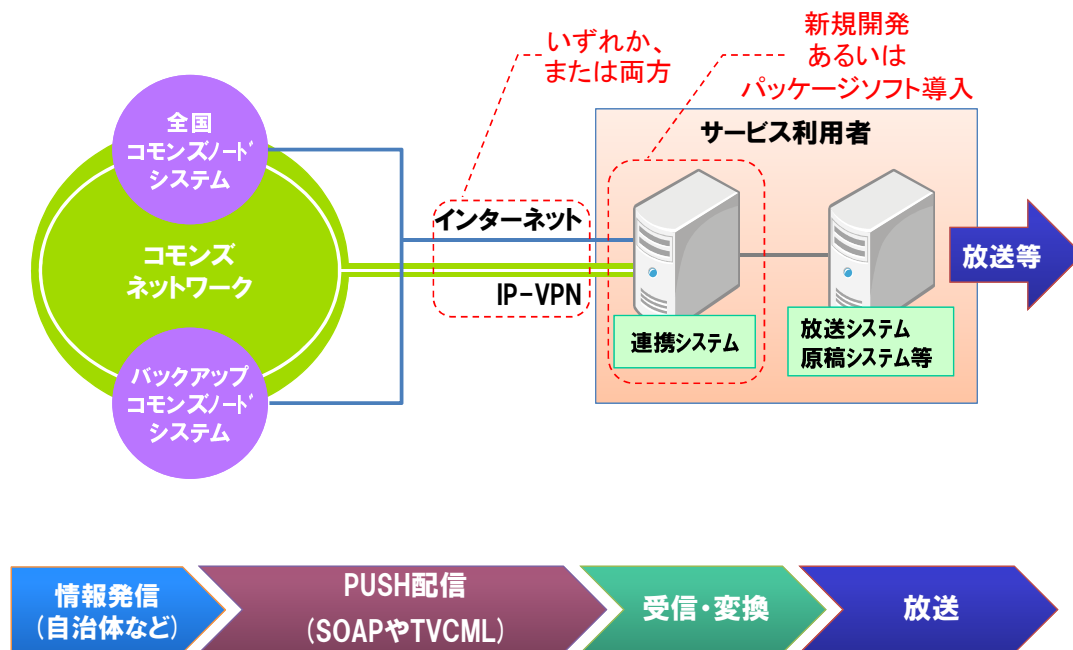


図 10: 公共情報コモンズの情報閲覧（自動連携）

けて —,” *Proc. SCIS2014*, 2C4-4, January 2014.

- [2] 辻井重男, “自由, 安心, プライバシーと三止揚 — MELT up — 放送・交流サイト・個人通信・組織通信の枠組みの中で —,” 民放 経営四季報, No. 101 (2013.9), 日本民間放送連盟 研究所.

- [3] Shigeo Tsujii, Hiroshi Yamaguchi, and Masahito Gotaishi, “Advanced Concept of Information Security as a Comprehensive Science,” *Transdisciplinary Journal for Engineering and Science*, Vol. 5, pp. 14-32, December, 2014.

- [4] Hiroshi Yamaguchi, “Secure and Privacy Preserving Data Services in Cloud, —Cryptosystem for Inter-Organizational Communications—,” *The SDPS (Society for Design & Process Science) -2013, San Paul, Brazil*

- [5] Hiroshi Yamaguchi, “Inter-Organizational Communication and Advanced Communication,” *The ATLAS (The Academy of Transdisciplinary Learning and Studies) -2014, Asia University, Taichung, Taiwan, June 8, 2014.*

- [6] 辻井重男, “組織間コミュニケーションの展望と課題,” 日本計画行政学会 第 37 回全国大会ワークショップ, 2014 年 9 月 13 日.

- [7] 吉田正彦, “公共情報コモンズ（Lアラート）の展望と課題 — 組織間通信の文脈で —,” 日本計画行政学会 第 37 回全国大会ワークショップ, 2014 年 9 月 13 日.

- [8] 総務省, “「災害時等の情報伝達の共通基盤の在り方に関する研究会」報告書,” 2014 年 8 月.

- [9] （一財）マルチメディア振興センター, “公共情報コモンズの概要と運用について— 地域の安心と安全のために—,” 2014 年 8 月 22 日.

尸位素餐（しいそさん）とは？
—ビッグデータ時代の日本語の論理性と国際性

このところ、NHKから「第2次大戦と暗号」について取材を受けている。我々、暗号研究者は、1970年代以降の、情報社会の基盤としての暗号を現代暗号と呼び、20世紀前半までの軍事・外交用暗号を古典暗号と呼んで区別している。

2016年からマイ・ナンバーの使用も始まり、本人確認やデジタル署名の基盤性は増すばかりだから、公開鍵暗号を初めとする現代暗号についても、メディアの関心を期待したいのだが、そうはいかないのが悩ましいところだ。

それはさておき、改めて古典暗号関連の資料を読み直してみると、それはそれで、これからの日本再生に向けて考えさせられることも多い。

冒頭にあげた「尸位素餐」という言葉をご存知だろうか。私も最近、小松啓一郎著「暗号名はマジック—太平洋戦争が起こった本当の理由」で初めて知ったのだが、「尸位素餐」とは、「高い官位にありながら、満身に職責を果たさず、給料だけを貰っていること」だそうである。

昭和16年（1941年）12月8日の真珠湾攻撃に始まる日米開戦を回避すべくアメリカとの折衝に当たっていた、野村吉三郎駐米大使は、遅々として進展しない外交交渉に失望し、「このまま、成果が挙がらないのに、給料を貰い続けるのは、心苦しいので、辞任したい」という気持ちを「尸位素餐」の四文字に託して、米国から日本に向けて暗号電文で送信したのである。この電文を傍受した、米国の暗号解読者は、日本人でも分からない漢語が分かる筈もなく、別の意味にすりかえて、米国政府に報告したのである。

これは一例であるが、暗号解読はできても、その翻訳ができず、好戦的に解釈された事例が多かったようである。こうした誤訳の積み重ねがなければ、日米開戦は避けられたかも知れないと、上記の著者は述べている。

古い話を持ち出したが、現在でも翻訳は難しい問題を孕んで

いる。例えば、情報通信分野の国際標準を日本語に、ニュアンスが正確に伝わるように翻訳するのも難しい作業である。

最近の流行語とも言えるビッグデータの正確な定義はないようであるが、多様で膨大な非定型なデータをさすことが多い。従って、文学・小説などは別として、産業、科学技術、法令などの多岐に亘る分野で、日本語で書かれた多くの文書も広い意味ではビッグデータと見なければならない。この場合、日本語の論理性と国際性が問題となる。

日本語の論理性が低いかどうかは、形態素解析（単語レベル）、構文解析（文法レベル）、意味解析、文脈解析という言語解析の4段階のどのレベルで考えるかによる。機械翻訳の場合は、形態素解析、構文解析の論理性が高い方が処理し易い。人間は、意味解析のレベルで理解できることが多いが、構文解析レベルでの論理性を高めないと誤解を招く可能性も生じる。

「赤いお墓の彼岸花」と聞けば、赤いのは、「お墓」でなく「彼岸花」だと理解できるが、

「眠れる森の美女」で、寝ているのは？と聞けば、「美女」と答える人が多い。だが、フランス語では、文法的に明確に「森」を指している。このような誤解を避ける上からも、形態素解析、構文解析レベルでの日本語の論理性を高める必要があるようである。

イギリスでは、17世紀末、シェークスピアから100年近く経った頃、「我々は、このような情緒的で感覚的な言語を使っているのか」と言う反省が起こり、言語の大改革を実行したそうである。

日本語の場合、平安時代は、和歌などが中心で、論理性よりも情緒性が重視されたが、鎌倉時代になると、事務処理も増え、論理性が高まったようである。近代に入り、明治の開国期、そして、第2次大戦後と言語の改革が進められたが、現在、改めて、言語について考える時期を迎えているように思われる。

尤も、これからは、ビジネス言語は英語で書かれるようになるだろうという意見も多いが、そういうケースが増えるにせよ、日本語について考える必要性が高まっていることは否定できない。

デジタルフォレンジックについては、米国における訴訟で、日本語文書の翻訳文を証拠書類として提出する際の、労力、コスト、時間が大きな問題となっており、法令や社内規則などの機械翻訳も要請されよう。ここで、法令分野について考えてみよう。

憲法のように理念を掲げる場合は別として、一般の法律や条

令などは、論理的に明解でなければならない。北陸先端科学技術大学院大学では、片山卓也前学長が研究リーダーとして推進した21世紀COE以降、法令工学プロジェクトを進めている。法令工学とは、法令を社会のソフトウェアと考えて、法令間の矛盾、たとえば、法律と条例との矛盾の検証や、法令文自体の論理性のチェック等の情報処理を進める学問である。法令工学では、日本語の論理性を高める研究、及び、自然言語処理に適するように、直観主義論理や様相論理などを用いて、論理学を多様化・高度化する研究という両面からのアプローチによる考察を深めている。

私も、2年ほど前から、暗号文を論理式で表すという発想で、研究を開始した関係で、論理学を俄か勉強していることもあって、法令工学の動向に興味深く見守っている。

いずれにしても、ビッグデータ時代には、日本語の論理性と国際性を高める研究は不可欠であろう。

招待講演 2

特許審査情報の世界への発信と産業日本語

大町 真義

特許庁審査第二部福祉機器 審査長

特許審査情報の世界への発信と 産業日本語 ーグローバル知財戦略支援の取組ー

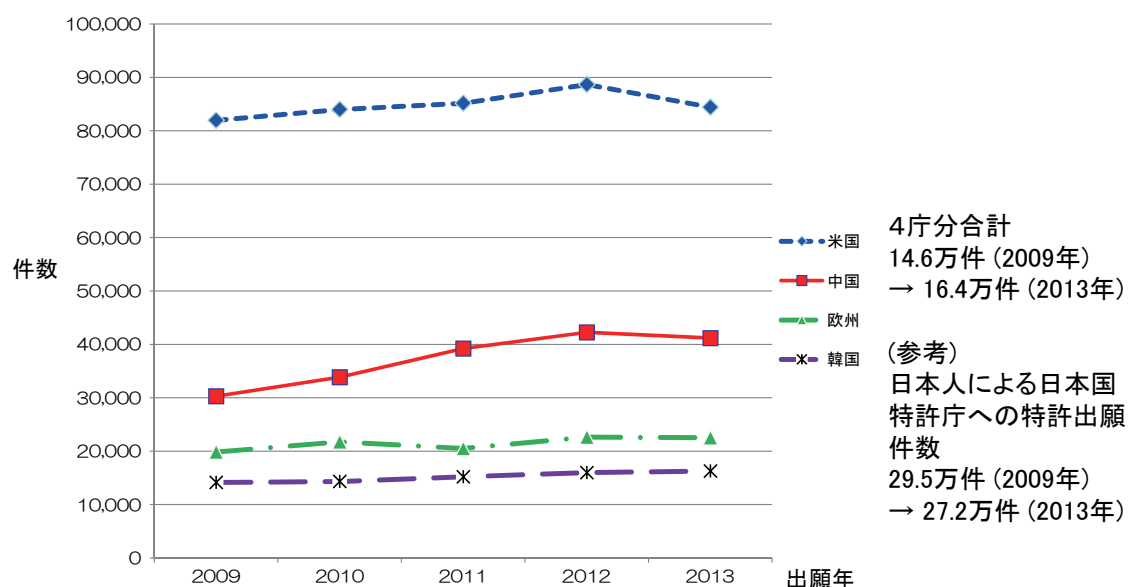
2015年2月24日

第6回 産業日本語研究会・シンポジウム

特許庁
審査第二部 福祉機器 審査長
大町 真義

はじめにー日本から海外への特許出願

■ 日本人による主要特許庁への特許出願件数の推移



データ出所: 特許庁「特許行政年次報告書 2014年版」

はじめに － 日本から海外への特許出願

1. 審査情報参照システム － その意義と発展

2. 拒絶理由通知の日英機械翻訳

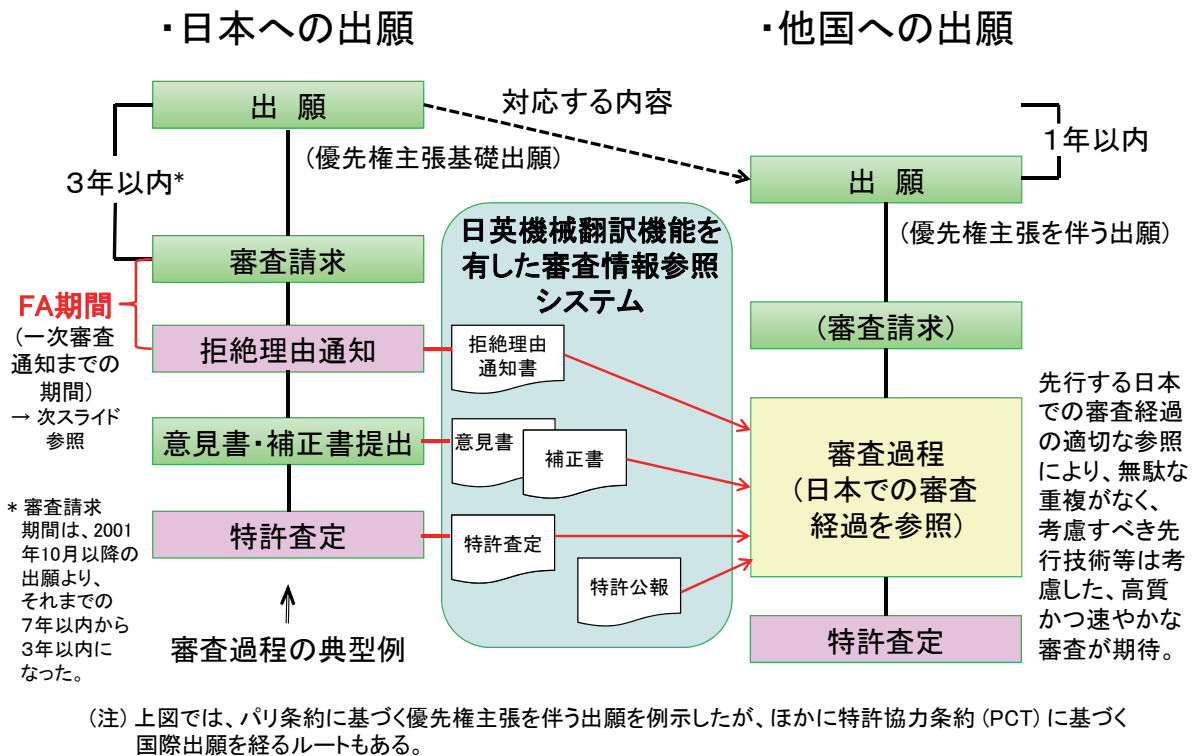
－ 定型性と構文解析：翻訳例を参照しながら

3. 特許庁審査部内での取組

－ 拒絶理由通知作成時の留意事項の取りまとめ

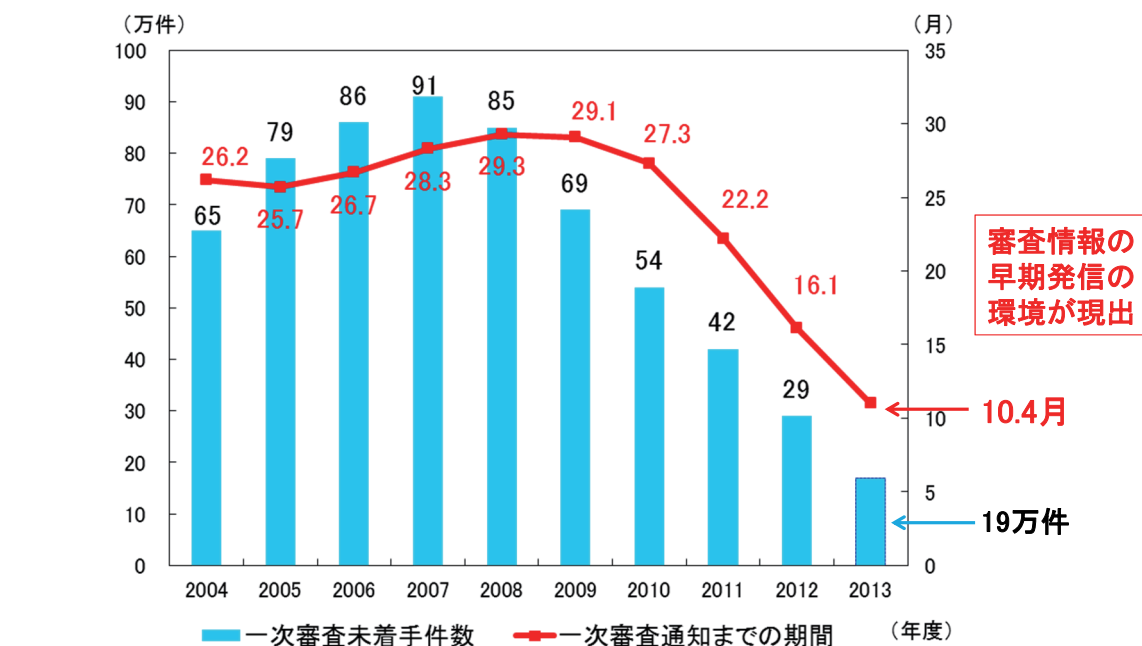
おわりに

1. 審査情報参照システム － その意義と発展



4

FA期間 (審査請求から一次審査通知までの期間) 短縮 → 10年目標の「FA11」達成



データ出所: 第1回審査品質管理小委員会 資料2

(注) 当該期間は、各年度末における値。

5

■ AIPN (Advanced Industrial Property Network; 高度産業財産ネットワーク)

日本国特許庁における特許出願の審査に関する情報を、海外の特許庁の審査官に対して参照可能とする、ウェブ・ベースのシステム。日英機械翻訳機能を有する。

- 具体的な提供情報: (1) 英語抄録 (PAJ; 人手翻訳), (2) パテント・ファミリー, (3) 法的状態 (リーガル・ステータス), (4) 公開公報, (5) 拒絶理由通知書, 意見書・補正書, 特許査定等、審査経過を示す書類, (6) 引用文献 (ファミリー文献を参照可能), (7) (特許されている場合) 特許公報
- 言語: 英語; 上記 (4)~(7) は、日英機械翻訳機能による (上記 (6) の引用文献については、日本語文献の場合、日英機械翻訳機能を利用可能)*。

(* 2013年3月より、他の種々の言語への機械翻訳の機能も利用可能。)

- 歴史: 2002年 開発開始、2004年10月 供用開始。

(参考: 拙稿「AIPN (高度産業財産ネットワーク) の生い立ちと発展 - 審査結果発信手段としての意義 -」 Japio YEAR BOOK 2009, 76-81頁)

- 利用庁: 世界 65の特許庁が利用可能。

■ ワン・ポータル・ドシエ (One Portal Dossier)

日米欧中韓の特許庁 (五庁) 間で、審査経過を示す書類を相互に参照可能とするシステム。日英・中英・韓英の機械翻訳機能を有する。

- 具体的な提供情報: 日本国特許庁からの提供情報は、AIPNにおけるものと同等。
- 言語: 英語。非英語原文も参照可能。
- 歴史: 2013年7月 供用開始。
- 展開: 2014年3月、世界知的所有権機関 (WIPO) が運営する審査情報参照システムであるWIPO-CASE (Centralized Access to Search and Examination) と接続することにより、日本国特許庁は、五庁以外の豪州、カナダ、英国、及びイスラエルの特許庁とも、相互に審査情報を参照できることに。

2. 拒絶理由通知の日英機械翻訳 — 定型性と構文解析：翻訳例を参照しながら

拒絶理由通知においてよく用いられる構成

■ 進歩性欠如の拒絶理由通知における典型的な流れ

主引用文献に記載された発明 (主引用発明) の認定 → 本願発明 (審査対象の発明) と主引用発明との対比 (一致点・相違点の認定) → 相違点についての判断 (副引用文献を引用)

■ 進歩性欠如の拒絶理由通知における典型的な文章表現

引用文献1には、次の発明 (以下、「引用発明」という。) が記載されている (段落[0041]—[0048]及び図5を参照)。

「……」

・一致点及び相違点

本願発明1と引用発明とを対比すると、両者は、「… と、 … と、 … とを有する〇〇装置」である点で一致し、次の点で相違する。

[相違点]△△機能を果たす手段が、本願発明1では「A」であるのに対し、引用発明では「B」である点

・相違点についての判断

引用文献2には、引用発明と同じ〇〇装置の技術分野において、△△機能を果たす手段として「A」を用いることが記載され、それにより構成の簡素化が可能になることも記載されている (段落[0112]及び図8を参照)。

そして、同等の機能を実現しつつ構成の簡素化を図ることは、当業者にとって自明の課題である。そうすると、引用発明において、△△機能を果たす手段として、「B」に代えて、引用文献2に記載された「A」を採用することは、本願出願時において、〇〇装置分野の当業者が容易に想到し得たことというべきである。

本願発明1と引用発明とを対比すると、両者は、「アイコンの機能説明を表示させる機能を実行させる機能説明表示手段、および所定の情報処理機能を実行させるための第2のアイコンを表示画面に表示させる表示手段と、前記表示手段の表示画面上に表示されたアイコンを指定する指定手段と、前記指定手段による、機能説明表示手段の指定に引き続く第2のアイコンの指定に応じて、前記表示手段の表示画面上に前記第2のアイコンの機能説明を表示させる制御手段とを有することを特徴とする情報処理装置」である点で一致し、次の点で相違する。

[相違点]アイコンの機能説明を表示させる機能を実行させる「機能説明表示手段」が、本願発明1では「アイコン」であるのに対し、引用発明では「スクリーン／メニュー・ヘルプ」アイテムである点

(出所：知財高判 平成17年9月30日 (平成17年 (ネ) 第10040号特許権侵害差止請求控訴事件) に基づき若干改変；「」内は原文のまま)

文例に現れている構造上の特徴点－定型性

■ 一致点の認定部分

かぎ括弧 (引用括弧) で括られた形で、本願請求項の発明特定事項の一部が記載されている。

- かぎ括弧内が一致点であることは、閉じかぎ括弧の直後に示されている。
(「…情報処理装置」である点で一致し …)
- かぎ括弧内が具体的に何であるかは、閉じかぎ括弧の直前の名詞として表れている。(「…情報処理装置」) なお、通常、当該名詞は本願請求項の末尾の名詞と一致する。
- かぎ括弧内においては、発明の構成要素が、並列構造をもって列記されている。
(「…表示手段と、…指定手段と、…制御手段とを有する …」)

■ 相違点の認定部分

体言止めの形で、本願の請求項に係る発明と主引用文献に記載された発明との間の相違点が対比的に記載されている。

- 相違点の認定部分の直前に、相違点を記す旨の記載がなされている。
(「…次の点で相違する。[相違点]…」)
- 相違の内容が、「…では…であるのに対し、…では…である点」という表現により示されている。(「…機能を実行させる「…手段」が、本願発明1では「…」であるのに対し、引用発明では「…」である点」)

When the invention in this application 1 and a cited invention are contrasted, [both]

"The displaying means which displays on a display screen the second icon for performing the functional description displaying means which performs the function on which the functional description of an icon is displayed, and a predetermined information processing function,

The setting means which specifies the icon displayed on the display screen of the aforementioned displaying means,

According to specification of the second icon which follows specification of the functional description displaying means by the aforementioned setting means, on the display screen of the aforementioned displaying means, it corresponds at the point which is the information processing apparatus having a control means on which the functional description of the above-mentioned second icon is displayed", and is different the following point.

[Point of difference] Point that the "functional description displaying means" which performs the function on which the functional description of an icon is displayed is a screen / "menu help" item in a cited invention to being an "icon" in the invention in this application 1

, while it is

12

そうすると、引用発明において、アイコンの機能説明を表示させる機能を実行させる「機能説明表示手段」として、「スクリーン／メニュー・ヘルプ」アイテムに代えて、引用文献2に記載された「アイコン」を採用することは、本願出願時に情報処理装置分野の当業者が容易に想到し得たことというべきである。

(出所: 同じく 知財高判 平成17年9月30日 (平成17年 (ネ) 第10040号) に
基づき改変)

If it does so, [as a "functional description displaying means" which performs the function on which the functional description of an icon is displayed in a cited invention]

It should be called things at which a person skilled in the art could have easily arrived of the information processing apparatus field at the time of submission of this application to replace with a screen / "menu help" item, and to adopt the "icon" described in cited document 2.

13

3. 特許庁審査部内での取組 — 拒絶理由通知 作成時の留意事項の取りまとめ

14

取組の概要

■ 取組主体

特許庁の審査部（審査第二部）内における、若手審査官4名のワーキング・グループ。

■ 取組内容

日本語としてわかりやすくかつ機械翻訳された場合にも理解されやすい拒絶理由の起案を行うために留意すべき事項の取りまとめと、部内審査官への周知。

■ 活動期間

2014年10月～12月

■ 活動成果

2つの「鉄則」と4つの「留意事項」の取りまとめ、及び、説明会を通じた周知。

- 鉄則1：長文を避ける
- 鉄則2：入れ子構造を避ける
- 留意事項1：修飾・被修飾の関係を明確にする（読点の効果的な使い方）
- 留意事項2：あいまいな接続語を使わない
- 留意事項3：主語・述語の関係を明確にする
- 留意事項4：多義的な表現を避ける

15

鉄則1：長文を避ける

	起案文	AIPN機械翻訳 ※[]は翻訳不良の場合に現れます。
問題のある表現	引用文献1に記載された発明と引用文献2に記載された発明は、同一の技術分野に属し、かつ、強固な固定を実現するという共通の課題を持つから、引用文献1に記載された発明において、両部材の固定手段として引用文献2に記載された発明の固定手段を採用することは当業者が容易に成し得るものである。	[the invention described in cited document 1, and the invention described in cited document 2] It belongs to the same technical field, and since it has the common problem that firm fixing is realized, in the invention described in cited document 1, a person skilled in the art can accomplish easily adopting the fixing means of the invention described in cited document 2 as a fixing means of both members.
改善例	引用文献1に記載された発明と引用文献2に記載された発明は、同一の技術分野に属し、かつ、強固な固定を実現するという共通の課題を持つ。 <u>したがって</u> 、引用文献1に記載された発明において、両部材の固定手段として引用文献2に記載された発明の固定手段を採用することは当業者が容易に成し得るものである。	The invention described in cited document 1 and the invention described in cited document 2 belong to the same technical field, and has the common problem that firm fixing is realized. <u>Therefore</u> , in the invention described in cited document 1, a person skilled in the art can accomplish easily adopting the fixing means of the invention described in cited document 2 as a fixing means of both members.

16

鉄則2：入れ子構造を避ける

	起案文	AIPN機械翻訳
問題のある表現	引用文献1には、 <u>コネクタがピンを有するハウジングを有する構成</u> が記載されている。	The composition which has <u>a housing in which a connector has a pin</u> is described in cited document 1.
改善例	引用文献1には、 <u>コネクタがハウジングを有し、当該ハウジングがピンを有する構成</u> が記載されている。	The composition in which <u>a connector has a housing and the housing concerned has a pin</u> is described in cited document 1.

問題のある表現

パターン1：コネクタが〔ピンを有するハウジング〕を有する

パターン2：〔コネクタがピンを有する〕ハウジングを有する

入れ子があるために、用語同士のつながりが明確でない。

改善例

〔コネクタがハウジングを有し〕〔当該ハウジングがピンを有する〕

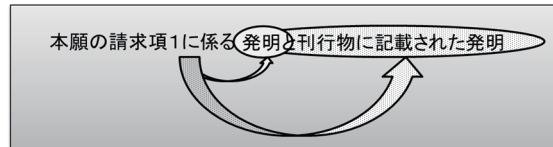
2つの文に分けることで、用語同士のつながりが明確となった。

17

留意事項1：修飾・被修飾の関係を明確にする

	起案文	AIPN機械翻訳
問題のある表現	本願の請求項1に係る発明と 刊行物に記載された発明とを 対比すると、	If the invention and invention disclosed in a publication concerning Claim 1 of an application concerned are contrasted,
改善例	本願の請求項1に係る発明と、 刊行物に記載された発明とを 対比すると、	If <u>the invention concerning Claim 1 of an application concerned</u> and an invention disclosed in a publication are contrasted,

問題のある表現



どの語を修飾しているのか明確でない。

改善例



「、」を付与することで、修飾している対象が明確となった。

18

留意事項2：あいまいな接続語を使わない

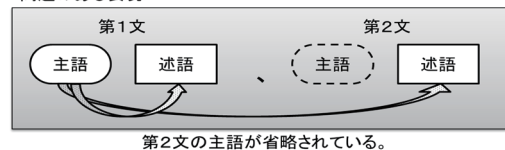
	起案文	AIPN機械翻訳
問題のある表現	しかしながら、請求項1の各 工程の主体が特定されてお らず、かかる工程は「医師が 行う工程」を含み得る。	However, the subject of each process of Claim 1 is not specified, <u>but</u> this process may include "the process which a doctor performs."
改善例	しかしながら、請求項1の各 工程の主体が特定されてい ないため、かかる工程は「医 師が行う工程」を含み得る。	However, <u>since</u> the subject of each process of Claim 1 is not specified, this process may include "the process which a doctor performs."

19

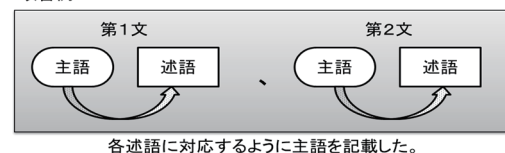
留意事項3: 主語・述語の関係を明確にする

	起案文	AIPN機械翻訳
問題のある表現	用紙がレジスト部に突き当たるようにすることは、当業者であれば適宜なし得る設計的事項であり、格別の作用効果を奏するものとは認められない。	Making it a paper run against a resist part is a matter of design which can be made suitably, if it is a person skilled in the art. <u>Generating an exceptional operation effect is not admitted.</u>
改善例	用紙がレジスト部に突き当たるようにすることは、当業者であれば適宜なし得る設計的事項であり、 <u>かかる構成が</u> 格別の作用効果を奏するものとは認められない。	Making it a paper run against a resist part is a matter of design which can be made suitably, if it is a person skilled in the art. <u>This composition</u> is not accepted to generate an exceptional operation effect.

問題のある表現



改善例

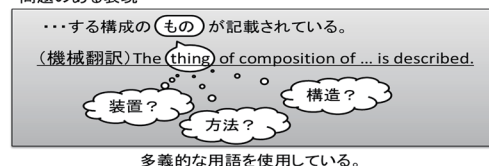


20

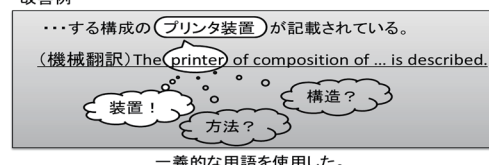
留意事項4: 多義的な表現を避ける

	起案文	AIPN機械翻訳
問題のある表現	引用文献1(特に、段落0084を参照)には、印字部への給紙速度を用紙の大きさに応じて制御する構成の <u>もの</u> が記載されている。	The <u>thing</u> of composition of controlling the paper-feeding speed to the printing unit according to the size of a paper is described in cited document 1 (see the paragraph 0084 especially).
改善例	引用文献1(特に、段落0084を参照)には、印字部への給紙速度を用紙の大きさに応じて制御する構成の <u>プリンタ装置</u> が記載されている。	The <u>printer</u> of composition of controlling the paper-feeding speed to the printing unit according to the size of a paper is described in cited document 1 (see the paragraph 0084 especially).

問題のある表現



改善例



21

日本語として問題はないが機械翻訳結果が悪化する例 (1)

	起案文	AIPN機械翻訳
機械翻訳が悪くなる例	また、本願の請求項3に係る発明の効果は引用文献1及び2の記載から予測し得る程度のものである。	The effect of the invention concerning Claim 3 of an application concerned could be projected based on <u>cited document 1 and the description of 2.</u>
機械翻訳が良好になる例	また、本願の請求項3に係る発明の効果は引用文献1及び引用文献2の記載から予測し得る程度のものである。	The effect of the invention concerning Claim 3 of an application concerned could be projected based on <u>the description of cited document 1 and cited document 2.</u>
	また、本願の請求項3に係る発明の効果は引用文献1－2の記載から予測し得る程度のものである。	The effect of the invention concerning Claim 3 of an application concerned could be projected based on <u>the description of cited document 1-2.</u>

22

日本語として問題はないが機械翻訳結果が悪化する例 (2)

	起案文	AIPN機械翻訳
機械翻訳が悪くなる例	(例えば、第27ページ下から第2行～第29ページ第6行、図2を参照。)	(see <u>the 2nd line</u> - 29th page the 6th line and Fig.2 for example, <u>from under the 27th page.</u>)
	(例えば、第27ページの下から第2行～第29ページ第6行、図2を参照。)	(see <u>the 2nd line</u> - 29th page the 6th line and Fig.2 <u>from under the 27th page.</u>)
機械翻訳が良好になる例	(例えば、第27ページ第24行～第29ページ第6行、図2を参照。)	(see <u>27th page the 24th line</u> - 29th page the 6th line and Fig.2.)

	起案文	AIPN機械翻訳
機械翻訳が悪くなる例	例示するまでもなく周知技術である。	It is not necessary to illustrate and is well-known technology.
機械翻訳が良好になる例	例示するまでもない周知技術である。	It is the well-known technology which does not need to be illustrated.

23

- 「周知技術を示す文献」→ “Document showing a well-known technique”
- 「周知文献」→ “Well-known document”
- 「周知例」→ “Well-known example”
- なお、「周知文献」や「周知例」の場合も “Document showing a well-known technique” と訳出されるように、訳語登録。

1. 審査情報参照システム — その意義と発展

- 海外の特許庁において日本での特許審査経過が適切に参照されることにより、無駄な重複がなく、考慮すべき先行技術等は考慮された、高質かつ速やかな審査が行われることが期待 (企業等のグローバル知財戦略の支援)。
- 日本での審査が迅速化し、FA期間は 29.3月 (2008年度) から 10.4月 (2013年度) に短縮。審査情報を早期に海外に発信する環境が現出。
- 日英機械翻訳機能を有した審査情報参照システムとして、AIPN等が稼働。

2. 拒絶理由通知の日英機械翻訳 — 定型性と構文解析：翻訳例を参照しながら

- 拒絶理由通知には、定型性がある。それを構文解析に一層活用することが機械翻訳の品質向上につながるものと考えられる。

3. 特許庁審査部内での取組 — 拒絶理由通知作成時の留意事項の取りまとめ

- 日本語としてわかりやすく、かつ、(現状のAIPN日英機械翻訳でも) 理解されやすい機械翻訳結果が得られる起案を行うために、2つの「鉄則」と4つの「留意事項」を取りまとめ。

第二部

活動報告 1

特許版・産業日本語の活動報告

河合 弘明

一般財団法人日本特許情報機構 調査研究部長

／特許版・産業日本語委員会 委員

特許版・産業日本語の 活動報告

一般財団法人 日本特許情報機構
河合 弘明



はじめに(目次)

- ◆ 「産業日本語」とは
- ◆ 「特許版・産業日本語」の背景
- ◆ 「特許版・産業日本語」の枠組み
 - 36条ルール化検討G
 - 構造化クレームTF
 - 「特許ライティングマニュアル」改訂TF
- ◆ 公式ウェブサイトのご案内
- ◆ おわりに

「産業日本語」とは

「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」を意味する造語です。

- 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目指します。
- 「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します。
- これにより、以下を実現していきます。
 - (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

3

「特許版・産業日本語」の背景

産業分野における「特許文書」は、技術文書であり、法律文書(権利文書)でもあります。

- 特許文書は、「願書」・「特許請求の範囲」・「明細書」・「図面」・「要約書」から構成されます。
- 特許文書の作成は高度な正確性が求められるため、機械言語処理の支援が有益です。
- 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です。
- しかし、特許文書に特有の課題があります。
 - 一文が長すぎる。文章の係り受けが複雑。
 - 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

「特許版・産業日本語」の枠組み

特許文書に対する日本語改善のアプローチとして、「特許版・産業日本語」の活動を支援しています。

特許版・産業日本語委員会(事務局：Japio)

委員長：橋田浩一（東京大学教授）

委員：23名（平成26年度）

特許全般（企業知財担当、弁理士）・特許情報・特許翻訳・自然言語処理・情報工学ほか各分野の専門家より構成

主な活動：特許版・産業日本語の調査研究、ワークショップ開催、報告書のとりまとめ・公表。また、以下の研究活動を支援。

- 36条ルール化
- 構造化クレームを用いる請求項文ライティング
- グラフ構造形式による特許文書ライティング支援
- 「特許ライティングマニュアル」改訂

5

36条ルール化検討会議

「明細書及び特許請求の範囲の記載要件」（特許法36条関連）について検討をしています。

主査：谷川英和（IRD国際特許事務所所長・弁理士／情報学博士）

委員：7名（主査を含む）

主な活動：拒絶理由通知書（特許法36条関連）の収集・分析、検討会議の開催、報告書のとりまとめ、

「特許版・産業日本語ワークショップ」等での発表

主な論点：

- 拒絶理由通知書の適用条項の分析（審査年別・技術分野別）
- 明確性要件（特許法36条6項2号）に基づく違反類型化
- 産業上利用可能性（特許法29条1項柱書）への応用
- 拒絶理由通知書の自動解析＜今後の課題＞
- 特許文書作成の支援ツール化＜今後の課題＞

6

構造化クレームTF

「構造化クレーム」のツールとしての有用性、機械翻訳への適用可能性等について議論をしています。

リーダー：横井俊夫（NPO法人ISeC理事／東京工科大学名誉教授）

メンバー：8名（リーダーを含む）

主な活動：タスクフォース会合の開催、報告書のとりまとめ

主な論点：

- 「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」がツールとして有用であるか
- 「構造化クレーム」を介することにより、機械翻訳システムへ適用できるか
- 「和文構造化クレーム」の作成
- 「和文構造化クレーム」から「英文構造化クレーム」への翻訳

7

「特許ライティングマニュアル」改訂TF

「特許ライティングマニュアル」（初版：25年度刊行）の改訂に向けた議論をしています。

リーダー：早川貴之（Japio特許情報研究所）

メンバー：8名（リーダー・参加予定者を含む）

主な活動：「特許ライティングマニュアル」（初版：25年度刊行）の配布、有識者ヒアリング、ルール改訂準備

主な論点：

- 翻訳専門家を通じた特許ライティングの課題収集
- 「特許ライティングルール」に係る有識者ヒアリング
- 「仮想特許明細書」の作成・英訳と分析
- 「特許ライティングルール」の分析と再構築
- 「特許ライティングマニュアル」（第二版）への改訂準備
- 「特許ライティングリファレンス」の整備＜今後の課題＞

8

公式ウェブサイトのご紹介

<http://japio-tjp.org/>

特許版・産業日本語

Technical Japanese for Patents

[ホーム](#)
[概要](#)
[公開資料](#)
[リンク](#)
[ログイン](#)

[サイトマップ](#)

本サイトでは、特許情報のための産業日本語に関するホットな話題を提供いたします。

◆トピックス

- ・2014-4-25 平成25年度 特許版・産業日本語委員会 報告書を掲載いたしました。
- ・2014-3-24 特許ライティングマニュアル(初版)第2刷を掲載いたしました。
- ・2013-8-29 特許ライティングマニュアル(初版)を掲載いたしました。
- ・2013-4-26 平成24年度 特許版・産業日本語委員会 報告書を掲載いたしました。
- ・2013-1-30 第4回産業日本語研究会・シンポジウム開催の案内がアナウンスされました。皆様のご参加をお待ちしています。
- ・2012-6-1 特許・情報フェア&コンファレンス2012の開催(11月)がアナウンスされました。
- ・2012-6-1 特許版・産業日本語のサイトをリニューアルいたしました。

平成26年度版も掲載予定

初版(PDF)を入手可能

9

おわりに

ご静聴ありがとうございました。

グローバル化が進むなか、正確かつ円滑な情報発信力が求められ、「産業日本語」の役割は益々重要になるものと考えます。

日本特許情報機構は、皆様のご支援・ご協力を賜りながら、今後も引き続き、「特許版・産業日本語」活動への支援を続けてまいります。

＜本件問合せ先＞
 一般財団法人 日本特許情報機構
 調査研究部長 河合 弘明
 TEL: 03-3615-5513
 E-mail: hiroaki_kawai@japio.or.jp



一般財団法人
日本特許情報機構
 Japan Patent Information Organization

＜Website＞ <http://www.japio.or.jp/>

10

活動報告 2
日本人のための日本語マニュアル

横井 俊夫

東京工科大学 名誉教授

／一般財団法人日本特許情報機構特許情報研究所 顧問

日本人のための日本語マニュアル

2015 年 2 月 24 日

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」が必要である

日本人は、日本語を使えるが日本語の仕組みを十分に知っているわけではない。日本語文章の分析能力を養成するための「日本人のための日本語マニュアル」が必要である。本年 9 月上梓を目標に「日本語マニュアルの会」がマニュアル制作を進めている。

日本語マニュアルの会

(とりあえず年齢順に)

横井俊夫 (東京工科大学名誉教授、Japio 特許情報研究所顧問)

石崎 俊 (慶応大学名誉教授)

佐野 洋 (東京外国語大学教授)

石黒 圭 (一橋大学教授)

猪野真理枝 (翻訳家、語学教材作家)

烏日哲 (ウリジャ) (一橋大学国際教育センター非常勤講師)

「日本人のための日本語マニュアル」の構成

『日本人のための日本語マニュアル

- 言葉の仕組みを学び、外国語との対照を通じて日本語スキルを磨く - 』

はじめに マニュアルの目的、マニュアルの構成と利用の仕方

1 章 文書・文章ライティングのモデルプロセスを学ぶ

2 章 情報を表し情報を伝える言葉の仕組みを学ぶ

- 日本語と外国語とを照らし合わせ -

3 章 「表す日本語」へと言い換え、「伝える日本語」へと言い換える

4 章 「訳せる日本語」へと言い換える

5 章 コンピュータの支援機能を活用する - 文章校正ソフトと機械翻訳ソフト -

おわりに さらなるスキルアップのために - 参考資料 -

それぞれがそれぞれのライティングマニュアルを

産業日本語を定着させるためには、それぞれの機関・組織・業界・分野がそれぞれの文書目的に沿ったライティングマニュアルを整備しなければならない。本マニュアルが各所での整備作業の一助となることを願うと共に、そうなるための効果的な仕組を用意する。

産業日本語活動のための仕組作りに向けて

(1) ライティングルール（多くは、言い換えルール）記述の共通フォーマット

- 人が理解し易いルール記述から、コンピュータ用の形式的ルール記述まで
- マニュアルのルール記述から、文書処理ソフトのインタフェース記述まで
- スキーマ的なルール記述から、事例的なルール記述まで

(2) 「日本人のための日本語マニュアル」のライティングルールの公開

- 日本語マニュアル制作のソーシャルネットワーク
- ライティングルールの共有知化
- 老若協創のモデル作り

(3) 産業日本語ガイドライン策定の提案

- それぞれのライティングマニュアル整備のためのガイドライン
- ライティング支援ツール開発のためのガイドライン
- 統合的な文書作成・再利用環境整備のためのガイドライン

なぜ、日本人に日本語のマニュアルが必要なのか

- 「使える」と「知っている」は別、「知って使える」のが一番
- 身体能力の基本は暗黙知、使えるが知っているわけではない
 - 身体能力の高度化、障害の克服、コンピュータ化に対応するには、身体能力の仕組を知らねばならない
 - 言語能力（特に、会話能力）も身体能力のひとつ
- 言語教育における課題
 - 国語教育は、重要な日本語文化・日本文化の教育が主眼、日本語のツールスキルとしての教育が希薄

- 国語教育と英語教育との間に、言葉のスキル教育としての連携が希薄

なぜ、今、日本語、そして、日本語マニュアルなのか

(1) 潜在的ニーズの高まり

- 産業活動のグローバル化 →世界へ日本語を開く
- 産業活動の構造変化 →ジャーゴン化を排し、分野間に日本語を開く
- 多様なメディア連携の環境 →メディア間に日本語を開く
- 産業活動の共通インフラとなる ICT 環境 →コンピュータに日本語を開く

(2) 潜在的ニーズを顕在的ニーズへ

- 日本語か英語かの二者択一を排し、複合的な社内言語システムの整備へ
- 日本語で考え、英語で伝え、さらに、中国語等々で伝える
- 実感できるメリットを提示する

(3) 日本語マニュアル整備への新たな基盤

- 言語学における進展、日本語教育における進展、言語処理における進展
- 諸成果を踏まえることによって、日本語トレーナーのスキル確立

産業文書・ビジネス文書

(1) 顧客サービス文書

ユーザマニュアル、社外 Web

(2) 業務文書

業務連絡文書、業務報告書、業務提案書、議事録、社内 Web、社内メール、社内 SNS

(3) 技術文書

開発文書、障害対応文書、マニュアル（業務、操作、運用）、成果文書（技術報告書、発明提案書、学術論文）

(4) 法的文書

開示知財文書（特許、実用新案、商標、意匠）、守秘知財文書、契約書、規約・定款、コンプライアンス関連文書

文書と文章

文書 (document) の構造

それぞれの目的に沿った文書記載項目の構造、そして、記載項目を埋めて項目の内容となるのが文章

文章 (text) の構造

文書の記載項目の規定によって、一文、一語句の場合もあるが、基本の構造は、

文章 (text)

段 (paragraph)

文 (sentence)

語 (word)

文章特性に則した日本語ライティング

(1) 印象深さを重視する

読み手が共通に持つ知識や推論に大きく委ねることを前提に印象深くテンポ良く伝える日本語。ユーザマニュアルや報道記事などの顧客サービス系の文章特性。

(2) 正確さを重視する

読み手の知識や推論に依存する部分を確実なものに限定し、誤解が生じないように正確に効率よく伝える日本語。学術文書、技術文書、業務文書等々の文章特性。

(3) 厳格さを重視する

主旨に反する読み方ができないよう解釈を限定し厳格に伝える日本語。法的文書の文章特性。

(4) 情感の豊かさを重視する

読み手の感興を刺激し、解釈を読み手の創造性に委ね、情感豊かに伝える日本語。文芸作品の文章特性。

(5) 対話性を重視する

読み手からの即時応答を前提とし書き言葉の記録性と話し言葉の即応性を併せ持つ日本語。電子メールや SNS メッセージなどの文章特性。

文書・文章ライティングのモデルプロセス (図 1 を参照)

「試みる日本語」

思考のツールとしての機能に重きを置いた日本語である。思考のダイナミズムを加速する柔軟さ、時には、いい加減さが求められる。

「表わす日本語」

思考の精密化というプロセスを伴って、表現対象に注視し、表現対象を適切に確定し、表現対象を的確に表現する日本語である。

「伝える日本語」

読み手（読者）の知識や読み手の立場を考慮し、読み手が効率良く間違いなく読み取れるようにする日本語である。

「訳せる日本語」

外国語へ直訳できる日本語である。文を超える構成成分に関しては線形翻訳できるとし、連語以下の非線形成分は対訳辞書に登録されているとし、焦点を文レベルに絞る。

「機械が訳せる日本語」

容易な後編集を伴って正しく機械翻訳できる日本語である。短文化と適切な辞書登録が必要である。

「機械が探せる日本語」

「機械が約せる日本語」

情報を表わし伝える言葉の仕組

言葉の仕組を、言語学の解説としてではなく、文章実務のための文章分析スキルとなるように説く

それぞれの言語に共通となる仕組

「絵にも描けない美しさ」、この美しさを絵に描くことは出来ないが、
「言葉に言い表せない驚き」、この驚きは言葉で言い表せている。何故か？

-
- ① 記号性・分節性・概念性・範疇性・高階性
 - ② 線状性（音声言語が優位）
→情報（四次元）と言語表現（一次元）間の相互マッピング
 - ③ 表現の経済性（読み手（聞き手）の知識を前提とした表現）
 - ④ 表現メディア間の連携性（多様な表現メディアを結びつける機能）
 - ⑤ 柔軟な規則性（プロトタイプと周辺事例）

それぞれの言語に特有の仕組

「今日書店で先生が本を買う。」
“The teacher buys a book in the bookstore today.”
「老師今天在書店買書。」
これら3文の違いは？

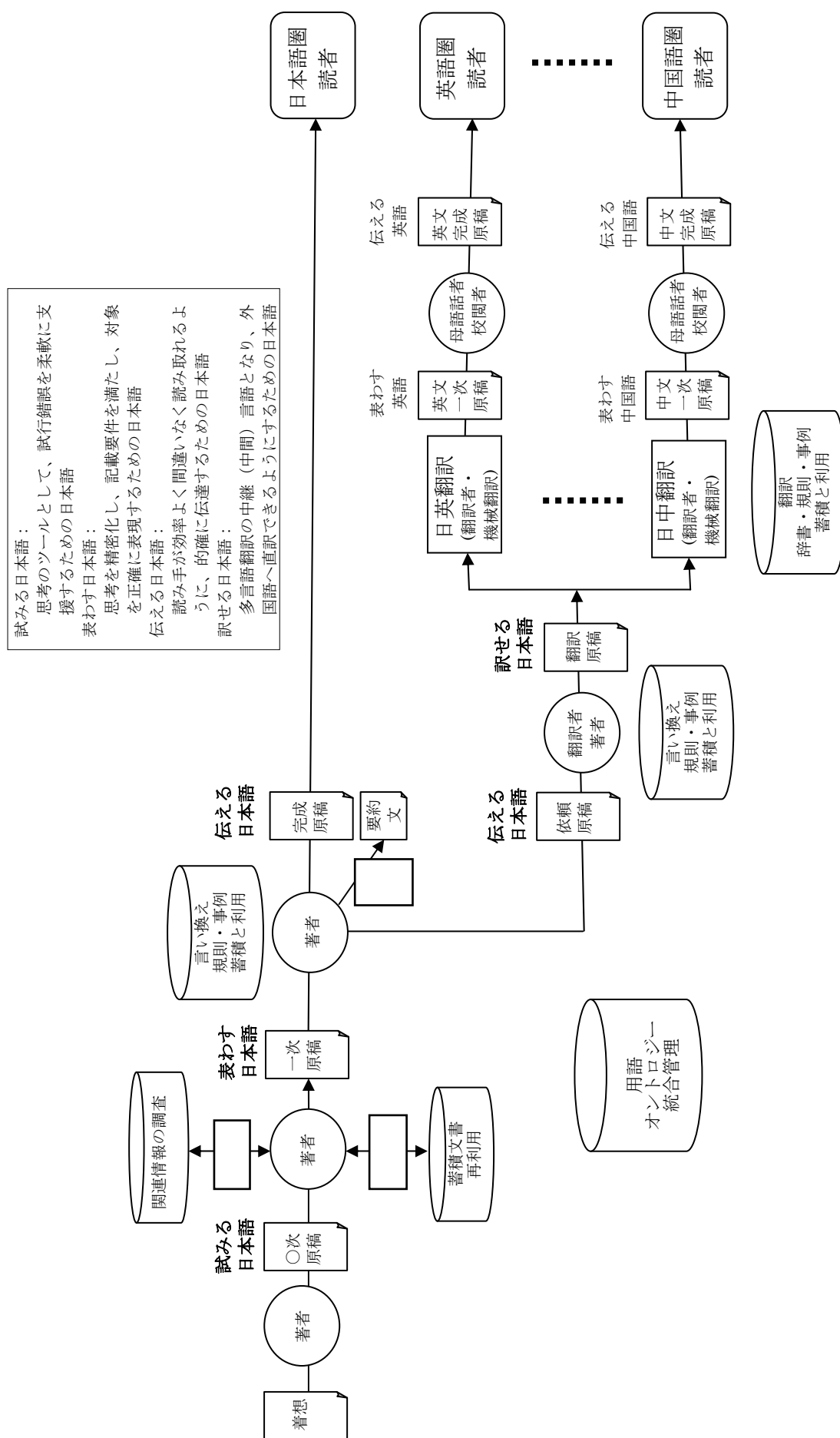
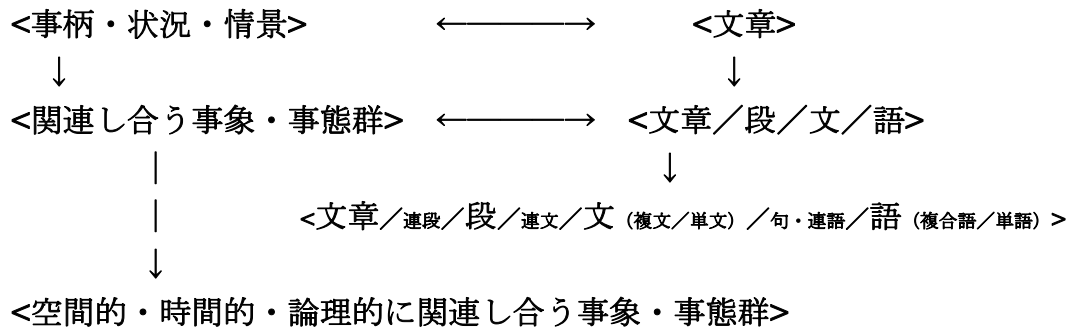


図 1. 文書・文章ライティングのモデルプロセス

それぞれの言語に共通となる仕組

情報：四次元に展開するひとつの事柄（物事の様子・事情・内容）

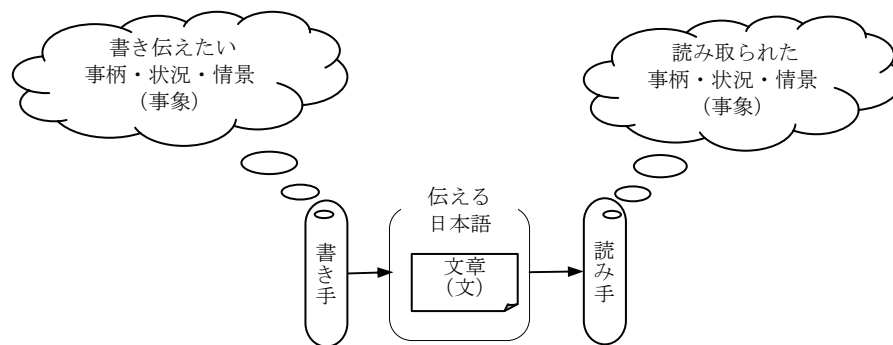
表現：一次元（書き言葉は二次元的表現要素を含む）に展開する文章



まずは、

<事象・事態> ←→ <単文>

明晰な情報伝達



<書き伝えたい事柄・状況・情景（事象）>=<読み取られた事柄・状況・情景（事象）>
しかも、効率良く読み取れること

事象を構成するもの

事象そのもの（事象の客観的側面）

<コト>+<係わるモノ>+<状況を構成するモノ>+<アスペクト性>+<テンス性>

書き手の捉え方（事象の主観的側面）

<モダリティ性>+<取り立て>+<強調・焦点>+<否定>+<読み手の位置づけ方>

読み手への伝え方（事象の談話的側面）

<既出情報・旧情報>+<未出情報・新情報>

それぞれの言語に特有の仕組

まずは、事象の言語表現に関して

事象そのものの表現法

通常、日本語は SOV、英語や中国語は SVO といわれていること

書き手の捉え方の表現法

通常、態度表現、待遇表現等々といわれていること

読み手への伝え方の表現法

通常、情報構造といわれていること

事象そのものの表現法

日本語

<何時><何処>で<誰>が<何>を<どう>する／<どう>した／<どう>している。

英語

<誰> <どう>／<どう>ed／will <どう>／be <どう>ing <何> in <何処> <何時>.

中国語

<誰><何時>在<何処><どう>／<どう>了<何>。

読み手への伝え方の表現法

<既出情報・旧情報>を前方に、<未出情報・新情報>を後方には言語共通

日本語

<主題>：<既出情報・旧情報>←場の設定

<題述>：<未出情報・新情報>←場における事象

英語

<主語>：<既出情報・旧情報>←定冠詞

<述部>：<未出情報・新情報>←不定冠詞

主語卓立言語と題目卓立言語

[+題目卓立 ー主語卓立] 言語：中国語、ラフ語、リス語

[ー題目卓立 +主語卓立] 言語：印欧語族、ニジェール・コンゴ語族、セム語族

[+題目卓立 +主語卓立] 言語：日本語、韓国語

[ー題目卓立 ー主語卓立] 言語：タガログ語、イロカノ語

資料 1 : 共通に参照する資料

日本語マニュアルの会のメンバーがマニュアル制作に際して、共通に参照する資料であり、より高いスキルアップを目指すマニュアルユーザーへの推奨資料である。

(1) 日本語マニュアルの会のメンバーの執筆文献

- 横井俊夫: 翻訳できる日本語、Japio Year Book 2010、一般財団法人日本特許情報機構、pp.154-159 (2010 年 11 月)
- 横井俊夫: 産業日本語における日本語の諸相 - 文書の作成・利用を高度化する新たな文書工学の仕組 - 、Japio Year Book 2011、一般財団法人日本特許情報機構、pp.286-295 (2011 年 11 月)
- 横井俊夫: 言葉をデザインする - デザインされた産業日本語 - 、Japio Year Book 2012、一般財団法人日本特許情報機構、pp.298-308 (2012 年 11 月)
- 横井俊夫: 構造化言語 - 知を構造化する言葉の構造化技術 - 、ISEC10 周年記念シンポジウム予稿集、pp.25-61 (2013 年 6 月)
- 横井俊夫: 産業日本語のガイドライン策定に向けて、Japio Year Book 2013、一般財団法人日本特許情報機構、pp.302-307 (2013 年 11 月)
- 横井俊夫: 日本人のための日本語マニュアル、Japio Year Book 2014、一般財団法人日本特許情報機構、(2014 年 11 月)

- 石黒 圭: よくわかる文章表現の技術 I 表現・表記編【新版】、明治書院 (2012 年 11 月)
- 石黒 圭: よくわかる文章表現の技術 II 文章構成編【新版】、明治書院 (2009 年 11 月)
- 石黒 圭: よくわかる文章表現の技術 III 文法編、明治書院 (2014 年 3 月)
- 石黒 圭: よくわかる文章表現の技術 IV 発想編、明治書院 (2006 年 9 月)
- 石黒 圭: よくわかる文章表現の技術 V 文体編、明治書院 (2007 年 10 月)
- 石黒 圭: 文章は接続詞で決まる、光文社新書、光文社 (2008 年 9 月)
- 石黒 圭: スッキリ伝わるビジネス文書「一読必解」21 のルール、光文社 (2010 年 2 月)
- 石黒 圭: 「予測」で読解に強くなる!、ちくまプリマー新書、筑摩書房 (2010 年 7 月)
- 石黒 圭: この 1 冊できちんと書ける! 論文・レポートの基本、日本実業出版社 (2012 年 2 月)
- 石黒 圭: 正確に伝わる! わかりやすい文書の書き方、日本経済新聞出版社 (2012 年 12 月)
- 石黒 圭: 「うまい!」と言わせる文章の裏ワザ、河出書房新社 (2014 年 5 月)

- 猪野真理枝、佐野 洋著、馬場 彰監修：英作文なんかこわくない - 日本語の発想でマスターする英文ライティング、東京外国語大学出版会（2011 年 4 月）
- 猪野真理枝、佐野 洋著、馬場 彰監修：英作文なんかこわくない II 連体修飾編 - 日本語の発想でマスターする英文ライティング、東京外国語大学出版会（2014 年 4 月）
- 猪野真理枝、佐野 洋著、馬場 彰監修：英作文なんかこわくない III 連用修飾編 - 日本語の発想でマスターする英文ライティング、東京外国語大学出版会（2014 年 予定）
- 猪野真理枝、佐野 洋著、馬場 彰監修：英作文なんかこわくない IV パラグラフ編 - 日本語の発想でマスターする英文ライティング、東京外国語大学出版会（2015 年 予定）

- 安西、石崎他：認知科学ハンドブック、共立出版（1995 年 10 月）
- 石崎：自然言語処理、昭晃堂（1995 年 5 月）
- 天野、石崎他：自然言語処理、オーム社（2007 年 10 月）
- 言語処理学会編(石崎編集委員長)：言語処理学事典、共立出版（2009 年 12 月）
- 言語処理学会編(石崎編集委員長)：デジタル言語処理学事典、共立出版（2010 年 6 月）

（2）代表的なライティング指南書

- 木下是雄：理科系の作文技術、中公新書、中央公論新社（1981 年 1 月）
- 本田勝一：日本語の作文技術、朝日文庫、朝日新聞出版（1982 年 2 月）
- 阿部圭一：明文術－伝わる日本語の書き方、NTT 出版（2006 年 8 月）
- メアリ・K・マカスキル、片岡秀樹訳・解説：NASA に学ぶ英語論文・レポートの書き方－NASA SP-7084 テクニカルライティング、共立出版（2012 年 2 月）

（3）日本語に関する事典・辞典

- 北原保雄監修：岩波 日本語使い方考え方辞典、岩波書店（2003 年 5 月）
- 森田良行：助詞・助動詞の辞典、東京堂出版（2007 年 9 月）
- 森田良行：動詞・形容詞・副詞の事典、東京堂出版（2008 年 10 月）
- 言語処理学会編：言語処理学事典、共立出版（2009 年 12 月）
- 辻 幸夫編：新編 認知言語学キーワード事典、研究社（2013 年 10 月）
- 日本語文法学会編：日本語文法事典、大修館書店（2014 年 7 月）
- 佐藤武義、前田富祺編集代表：日本語大事典、朝倉書店（2014 年 11 月）

（4）外国人のための日本語教育の教師用ハンドブック

- 松岡 弘監修、庵 功雄、高梨信乃、中西久美子、山田敏弘著：初級を教える人のための日本語文法ハンドブック、スリーエーネットワーク（2000 年 5 月）
- 白川博之監修、庵 功雄、高梨信乃、中西久美子、山田敏弘著：中上級を教える人のた

めの日本語文法ハンドブック、スリーエーネットワーク（2001 年 10 月）

（5）日本語と外国語（特に、英語）との対照

- 角田太作：世界の言語と日本語 改訂版 - 言語類型論から見た日本語、くろしお出版（2009 年 5 月）
- 吉村公宏：英語世界の表現スタイル 「捉え方」の視点から、青灯社（2011 年 5 月）
- 菅井三実：英語を通して学ぶ日本語のツボ、開拓社言語・文化選書、開拓社（2012 年 10 月）
- 岡 智之：場所の言語学、ひつじ研究叢書＜言語編＞、ひつじ書房（2013 年 2 月）
- 影山太郎編：日英対照 動詞の意味と構文、大修館書店（2001 年 3 月）
- 影山太郎編：日英対照 形容詞・副詞の意味と構文、大修館書店（2009 年 4 月）
- 影山太郎編：日英対照 名詞の意味と構文、大修館書店（2011 年 11 月）

（6）中国語文法への入門書

- 瀬戸口律子：完全マスター 中国語の文法、語研（2003 年 6 月）
- 相原 茂：はじめての中国語「超」入門、ソフトバンク新書、ソフトバンク・クリエイティブ（2007 年 4 月）

（7）日本語の表記に関する規範

- 小学館辞典編集部編：句読点、記号・符号活用辞典、小学館（2007 年 9 月）
- 日本エディタースクール編：日本語表記ルールブック 第 2 版、日本エディタースクール出版部（2012 年 3 月）
- 三省堂編集所編：新しい国語表記ハンドブック第 6 版、三省堂（2013 年 9 月）

資料 2：日本語文章力に関する出版物

【単行本】

2013 年

日垣 隆：文章力飛躍最終兵器 第 1 弾 基礎編(2013/8/17)
日垣 隆：文章力飛躍最終兵器 第 2 弾 実践編(2013/8/17)
日垣 隆：文章力飛躍最終兵器 第 3 弾 仕上げ編(2013/8/17)
米村貴裕：天使の法則：役立つ文章の世界へ(2013/8/9)
坪田知己：共感文章術シリーズ 1 読まない人に読ませる共感文章術 ネット時代の文章術・基礎編(2013/8/6)
入部明子：パワーライティング入門：説得力のある文章を書く技術(2013/8/5)
木山泰嗣：頭が 10 倍よく見える文章の書き方：弁護士が書いた「伝わる」文章術 (知的生きかた文庫)(2013/7/22)
福岡隆史："ふくしま式 200 字メソッド"で「書く力」は驚くほど伸びる!(2013/7/19)
高橋俊一：決定版! すっきり書ける文章のコツ 80(2013/6/28)
小林洋介：デキる大人の文章力教室(2013/6/27)
木暮太一：伝え方の教科書(2013/6/25)
中野 巧：6 分間文章術一想いを伝える教科書(2013/6/21)
イケダハヤト：武器としての書く技術(2013/6/19)
小野泰央：創造するための文章(2013/6)
中谷彰宏：「ひと言」力。サッと書いて、グッとくる 99 の方法(2013/5/28)
前田安正：きっちり! 恥ずかしいくない! 文章が書ける(2013/5/23)
高橋廣敏：7 日で身につく 正しい文章の書き方(2013/5/21)
出口 汪：あなたも突然、上手に書けるようになる 7 日間「大人の論理エンジン」で本物の文章力が身につく(経済界新書)(2013/5/10)
安藤智子：電子書籍を出す人は知っておきたい文章術 どんな原稿も必ず格上げする「秘伝」公開(2013/5/9)
新星出版社編集部：図解まるわかり 文章術の基本―ビジネス力をグンと上げる(2013/5)
平野友朗：ちょっとした工夫で仕事がぐんぐんはかどるビジネスメール術―仕事ができる人がやっている 43 のルール(2013/4/19)
山口拓朗：文章が変わると人生が変わる! 文章力アップ 33 の方法(2013/4/17)
結城 浩：数学文章作法 基礎編(ちくま学芸文庫)(2013/4/11)
杉山美奈子：ビジネスメールの作法と新常識 会社では教えてくれない気くばりメール術 (アスキー新書)(2013/4/10)
高橋源一郎：はくらの文章教室(2013/4/5)
村岡貴子、因 京子、仁科 喜久子：論文作成のための文章力向上プログラム「アカデミック・ライティング」の核心をつかむ(2013/4/5)
神垣あゆみ：言いたいことが 5 秒で伝わるメール術(2013/4)
扇田麻里子：読まれるための文章読本(2013/3/29)
佐渡島紗織、太田 裕子：文章チュータリングの理念と実践―早稲田大学ライティング・センターでの取り組み(2013/3/28)
赤羽博之：書くスキル UP すぐできる! 伝わる文章の書き方 確実に文章力がつく! 7 つのステップ(2013/3/20)
外岡秀俊：「伝わる文章」が書ける作文の技術(2013/3/12)
銅畑信子、坂東実子：大学生のための文章表現&口頭発表練習帳(2013/3/8)
園部俊晴：医療・福祉で役立つ「効果的な文章の書き方」入門講座(医療・福祉で働く人のスキルアップシリーズ)(2013/3/1)
前田めぐる：ソーシャルメディアで伝わる文章術(2013/3)
山口拓朗：ダメな文章を達人の文章にする 31 の方法 なぜあなたの文章はわかりにくいのか? 文章の書き方が分かる本 (縦組版) (2013/2/18)
吉岡友治：いい文章には型がある (PHP 新書)(2013/2/17)
中川路幸紀：ビジネスメール文章術(2013/2/16)
竹内政明：「編集手帳」の文章術(文春新書)(2013/1/20)

2012 年

福澤一吉、にしかわ たく：論理的に読む技術 文章の中身を理解する"読解力"強化の必須スキル (サイエンス・アイ新書)(2012/12/18)
岡本 真：ウェブでの<伝わる>文章の書き方(講談社現代新書)(2012/12/18)
石黒 圭：正確に伝わる! わかりやすい文章の書き方(2012/12/13)
うすたか、江(平総理)：ハッター 文章術(ソフトバンク文庫)(2012/12/12)
出口 汪：出口 汪の論理的に書く技術(ソフトバンク文庫)(2012/11/22)
倉島保美：論理が伝わる 世界標準の「書く技術 (ブルーバックス)(2012/11/21)
外岡秀俊：「伝わる文章」が書ける作文の技術 名文記者が教える 65 のコツ(2012/10/19)
久恒啓一：図で考えれば文章がうまくなる(2012/10/5)
プレジデントムック、書き方の基本 まねて書ける! OK 文例集―ビジネス文書からメール、ツイッター、フェイスブックまで(2012/10/1)
山口拓朗：ダメな文章を達人の文章にする 31 の方法 なぜあなたの文章はわかりにくいのか? 文章の書き方が分かる本 (横組版) (2012/10/1)
栗倉敏貴：介護職の文章作成術(2012/9/10)
阿部敏久：シンプルに書く! 伝わる文章術(2012/9/6)
校條 剛：朝 5 分! 読むだけで文章力がグッと上がる本 (ナガオカ文庫)(2012/7/20)
山崎政志：文章力の「基本」が身につく本：伝わる文章が書ける 76 の簡単テクニック(2012/6/26)
石崎秀穂：あなたの文章が(みるみる) わかりやすくなる本(2012/5/8)
横沢紫苑：SNS の超プロが教える ソーシャルメディア文章術(2012/4/4)
日野るな、戸田美紀、小谷俊介、カトウナオコ：世界一やさしい「人脈」と「収入」をザクザク生み出すブログ文章術(2012/3/26)
メアリ・K・マカスキル著、片岡英樹訳・解説：NASA に学ぶ英語論文・レポートの書き方 - NASA SP7084 テクニカルライティング - (2012/2/25)
木暮太一：誰にでも伝わる 文章力のつくり方(2012/2/24)
石黒 圭：この 1 冊できちんと書ける! 論文・レポートの基本(2012/2/23)
山本ゆうじ：IT 時代の実務日本語スタイルブック(2012/2/16)
木山泰嗣：センスのよい法律文章の書き方(2012/2/8)
西田みどり：「型」で書く文章論―誰でも書けるレポート講座(2012/2)
古賀史穂：20 歳の自分に受けさせたい文章講義(星海社新書)(2012/1/26)
小笠原信之：伝わる! 文章力が豊かになる本(2012/1/7)

2011 年

高橋俊一：すっきり! わかりやすい! 文章が書ける(2011/12/20)
平野友朗：お客様から選ばれるウェブ文章術(2011/11/19)
苦米地英人：人を動かす【超】書き方トレーニング 劇的な成果が手に入る 驚異の作文術(2011/11/16)
堀内伸浩：あたりまえだけどなかなか書けない 文章のルール(アスカビジネス)(2011/10/18)
野内良三：伝える! 作文の練習問題(2011/9/30)
近藤勝重：書くことが思いつかない人のための文章教室 (幻冬舎新書)(2011/9/29)
上阪 徹：文章は「書く前」に 8 割決まる(2011/9/21)

堀井憲一郎：いますぐ書け、の文章法(ちくま新書)(2011/9/5)
木山泰嗣：もっと論理的な文章を書く(2011/8/27)
西村充己：「ポイント図解」論理的な文章の書き方が面白いほど身につく本「わかりやすい文章」を書くための基本ポイント 35(2011/8/26)
西川真理子：図解 栄養士・管理栄養士をめざす人の文章術ハンドブック：ノート・レポート、手紙・メールから、履歴書・エントリーシート、卒論まで(2011/8/14)
吉村公宏：英語世界の表現スタイル―「捉え方」の視点から(2011/5/30)
藤原智美：文は一行目から書かなくていい―検索、コピー時代の文章術(2011/5/27)
小田順子：誰も教えてくれなかった公務員の文章・メール術(2011/5/19)
樋口裕一：文章力の鍛え方 (中経の文庫)(2011/4/27)
高橋フミアキ：150 字からはじめる「うまい」と言われる文章の書き方(2011/4/21)
高橋フミアキ：伝わる文章の書き方(2011/4/15)
猪野真理枝、佐野 洋：英作文なんかこわくない 日本語の発想でマスターする英文ライティング(2011/4/15)
山崎康司：入門 考える技術・書く技術(2011/4/8)
一般財団法人テクニカルコミュニケーター協会：日本語スタイルガイド 第 2 版(2011/4/1)
黒本登志夫：知的文章とプレゼンテーション―日本語の場合、英語の場合(中公新書)(2011/4)
酒井聡樹：100 ページの文章術―わかりやすい文章の書き方のすべてがここに―(2011/3/10)
村田喜代子：縦横無尽の文章レッスン(2011/3/4)
田中 豊：法律文書作成の基本 Legal Reasoning and Legal Writing(2011/2/20)
上原広嗣、石原知樹、寄藤平平：文章には「道」がある 10 代 20 代のための日本語を読む技術 Part1(2011/2/9)
黒田龍之助：大学生からの文章表現 無難で退屈な日本語から卒業する(ちくま新書)(2011/2/9)
杉山博亮：論文答案作成教室―法律的文章を書くコツ(2011/2)
小笠原信之：伝わる! 文章力が身につく本(2011/1/29)
飯間浩明：伝わる文章の書き方教室 書き換えトレーニング 10 講(ちくまプリマー新書)(2011/1/7)

2010 年

齋藤 孝：誰も教えてくれない人を動かす文章術(講談社現代新書)(2010/12/17)
上阪 徹：書いて生きていく プロ文章論(2010/11/26)
堀内伸浩：ビジネス文章 5 ステップ上達法―今の自分の文章がわかる診断テスト付き(2010/9)
阿部敏久：明快に書く! 心をつかむ文章力(知的生きかた文庫)(2010/9/21)
白井由紀：仕事の文章は 3 行でまとめなさい(2010/8/25)
成川豊彦：【決定版】成川式 文章の書き方(PHP ビジネス新書)(2010/8/19)
磯崎陽輔：分かりやすい公文書の書き方 改訂版(2010/8/6)
有田憲史：「売る」文章 51 の技―説得力あるキャッチコピーとロングコピーの作り方(2010/8/3)
阿部敏久：文章力の基本 100 題(2010/6/19)
スティーブン・D・スターク、小倉京子：訴訟に勝つ実践的文章術(2010/6/18)
松枝史明：「売れる・読ませる」文章が書ける 実践的ライター入門(2010/6/4)
日経 PC21：サクサク作成! エクセル文書ワザ 99(日経ビジネス人文庫)(2010/6/2)
藤田英時：メール文章力の基本 大切だけど、だれも教えてくれない 77 のルール(2010/5/27)
野内良三：日本語作文術 (中公新書) (2010/5/25)
高橋恵治：あたりまえだけどなかなかできない 文章のルール (アスカビジネス)(2010/5/17)
芦永奈雄：コミュニケーション力を高める文章の技術 (フォレスト 2545 新書)(2010/5/7)
橋本淳司：「簡条書き」を使ってまとまった量でもラクラク書ける文章術(2010/4/16)
田邊善治：電力社員の文章講座(2010/3)
内藤直人：心理学者が教える 思いどおり人に動かすブラック文章術(2010/2/23)
近藤勝重：早大院生と考えた文章がうまくなる 13 の秘訣(2010/1/25)
松下健次郎、山本高樹、白根ゆたんぼ：これは「効く!」Web 文章作成&編集術逆引きハンドブック(2010/1/21)

2009 年～

石黒 圭：よくわかる文章表現の技術 I～V(2009/11～2005/10)
原田豊太郎：理系のための英語「キー構文」46(2009/9/20)
阿部敏久：文章力の基本(2009/8/1)
三浦順治：英語流の説得力をもつ 日本語文章の書き方(2009/6/11)
中井浩一：正しく読み、深く考える 日本語論理トレーニング(講談社現代新書)(2009/2/19)
阿部圭一：明文術 伝わる日本語の書きかた(2006/7/25)
黒屋華子：ロジカル・ライティング 論理的にわかりやすく書くスキル(2006/4/6)
柳瀬和明：「日本語から考える英語表現」の技術(2005/3/20)
高橋昭男：日本語テクニカルライティング(2004/5/26)
藤沢晃治：「分かりやすい文章」の技術―読み手を説得する 18 のテクニック (ブルーバックス)(2004/5/21)
野内良三：実践ロジカル・シンキング入門 - 日本語論理トレーニング (2003/2/1)
原田豊太郎：理系のための英語論文執筆ガイド ネイティブとの発想のズレはどこに?(2002/3/20)
片岡英樹：技術英文作成に必須! テクニカル・ライティング 50 のルール (2001/5/1)
木下是雄：理科系の作文技術 (中公新書) (1981/1)

【雑誌】

PRESIDENT (プレジデント)
2013 年 8 月 19 日別冊：うまい文章の全技術
2013 年 4 月 1 月号：資料の作り方
2012 年 7 月 16 日号：1 億稼ぐ人の話し方
2012 年 4 月 16 日号：一流の思考法 落第思考法
2012 年 3 月 19 日号：「すべらない」書き方
週刊ダイヤモンド
2013 年 8 月 24 日号：伝える技術
2012 年 4 月 7 日号：「話し方」入門
2012 年 3 月 3 日号：身につく! 英語&中国語
2012 年 1 月 14 日号：ビジュアル活用仕事術
週刊東洋経済
2012 年 6 月 2 日号：脱 TOEIC の英語術
ニューズウィーク日本版
2013 年 7 月 23 日号：TOEFL 時代を制する英語術

第三部

言語処理分野の最前線 1

産業日本語としての医療カルテ文章

荒牧 英治

京都大学デザイン学ユニット 特定准教授

／独立行政法人科学技術振興機構 さきがけ研究員

産業日本語としての 医療カルテ文章

荒牧英治 京都大学



自己紹介

- 学部: 京大 総合人間学部
- 修士: // 言語メディア研
- 博士: 東大 情報理工

自然言語処理



- 東大 医学部附属病院
 - 助教 (2006-2008)
- 東大 知の構造化センター
 - 講師 (2008-2013) 研究主導者 (PI)
- 京大 デザイン学ユニット
 - 准教授 (2013-) 研究主導者 (PI)

医療分野での
言語処理研究に
従事



医療における 自然言語処理とは



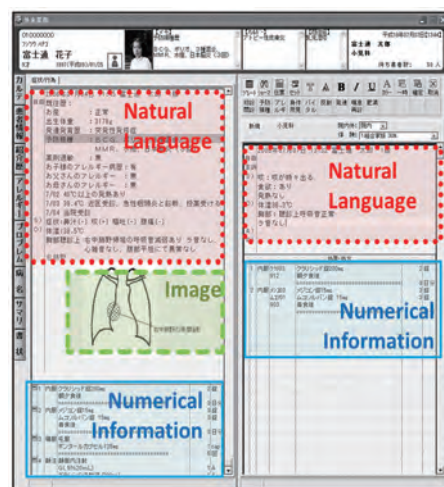
近所の医院 →

電子化 ≠ 標準化

電子カルテは **Natural Language**
(自然言語) を含んでいる



病院ごとに、各診療科ごとに、
(場合によっては) 医師ごとに
表現が異なる



様々な表現にバリエーションを扱うために
言語処理 (Natural Language
Processing; NLP) が必要

Marchesani Syndrome

マルケサーニ症候群
マルケサニ一症候群
マルケサニ症候群
マルケザーニ症候群
マルケザニ一症候群
マルケザニ症候群

ICD 10 =
Q871

Transliteration (翻字)

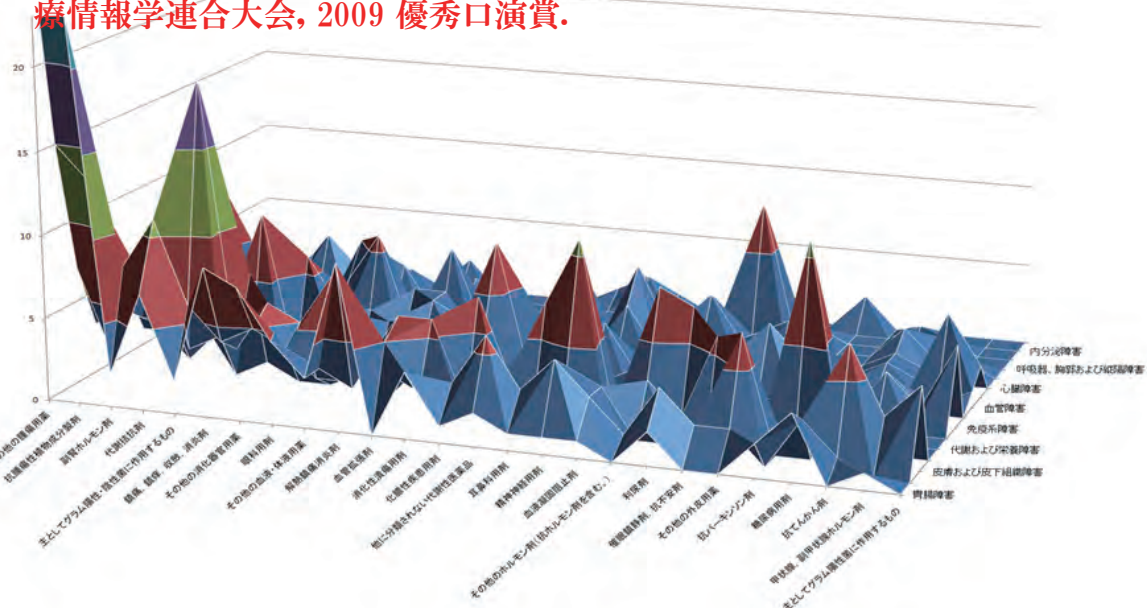
悪心増悪
胃のむかつき
悪心 (sickness)
食後悪心
嘔気 (vomiting)
吐き気
悪阻 (nausea)

ICD 10 = R16
悪心及び嘔吐
(sickness & vomiting)

Paraphrase (言い換え)

東大病院での副作用情報抽出

外池昌嗣, 大熊智子, 荒牧英治, 三浦康秀, 増市博, 大江和彦: 自然言語表現の現病歴情報を時系列表形式で表示するシステムとその評価, 第29回医療情報学連合大会, 2009 優秀口演賞.



医療言語処理システムの精度

工場に勤めている64歳の男性。2025年月8月2日(来院5日前)頃から腹痛が生じるとともに、食欲不振、嘔気・嘔吐出現した。体幹は温かいが、末梢は湿潤冷汗でショック状態。明らかな運動麻痺はみられず。翌日、意識障害出現し、腎機能障害の増悪を認めて徐々に尿量低下し、8月9日18時10分に心肺停止。8月9日21時44分死亡確認。

症状表現の抽出

22システム中1位
(2-stage CRF+辞書)

時間表現の抽出

15システム中3位
(2-stage CRF)

86%

87%

Composition	Precision	Recall	F1 score
BASELINE	87.87%	81.43%	84.53
SYMPDIS	87.46%	84.18%	85.79
MEDDIC	88.57%	83.45%	85.94
FULL	88.39%	84.76%	86.54

Composition	Tag	Precision	Recall	F1 score
BASELINE	a	86.67%	69.94%	77.23
	h	98.51%	88.00%	92.96
	t	90.42%	85.07%	87.66
	overall	91.26%	83.74%	87.34

NTCIR-MedNLP1 (2012)

[荒牧, 大江2009]

症例報告の有効活用に関する研究

- 二つの学会に同義／表記ゆれ検索機能をもった症例検索システムを提供



社団法人 日本内科学会

The Japanese Society of Internal Medicine



社団法人 日本循環器学会



現在データベースには、すでに約25,000件の演題が登録されており、毎年、新たに約3,100件の演題が追加登録

心筋梗塞

キーワード検索 ☒ 同義語検索

あらゆる年齢で ☐ 40 ☐ 歳以上 ☒ 歳以下

あらゆる性別で ☐ 男性 ☒ 女性

表示順: ☒ 日付(最新順) ☐ 地方会 ☐ 発表者 ☐ 発表者所属 ☐ 分類項目

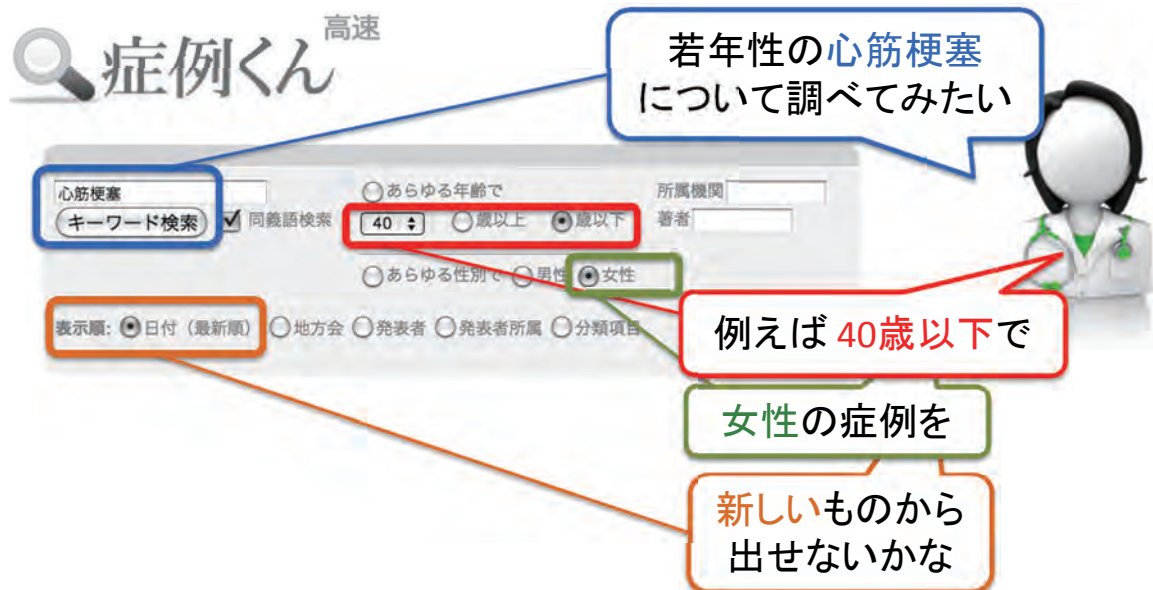
<http://member.naika.or.jp/member/content/ninsho1/search.html>

「症例くん」の検索画面



<http://uth.naika.or.jp>

特徴1) さまざまな要求に 耐えうる症例検索



<http://uth.naika.or.jp>

特徴2) 同義語拡張



同義語検索：「狭心症」以外に「虚血性胸痛」でも検索します。

バージャー病 →

閉塞性血栓血管炎
BUERGER病
ビュルガー病
血栓閉塞性動脈炎

約20,000の同義語
辞書による検索ク
エリ展開

狭心症 →

虚血性共通

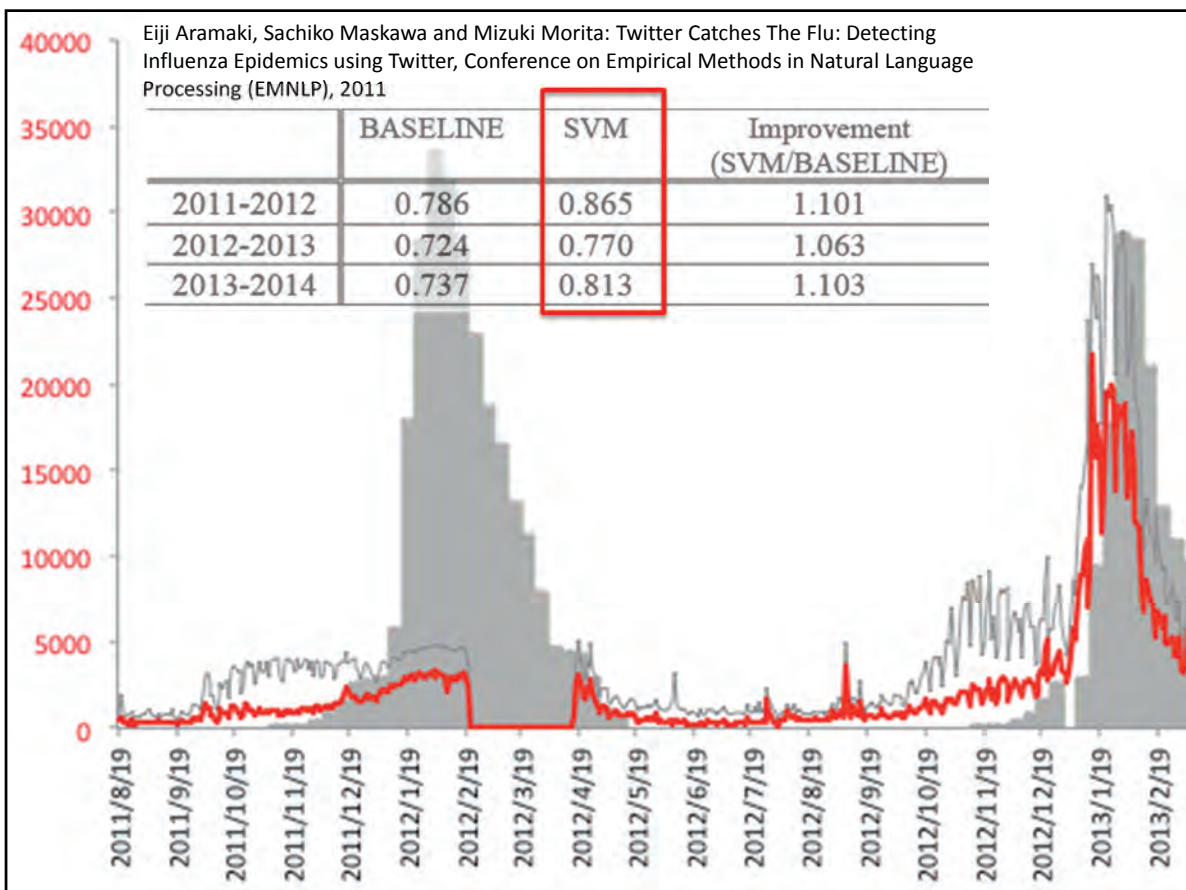
NASH →

非アルコール性脂肪性肝炎
非アルコール性脂肪肝炎

日々蓄積される大量の医学知識（上） と臨床知識（下）



ビッグデータで 病気を防ぐ



もう1つの医療記録 闘病記・闘病ブログ



悪心増悪
胃のむかつき
悪心 (sickness)
食後悪心
嘔気 (vomiting)
吐き気
悪阻 (nausea)

ICD 10 = R16
悪心及び嘔吐
(sickness & vomiting)

Paraphrase (言い換え)

吐き気
ムカつき
ムカムカ
気持ち悪さ
オエーとなる
リバースする



ICD 10 = R16
悪心及び嘔吐
(sickness & vomiting)

認知症者のブログ

発症

4年後

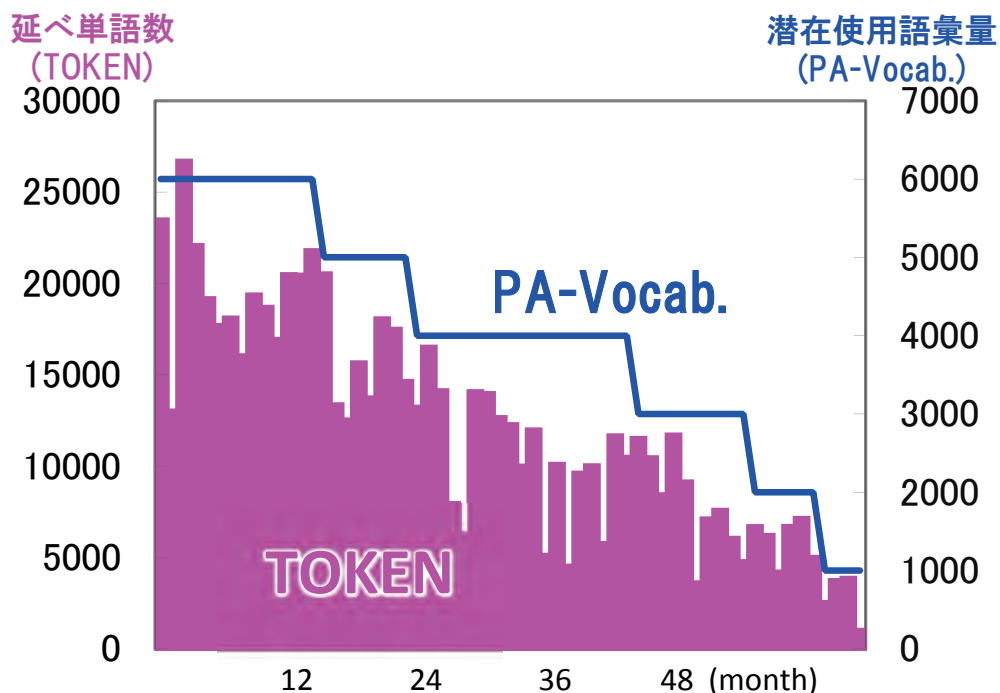
聞き取りをした後長谷川式簡易知能評価スケール（後に本で知る）のテストを受けた。
例貴方は何歳ですか？今日は何年何月何日ですか？私たちが今居るところはどこですか？これから言う言葉を言ってみてください。
桜 猫 電車 次に身長・体重・尿・血液検査をして一日目は終わった。
緊張から開放されると大きなため息がでた。
なんで桜や猫なんだ……と帰りの車の中でいらいらした。

2006 日記開始直後

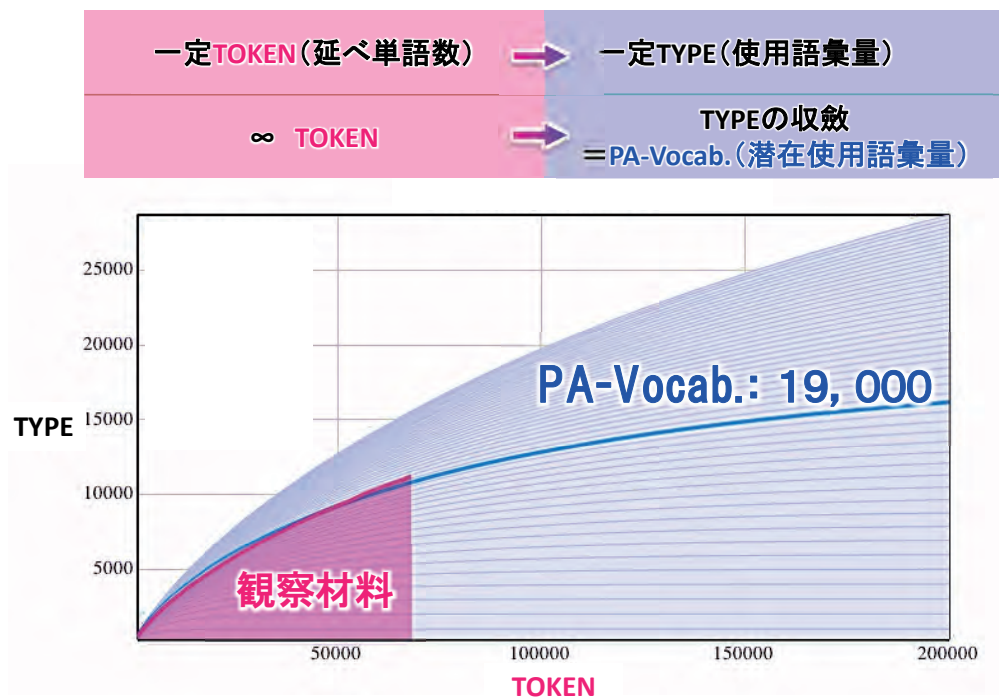
パソコンで文字がかけていたわたしですがなぜか文字の変換が出来ない今日です。
言葉をかくにはそれなりの言葉をさがし読んでいただける言葉にすれなはとはよういではなしのです。
とみにそのかいすが多くなった。
なんで・・・言葉が変換さかと私は私の頭をコツコツと叩いてみましたがそんなことでは反応がなかったです。

2010 日記停止1ヶ月前（原文ママ）

TOKENとPA-Vocab.の推移

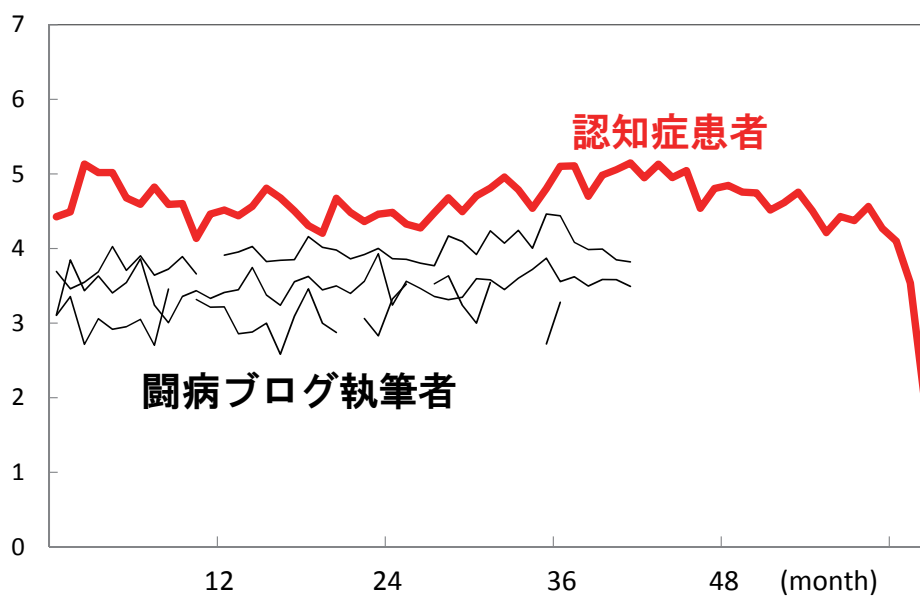


潜在使用語彙量：PA-Vocab. Aramaki + , 2012



25

闘病ブログとの対比 ② MT-Depth 値の推移

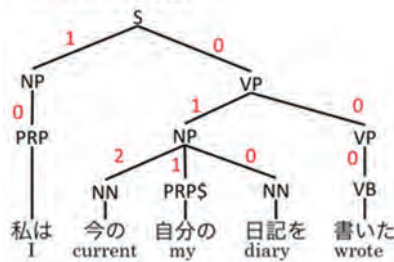


26

Frazier score (1985)



Yngve score (1960)



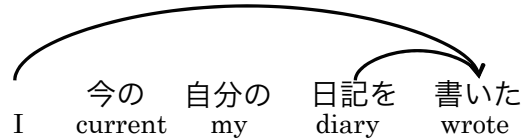
Dependency distance (1995)



文長にあまり依存しない 構文の複雑さの提案

Case Structure Probability (2006)

$$P(I_{\text{sub}}, \text{diary}_{\text{obj}} | \text{wrote}) = e^{-8} \quad P(\text{wrote}) = 0.002$$



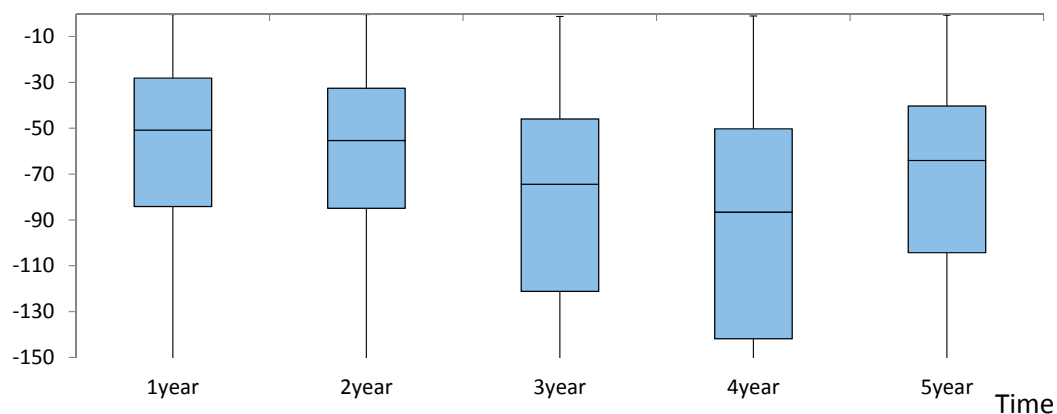
文確率 $\equiv$ 構文解析確率

$$= P(\text{書く}) P_{\text{sub}}(\text{私} | \text{書く}) P_{\text{obj}}(\text{日記} | \text{書く})$$

- 構文解析 $\text{ARGMAX}_{t \in T} \text{Pr}(t)$
 - あらゆる解析結果 T の中から、もっとも解析確率の高い解釈を選ぶ
- 文の生成確率 $\text{MAX}_{t \in T} \text{Pr}(t)$
 - $\equiv$ 構文の自然さ
 - その時の確率を文の自然さとみなす

構文能力も変化する！

Log(P(S))



言 秤

～コトハカリ～

言秤とは、話したことを、自動でリアルタイムに解析して、言語能力を測定することができる装置です。

何が測れるの？

言語能力が測れます。

おしゃべりをするだけで、言語能力のひとつ、**語彙能力**（単語や話の内容）を測定することができます。
語彙能力にはいろいろな種類があります

言秤～コトハカリ～では、6つの特徴が測定できます。

1. 使う語彙の難しさ
2. 使う語彙の量
3. 使う語彙の具体性
4. 使う語彙の特殊性
5. 使う文法の難しさ
6. 使う文法の丁寧さ

海外の研究では**認知症は言語能力と関係があると言われています**。^{1) 2)}
中でも、語彙能力が高い人が認知症になる割合は低い人に比べて大幅に低いと言われており、更にバイリンガルの人が認知症になる割合も、そうでない人とくらべて大幅に低いと言われています。
今日本では、言秤～コトハカリ～で、日々おしゃべりしながら言語能力を測る習慣をつけることで、認知症予防に役立てようという研究が始まっています。

1) Snowden JS, Hulstijn W (2003) Healthy aging and dementia: findings from the Bush Study. *Ann Neurol* 54: 1331-1339.
2) Bialik S, Bialik S. Bilingual delays age at onset of dementia, independent of education and immigration status. *Neurology* 2014 May 27; 82(11):1136.

ご意見・お問合せ先：京都大学 学際融合教育研究推進センター デザインユニット MedNLP
http://mednlp.jp E-MAIL: mednlp.office@gmail.com

修徳学区認知症を地域で見守るネットワーク主催

第3回 認知症 あんしん相談会

と き 平成27年3月15日(日) 午後2時～4時 (開場午後1時半)

と こ 修徳ふれあい福祉会館4階せんたんホール

京都大学木下先生と京都府立医科大学占部先生による第3回の認知症相談会です。
どなたでも参加していただける。認知症予防のイベントもたくさん！ どうぞお申込みください。

事前申込み 10名様

認知症の個別相談

大好評！セルフチェック出来る！
タブレットによる、もの忘れ診断

当日午後2時～受付終了

京都府お自見入！
言語能力測定「言秤」

当日午後2時～受付終了

～おはなしとエッセイサイズ～
転倒予防で認知症も恐くない！
一今日からできる体操とお部屋の工夫～

当日午後2時15分～

京都大学 木下彩香先生 (医芸部)
京都府立医科大学 占部美恵先生 (看護部)

ご自身やご家族のことについて、専用ブースで個別相談に役立ちます。
病状やお薬、対応についてなど、専門的なアドバイスが受けられます。

京都大学大学院生 野田希聖さん
最新のタブレット操作で、最新のものを忘れ診断ができます。今から準備中！

京都大学 四方美子さん (医芸部)
京都大学 高田真衣さん (看護部)

おしゃべりするだけで、あなたの言語能力をリアルタイムで自動診断。
語彙能力をあげることで、認知症の予防にもつながります！

京都大学 正木光裕 准教授
春になりお出かけの機会も増えますね。転倒しないよう身体とお部屋もちょっと動かしてみよう！

「個別相談」のお申込み、またイベントについてのお問い合わせは

高齢サポート・修徳

(京都市修徳地域包括支援センター)

☎075-351-2153

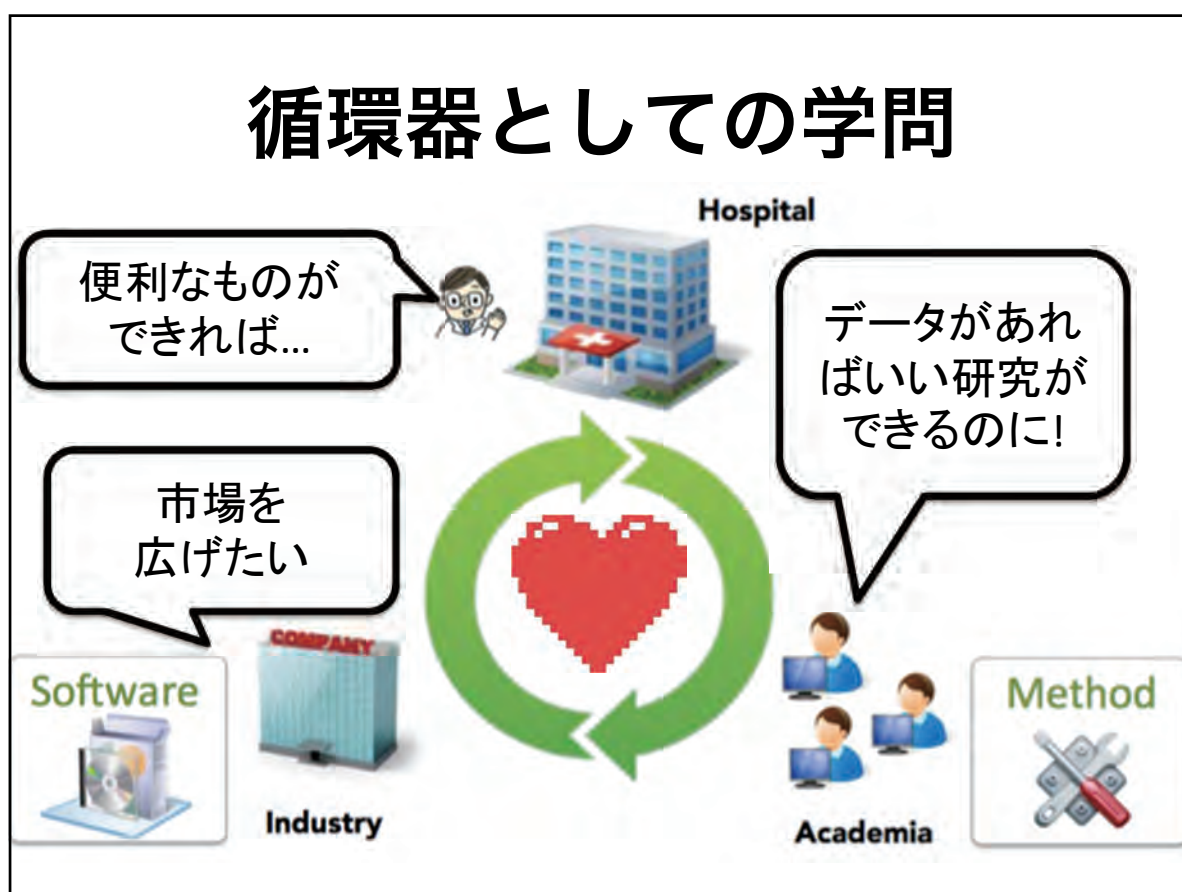
京都市下京区豊島町110-1 総合福祉施設 修徳 1F 地下鉄大宮駅(五島)駅より徒歩約7分

さいごに

One and Only non-English MedNLP Competition

Better NLP, Better Medicine

	NLP Researcher		Bioinformatics
	Medical NLP		Bioinformatics
	NLP Researcher		NLP Researcher (Fuji Xerox)
	Framework Tool sharing		Company viewpoint



仮想患者のカルテを配布

- (I) 模擬病歴報告 (70症例)
 - 仮想の患者を想定して、医師が記述した病歴報告
- (II) 授業用教材のデータ (50症例)

	(I)	(II)
消化管・腹壁・腹膜疾患	4	11
肝・胆・膵疾患	2	10
心臓・脈管疾患	12	10
内分泌・代謝・栄養疾患	5	9
腎・泌尿器疾患	4	9
免疫・アレルギー性疾患・膠原病	5	6
血液・造血器疾患	1	7
感染症	6	9
呼吸器・胸壁・縦隔疾患	11	11



MedNLP 参加者

国立保健医療科学院
National Institute of Public Health
トーマツ
Deloitte Touche Tohmatsu

北陸先端科学技術大学院大学
JAIST
北海道大学
Hokkaido University
京都大学
Kyoto University
岡山大学
Okayama Prefectural University
岡山県立大学
Okayama Prefectural University
東京大学
The University of Tokyo
奈良先端科学技術大学院大学
Nara Institute of Science and Technology
安田女子大学
Yasuda Women's College

国立中央大学 (台湾)
National Central University
朝陽科技大学 (台湾)
Chaoyang University of Technology
南京大学 (中国)
Nanjing University
中央研究院 (台湾)
Academia Sinica
ダブリン大学 (英国)
Dublin City University
NEC アメリカ (米国)
NEC USA

日本ユニシス
Nihon Unisys, Ltd
日立中央研究所
Hitachi, Ltd.
NTT研究所
NTT Science and Core Technology
Laboratory Group
数理システム
MSI Knowledge
富士ゼロックス
Fuji Xerox
日立中央研究所
Hitachi, Ltd.
NTTデータ
NTT Data

Task	MedNLP-1	MedNLP-2
De-identification	6 groups (11 systems)	-
NER	11 groups (15 systems)	10 groups (24 systems)
ICD-coding	-	9 groups (19 systems)
Free	1 groups (1 systems)	2 groups (2 systems)

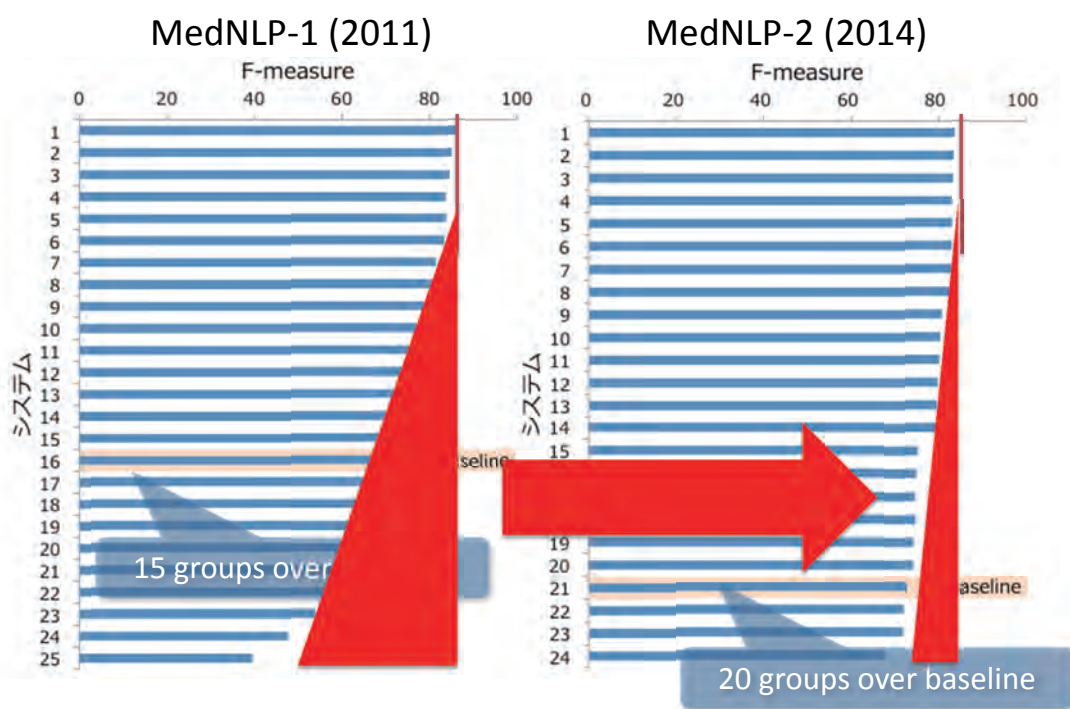
Milestone

What kind of Task is required?

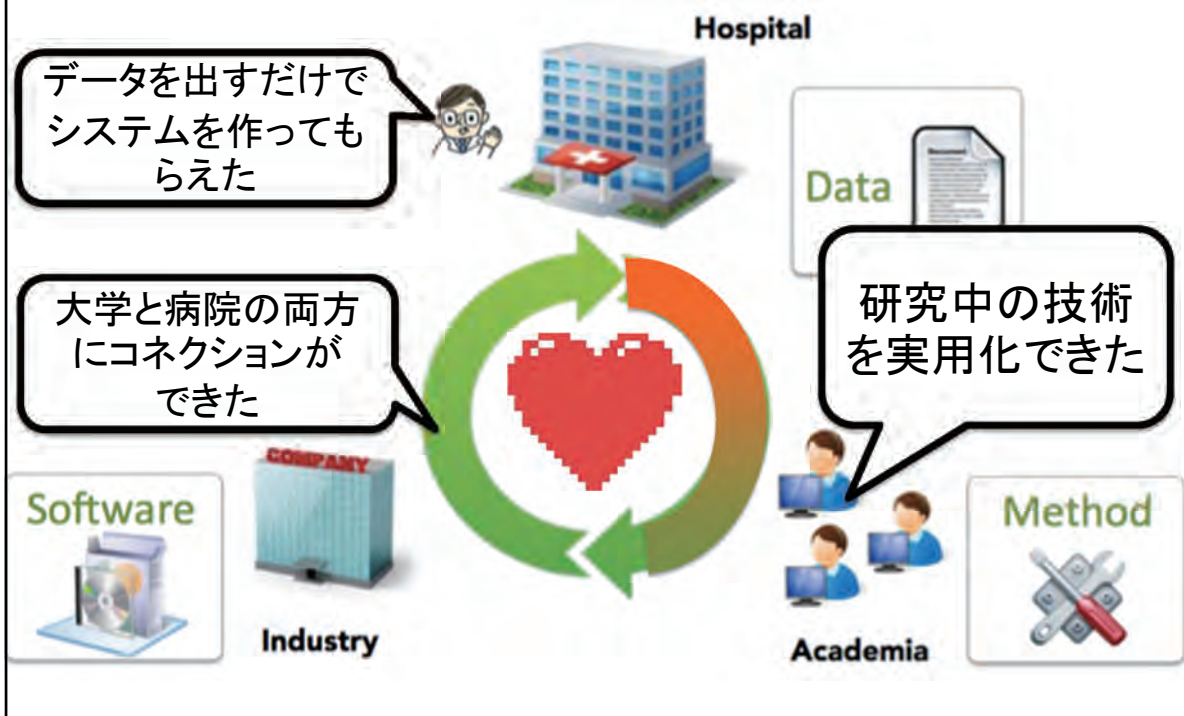
京大病院来院5日前から腹が痛むとのこと



MedNLP-1 << MedNLP-2



循環器としての学問



MedNLP: 医療言語処理プロジェクト

	原著論文	国際会議	受賞
2014年度	5編	4編	7賞
2013年度	4編	11編	4賞
2012年度	3編	6編	6賞
2011年度	2編	2編	4賞
累計	25編	48編	29賞

外部獲得資金（5年以内）

NTCIR MedNLP-2, NII共同研究費, 2013年度-2014年度, 研究代表者: 荒牧英治, 2,000千円.
 挑戦的萌芽研究, 2013年度~2015年度, 「テキストの安全な匿名化に関する研究」研究代表者: 荒牧英治, 2,800千円.
 基盤研究(A), 2011年度~2014年度, 「確率関係モデルによる医療臨床データの高度活用に関する研究」, 研究分担者: 荒牧英治(代表者: 麻生英樹), 47,970千円.
 若手研究(A), 2011年度~2014年度, 「表記ゆれ及びそれに類する現象の包括的言語処理に関する研究」, 研究代表者(個人型研究), 17,940千円.
 挑戦的萌芽, 2010年度~2012年度, 「ダミー診療録の構築および自動構造化に関する研究」, 研究代表者: 荒牧英治, 3,350千円.
 特定研究, 2009年度~2011年度, 「コミュニティ型コンテンツのコンテンツホール検索の研究」, 研究分担者: 荒牧英治(研究代表者: 瀬本明代), 5,000千円.
 若手研究(A), 2008年度~2011年度, 「非文法的かつ断片化したテキストからの情報抽出に関する研究」, 研究代表者: 荒牧英治, 9,980千円.

奨学寄付金, 2012年度, (株)カレン, 2012年度, 「ソーシャルメディア上の発言の分類器の構築に関する研究」, 研究代表者: 荒牧英治.
 奨学寄付金, 2012年度, (株)UTIX, 「気象データを用いた風邪流行予測サイト・カゼミルの高度化」, 研究代表者, 研究代表者: 荒牧英治.
 奨学寄付金, 2012年度, (株)ブルワーク, 「生命科学DB横断検索」, 研究代表者: 荒牧英治.
 奨学寄付金, 2012年度, Microsoft Research Asia(マイクロソフトアジア研究所), CORE8, 「患者モデルを用いた疾患観測モデルの構築に関する研究」, 研究代表者: 荒牧英治.
 奨学寄付金, 2010年度, Microsoft Research Asia(マイクロソフトアジア研究所), Microsoft Research Asia eHealth Theme Program, 研究代表者.
 奨学寄付金, 2010年度, (株)UTIX, 「ウェブからの疾病情報の大規模かつ即時的な抽出手法」, 研究代表者, 研究代表者: 荒牧英治.

プロジェクトの受賞（5年以内）

2014, グループウェアとネットワークサービスワークショップ2014, ベストプレゼンテーション (9%=2件/22件) (研究主導者(PI)として) .
 2014, グループウェアとネットワークサービスワークショップ2014, ベストペーパー賞 (9%=2件/22件) (研究主導者(PI)として) .
 2014, 第13回情報科学技術フォーラム (FIT2014), FIT奨励賞 (15%=1件/約7件) (研究主導者(PI)として) .
 2014, 京都大学 学際研究着想コンテスト, 優秀賞 (10%=3件/32件) (研究主導者(PI)として).
 2014, 日本災害食学会, (カゴメ賞 (企業賞)) (共著者として) .
 2014, 第33回医療情報学連合大会 (第14回日本医療情報学会学術大会), 優秀論文賞 (3%=数件/約200件) (研究主導者(PI)として) .
 2014, データ工学と情報マネジメントに関するフォーラム(DEIM), 最優秀インタラクティブ賞 (0.6%=1件/約160件) (研究主導者(PI)として) .
 2013, 京都大学 学際研究着想コンテスト, 奨励賞 (研究主導者(PI)として) .
 2013, 第12回情報科学技術フォーラム (FIT2013), FIT奨励賞 (12%=1件/約8件) (研究主導者(PI)として) .
 2013, 第32回医療情報学連合大会, 優秀口演賞 (3%=数件/200件) (共著者として).
 2013, 第31回社会言語科学会, 研究大会発表賞 (3%=1件/30件) (研究主導者(PI)として).
 2012, WebDBフォーラム, 企業賞 (チームラボ賞) (共著者として) .
 2012, 言語処理学会 第18回年次大会, 優秀賞 (2-3%=数件/約340件) (研究主導者(PI)として) .
 2012, 言語処理学会 第18回年次大会, 若手奨励賞 (2-3%=数件/約340件) (共著者として) .
 2012, テキストアノテーションワークショップ 2012, 奨励賞 (16%=3件/18件) (研究主導者(PI)として).
 2012, マルチメディア、分散、協調とモバイルシンポジウム(DICOMO), ヤングリサーチ賞 (10%=30件/300件) (研究主導者(PI)として).
 2012, 第31回医療情報学連合大会, 優秀賞 (1.5%=5件/300件) (共著者として).
 2011, CLIO Healthcare Awards 2011, Gold Awards (最高賞) (技術提供サイト「カゼミル」による) (世界最高峰の広告コンクール)
 2011, Spikes Asia Advertising festival (スパイクス アジア 広告祭) 2011, デジタル部門 Gold Awards (第二位) (技術提供サイト「カゼミル」による)
 2011, 第81回グループウェアとネットワークサービス研究会, 優秀発表賞 (研究主導者(PI)として) (10%=2件/19件) .
 2011, Emerald Literati Network 2011 Awards for Excellence, Outstanding Paper Award Winner (共著者として).
 2010, 言語処理学会 第16回年次大会, 優秀発表賞 (2%=5件/270件).
 2010, 第29回医療情報学連合大会, 研究奨励賞 (2%=数件/300件) (共著者として).
 2010, 第29回医療情報学連合大会, 優秀口演賞 (2%=数件/300件) (共著者として).
 2010, データ工学と情報マネジメントに関するフォーラム (DEIM), 最優秀論文賞 (共著者として) .



患者文書プロジェクト

言語処理分野の最前線 2
日本語を対象とした統計的機械翻訳の進展

Graham Neubig

奈良先端科学技術大学院大学 助教

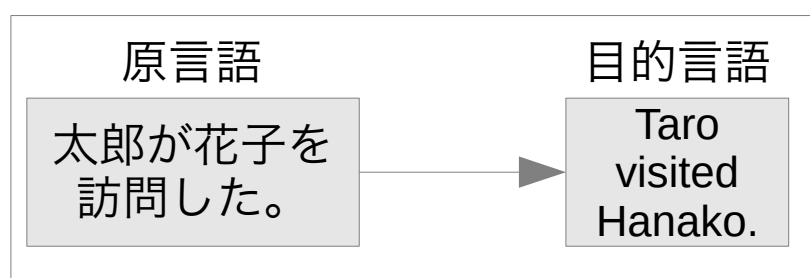
日本語を対象とした 統計的機械翻訳の進展

Graham Neubig
奈良先端科学技術大学院大学 (NAIST)
2015-2-24

1

機械翻訳

- 原言語から目的言語へと自動的に翻訳



- 近年に著しい発展と実用化

Google translate

excite

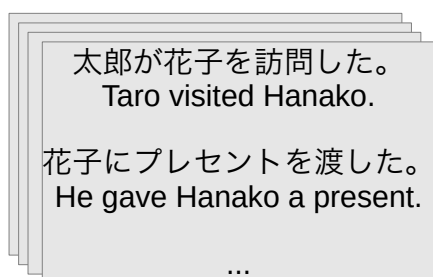


2

統計的機械翻訳 [Brown+ 93]

- 大量の学習データからシステムを自動的に学習

対訳文



モデル

翻訳モデル

並べ替えモデル

言語モデル

3

2010 年からの進展

- 1) 並べ替えの問題を克服
事前並べ替え、統語ベース統計翻訳
- 2) 原言語に含まれない単語の推定
助詞の挿入、ゼロ照応解析
- 3) 確率推定の高度化
ニューラルネットモデル

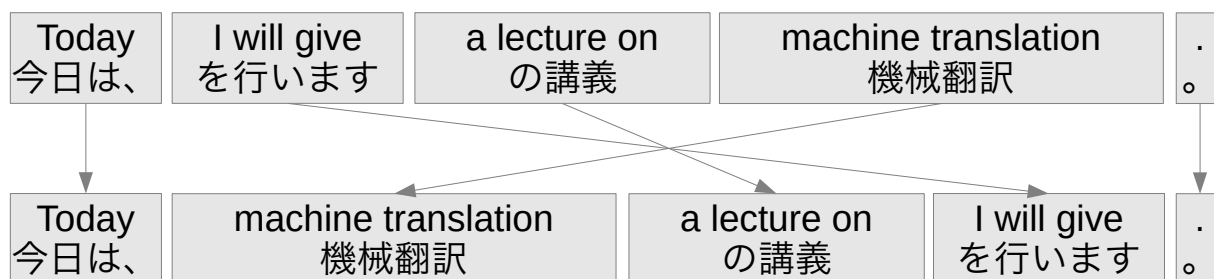
4

1) 並べ替えの問題の克服

フレーズベース機械翻訳 [Koehn+ 03]

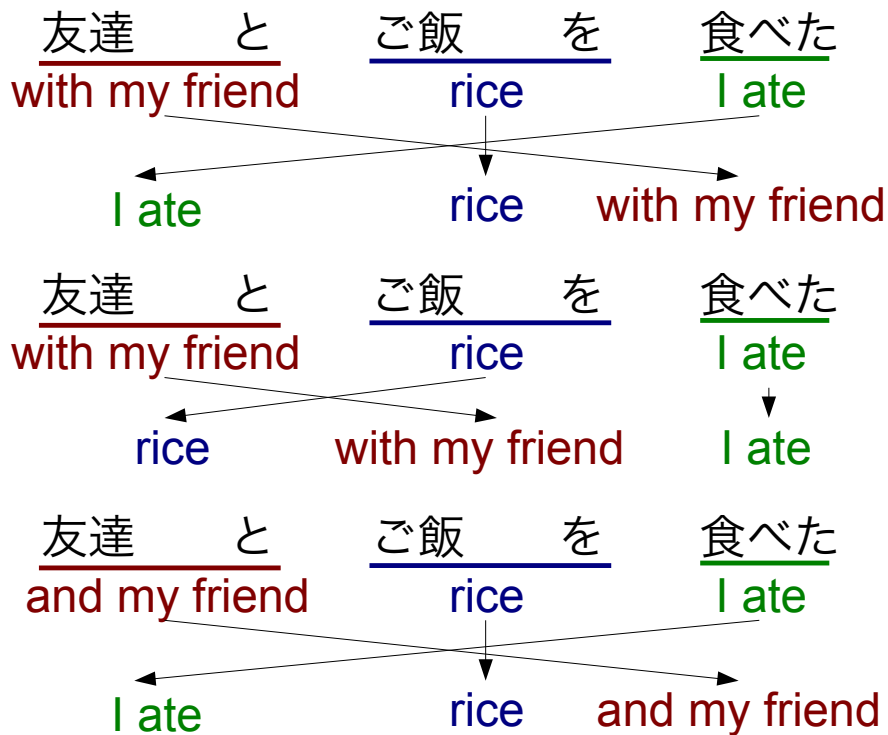
- 文をフレーズ（単語列）ごとに翻訳して、並べ替え

Today I will give a lecture on machine translation .



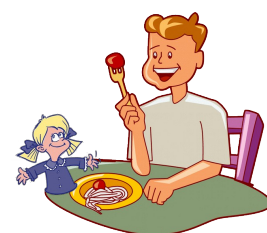
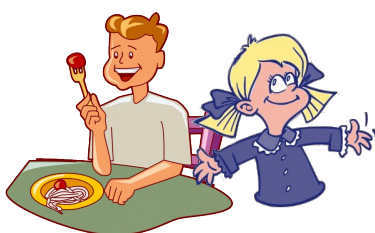
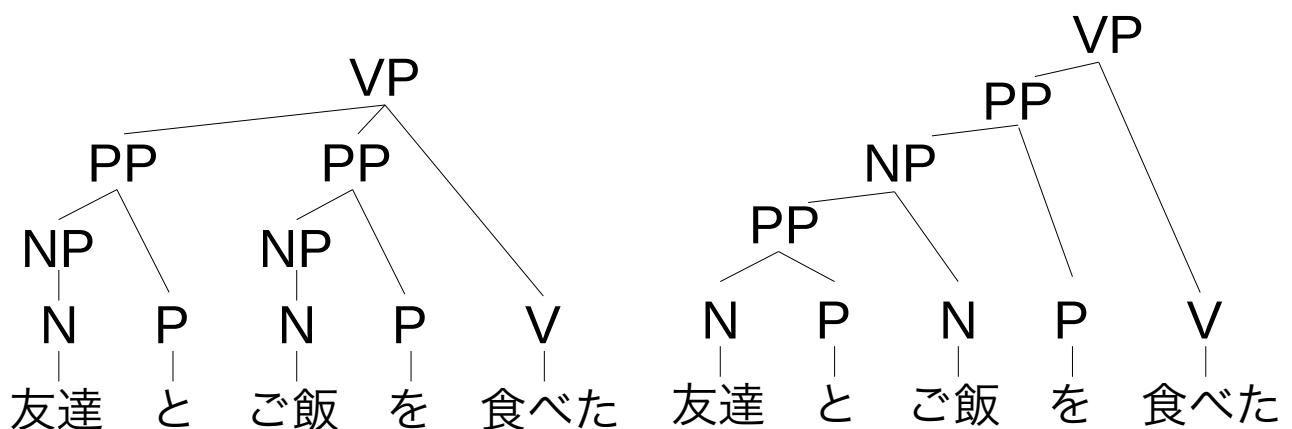
今日は、機械翻訳の講義を行います。

フレーズベース翻訳と並べ替え



7

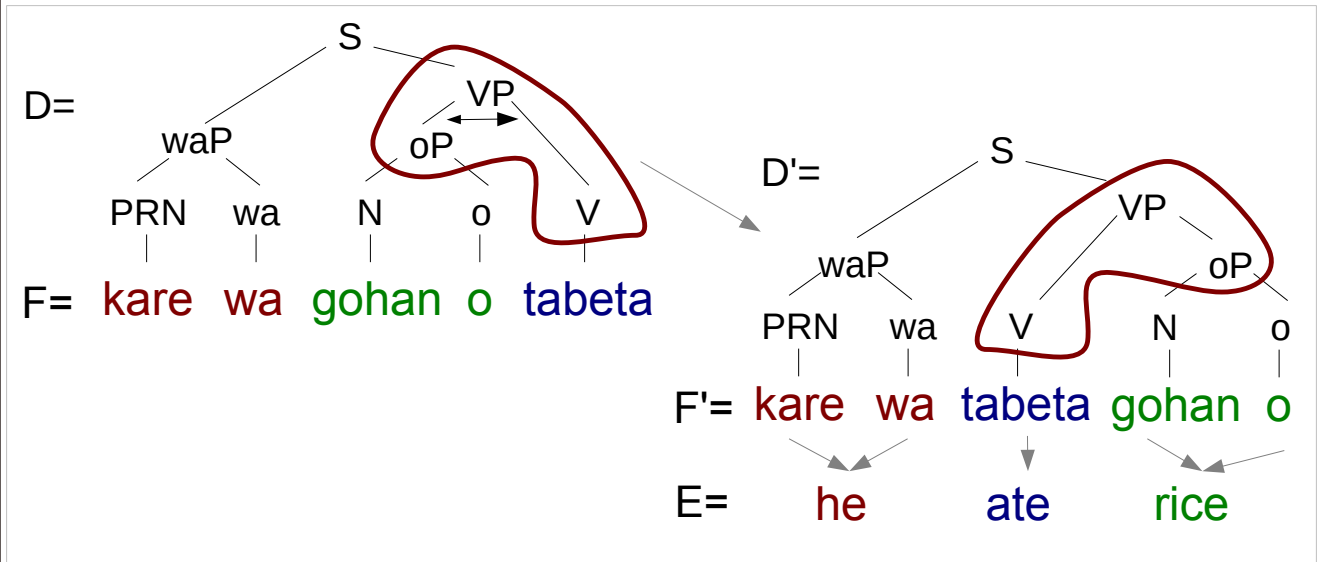
構文解析



8

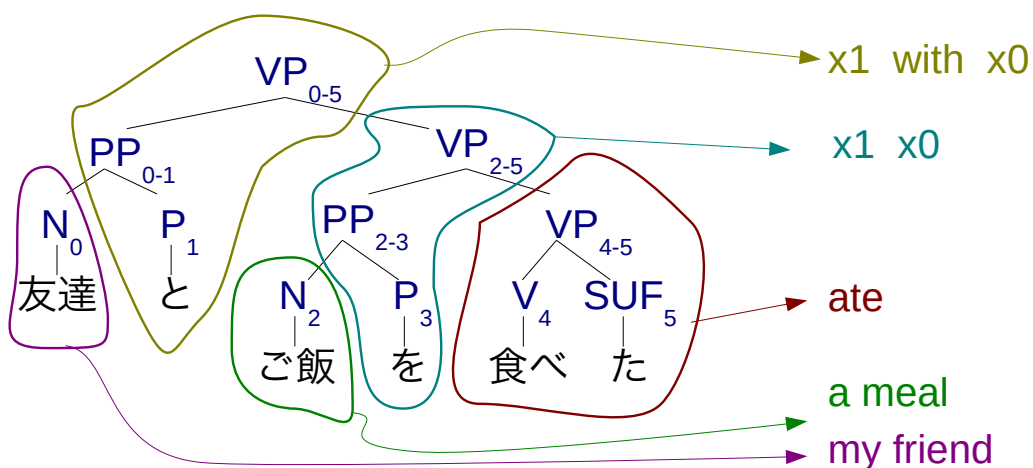
事前並べ替え [Xia+ 04, Isozaki+ 10]

- 原言語の構文木に対して並べ替えルールを定義



9

Tree-to-String 翻訳 [Liu+ 06, Neubig+ 14]



10

日本語における Tree-to-String 翻訳実験

- 少しの工夫で既存手法を大幅に上回る [Neubig+14]

入力 In equipment which generates CT by pressure waves, it was confirmed that development hours of CT bubbles increase in proportion to the 0.5th power of input power of pressure waves.

PB 気泡を生成することを確認した圧力波により c t 装置において, 0. 5 乗に比例して増加し, 圧力波の入力電力の c t の開発時間。

T2S 圧力波による c t を発生する装置において, c t 気泡の開発時間が圧力波の入力パワーの 0. 5 乗に比例することを確認した。

正解 圧力波により c t を発生させる装置では圧力波の入力パワーの 0. 5 乗に比例して c t 気泡の発達時間が増大することを確認した。

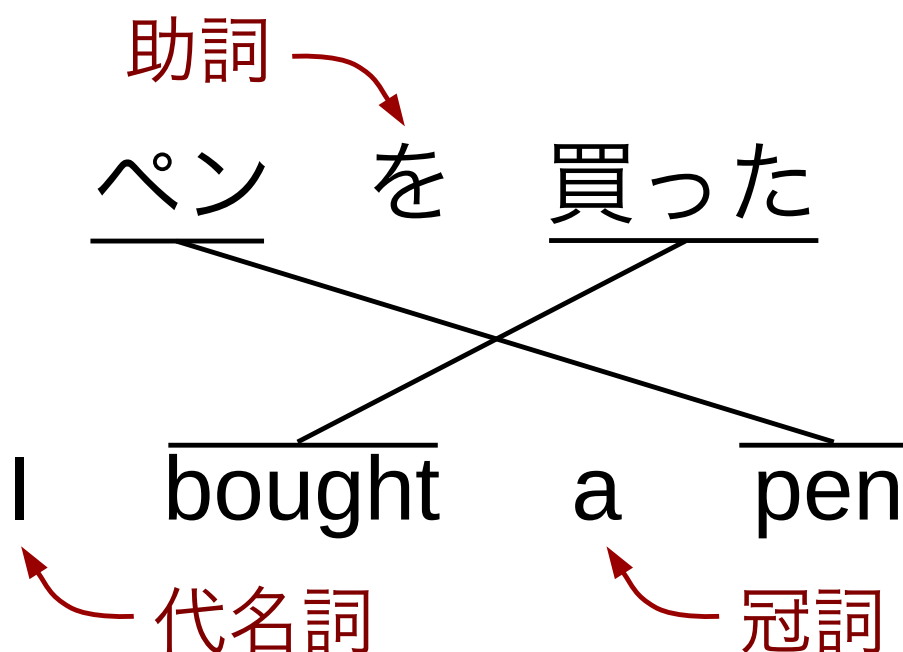
- オープンソース公開
<http://www.phontron.com/travatar>

11

2) 原言語に含まれない単語の推定

12

日英のヌル対応



13

助詞・冠詞の前処理 [Isozaki+ 10]

I bought a pen

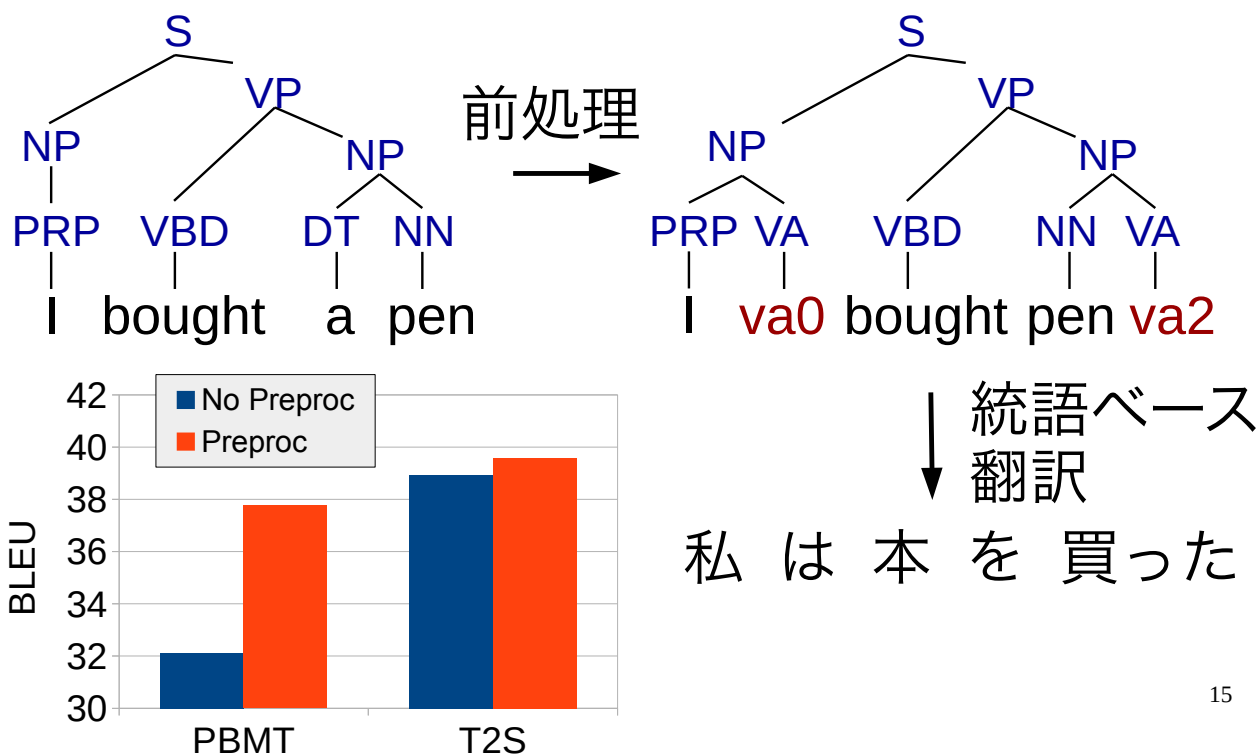
I va0_ pen va2_ bought

↓ ↓ ↓ ↓ ↓

私 は ペン を 買った

14

統語ベース翻訳における前処理 [Hatakoshi+ 14]



15

主語の補完・深層格解析 [Kudo+ 14]

- 同時に主語と深層格を解析し、事前並べ替えに利用

今日は {d=other} 酒が {d=obj} 飲める {v=potential, z=I}
I can drink alcohol today

ニュースが {d=subj} 伝えられた {v=passive, z=already_exist}
The news was conveyed

パスタは {d=obj} 食べましたか {v=active, z=you}
Did you eat pasta?

あなたは {d=subj} 食べましたか {v=active, z=already_exist}
Did you eat?

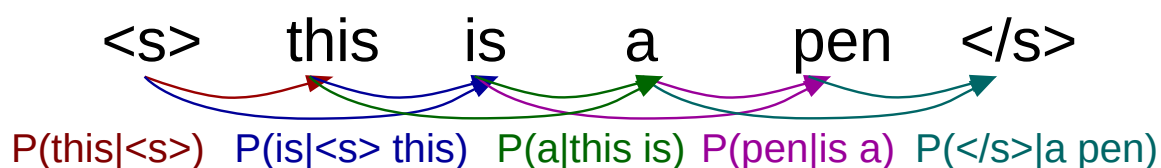
16

3) 確率モデルの高度化

17

n-gram 言語モデルとその弱点

- 次の単語の確率を前の n-1 単語で推定



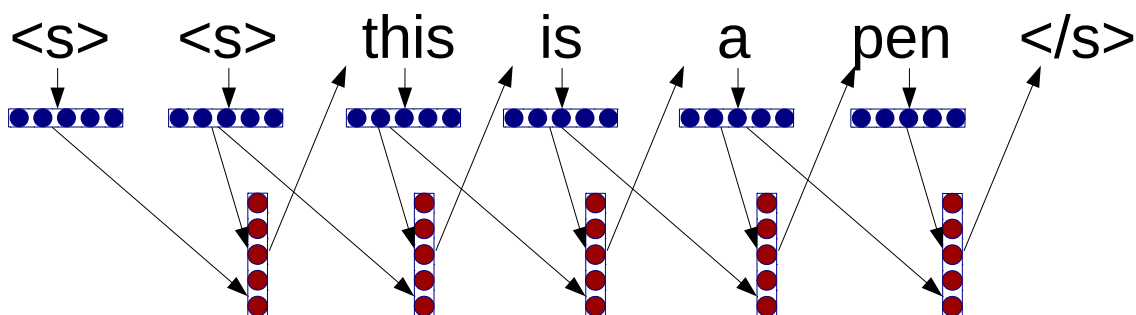
- 単語の類似性は未考慮
- 珍しい文脈で壊れる

my pet **cat**
 my pet **dog**
 my pet **wallaby**

my name is **john** smith
 my name is **jacob** smith
 my name is **sylvan** smith

18

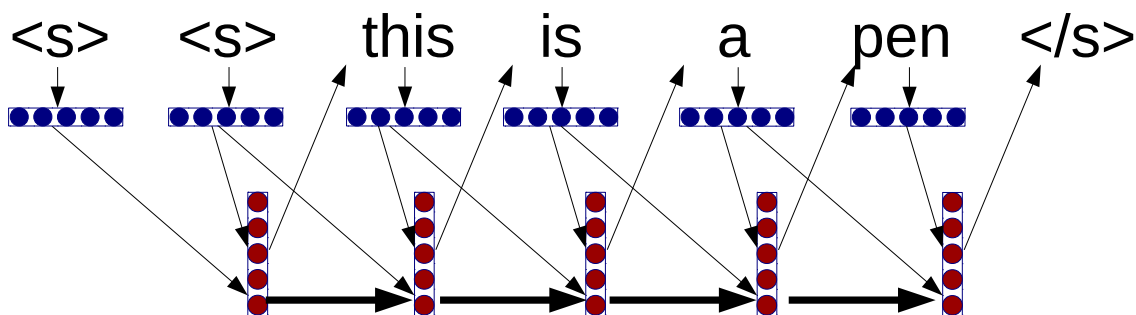
ニューラルネット言語モデル [Nakamura+ 90, Bengio+ 06]



- 低次元隠れ層で出力の類似性を考慮
- 単語表現で文脈の類似性を考慮
- 文脈のすべての単語を直接考慮するため、未知語を含めた文脈で壊れない

19

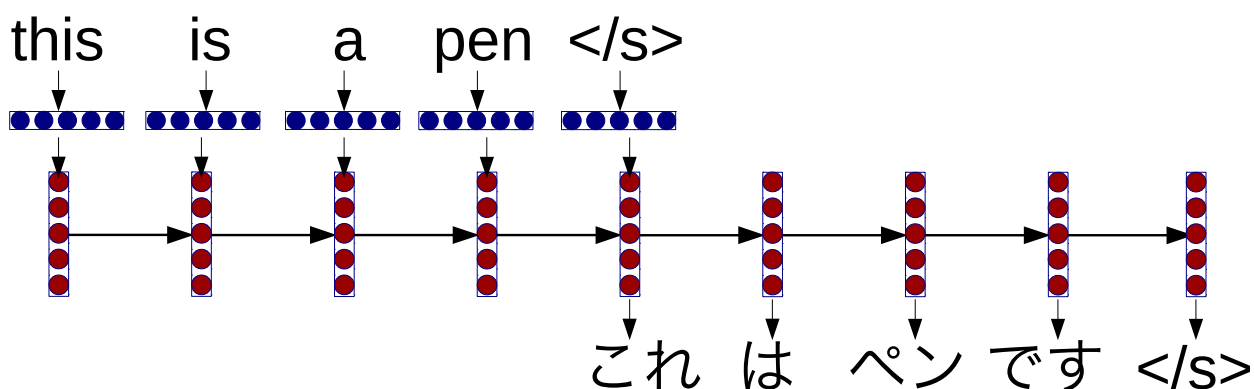
リカレントニューラルネット言語モデル [Mikolov+ 10]



- 以前の単語を「記憶」することが可能
- 4言語対の翻訳における BLEU 改善 [Neubig 14]
英日 : +0.71 日英 : +0.96 中日 : +0.79 日中 : +0.51

20

リカレントニューラルネット翻訳モデル [Sutskever+ 14]



- 独立、もしくは既存手法の一部として利用可能
- 簡単な割に高精度
- これから大きな波紋？

21

英日・日英翻訳の現状？

22

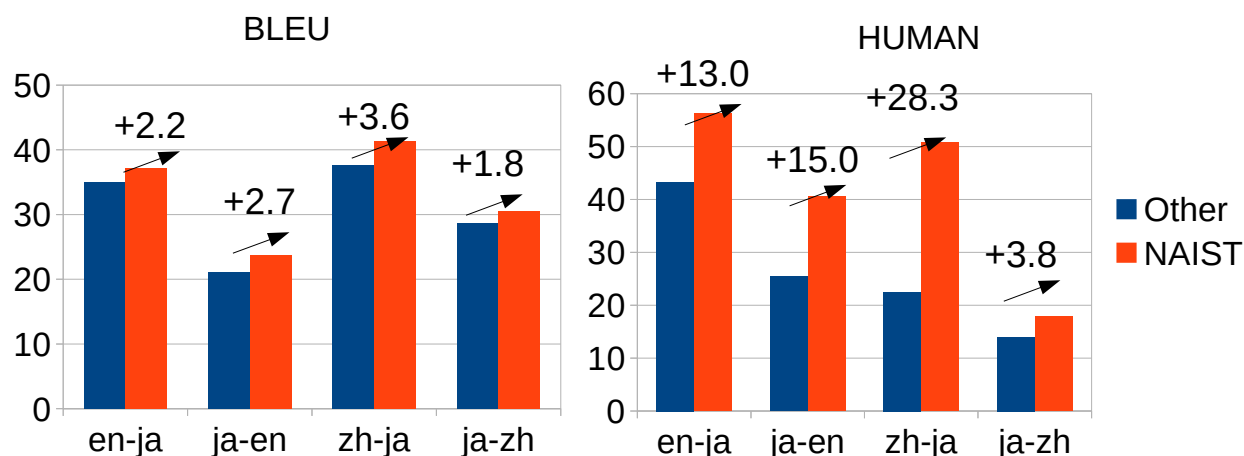
現状の印象

- ルールベース？統計ベース？
 - 英日一般分野：統計ベース > ルールベース
 - 日英一般分野：ルールベース > 統計ベース
 - 専門分野：統計ベース > ルールベース
- 問題の進捗度合い：
 - 高：並べ替え（長文でも）
 - 高：メジャーな慣用句・固有表現
 - 中：主語の補間
 - 中：口語、崩れた文
 - 低：曖昧な語彙の選択

23

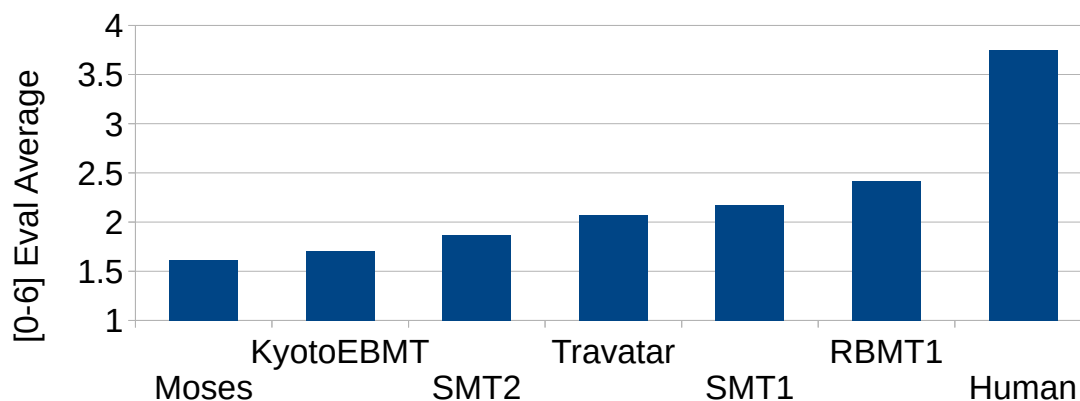
事例：Workshop on Asian Translation

- 科学論文の英日・日英・中日・日中翻訳
 - 専門分野、50~200 万文の豊富なデータ
- T2S+NNLM でルールベース・統計ベースを凌いで首位



事例：Project Next

- 現代日本語書き言葉均衡コーパス
- 日英の様々な分野、ぴったりのデータがない



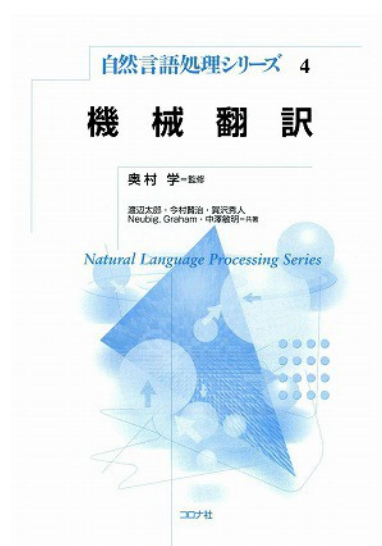
- すべてのシステムに共通の課題：
文脈依存の語彙選択、解釈が曖昧な文

25

ご清聴ありがとうございました！

更に勉強するには：

コロナ社「機械翻訳」



参考文献

- Y. Bengio, H. Schwenk, J.-S. Sencal, F. Morin, and J.-L. Gauvain. Neural probabilistic language models. In *Innovations in Machine Learning*, volume 194, pages 137–186. 2006.
- P. F. Brown, V. J. Pietra, S. A. D. Pietra, and R. L. Mercer. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19:263–312, 1993.
- Y. Hatakoshi, G. Neubig, S. Sakti, T. Toda, and S. Nakamura. Rule-based syntactic preprocessing for syntax-based machine translation. In *Proc. SSST*, 2014.
- H. Isozaki, K. Sudoh, H. Tsukada, and K. Duh. Head finalization: A simple reordering rule for SOV languages. In *Proc. WMT and MetricsMATR*, 2010.
- P. Koehn, F. J. Och, and D. Marcu. Statistical phrase-based translation. In *Proc. HLT*, pages 48–54, 2003.
- T. Kudo, H. Ichikawa, and H. Kazawa. A joint inference of deep case analysis and zero subject generation for Japanese-to-English statistical machine translation. In *Proc. ACL*, pages 557–562, 2014.
- Y. Liu, Q. Liu, and S. Lin. Tree-to-string alignment template for statistical machine translation. In *Proc. ACL*, 2006.
- T. Mikolov, M. Karafiat, L. Burget, J. Cernocký, and S. Khudanpur. Recurrent neural network based language model. In *Proc. 11th InterSpeech*, pages 1045–1048, 2010.
- M. Nakamura, K. Maruyama, T. Kawabata, and K. Shikano. Neural network approach to word category prediction for English texts. In *Proc. COLING*, 1990.
- G. Neubig and K. Duh. On the elements of an accurate tree-to-string machine translation system. In *Proc. ACL*, 2014.
- I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*, pages 3104–3112, 2014.
- F. Xia and M. McCord. Improving a statistical MT system with automatically learned rewrite patterns. In *Proc. COLING*, 2004.

27

言語処理分野の最前線 3
特許オントロジーの構築と
その特許検索への応用

難波 英嗣

広島市立大学大学院情報科学研究科 准教授

特許オントロジーの構築と その特許検索への応用

2015年2月24日

広島市立大学大学院情報科学研究科
難波英嗣

特許検索履歴の例

H01L23/00@A*[製品管理情報/TX+??製品管理
情報+特性/TX+??特性+性能/TX+??性能]*[バー
コード/TX+??バーコード]

上位—下位

対象—属性

研究の目的

オントロジーとは

文献検索や高度な言語処理を行なうための有用な情報源

オントロジー構築には非常にコストがかかる

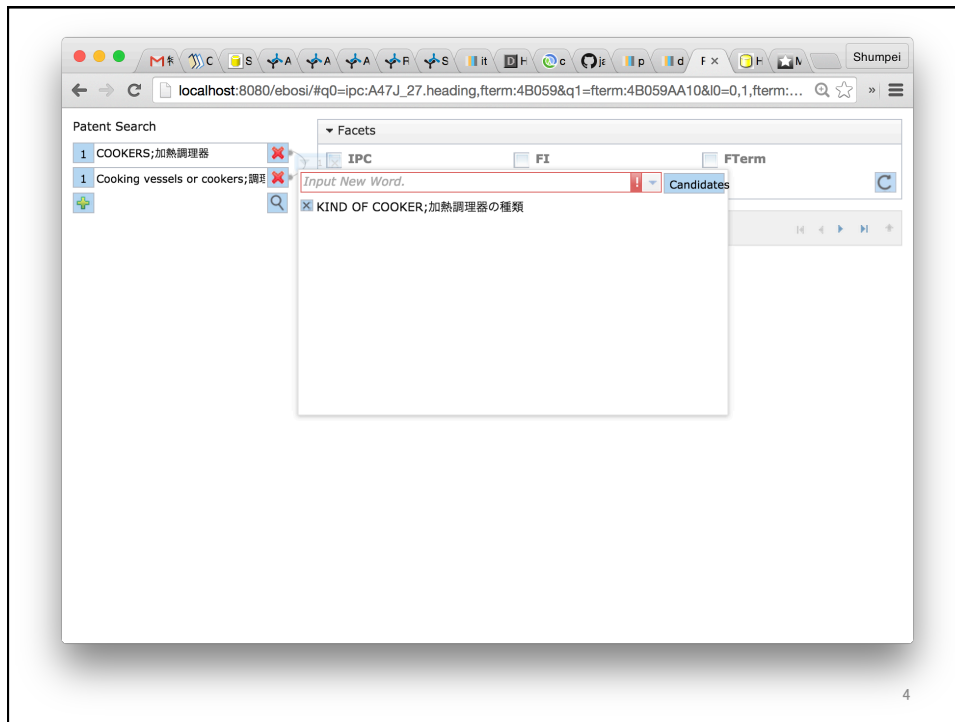
- 効率的なオントロジー構築手法の提案
- 特許検索への応用

2

表 1: 本手法により獲得されたインスタンス

カテゴリ	インスタンス
上位下位	(ポインティングデバイス, マウス) (入力装置, マウス)(入力装置, キーボード) (識別情報, ユーザ ID)(識別情報, パスワード) (記録媒体, IC カード)(表示装置, LCD)
部分全体	(プロセッサ, キャッシュメモリ) (記憶装置, 記憶領域)(情報処理装置, 入力装置) (IC カード, メモリ)(IC カード, IC チップ) (ページ, リンク)(メモリセル, トランジスタ)
定義域 属性	(書き込み, 手段)(最適化, 方法) (サービス, 手段)(サービス, 形態) (情報処理装置, 発明)(情報処理装置, 形態) (タッチパネル, 入力手段)(プリンタ, 設定)

3



4

基本方針

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、学術論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

5

オントロジー構築のイメージ

[入力]

間接翻訳 — ピボット方式 上位下位
 間接翻訳 — トランスファー方式 上位下位
 ...

Fタームオントロジーを
入力とし

ブートストラップ+
機械学習により

[出力]

情報機器 — イメージスキャナ 上位下位
 コンピュータ機器 — パソコン 上位下位
 ...

Fタームオントロジーにはない
新たな知識を公開公報から抽出

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

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

属性: 方式
定義域: 機械翻訳
値域: 直接翻訳、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₈

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

上位: 間接翻訳
下位: トランスファー方式、
ピボット方式

テーマ“5B091(機械翻訳)”のFタームコードの例₉

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	属性: 利用技術
BA17	・ ・ ・ ・ ・ ピボット方式	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₁₀

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	属性: 利用技術 定義域: 意味解析 値域: 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₁₁

オントロジー構築のイメージ

[入力]

間接翻訳 — ピボット方式 上位下位
 間接翻訳 — トランスファー方式 上位下位
 ...

Fタームオントロジーを
入力とし

ブートストラップ+
機械学習により

[出力]

情報機器 — イメージスキャナ 上位下位
 コンピュータ機器 — パソコン 上位下位
 ...

Fタームオントロジーにはない
新たな知識を公開公報から抽出

(ステップ1) ブートストラップ

パターン抽出とインスタンス獲得を繰り返して、少量のシードインスタンスを反復的に増やす

初期シード

半導体メモリ
ROM

ROMなどの半導体メモリ

パターン抽出

などの

入力装置
キーボード

キーボードなどの入力装置
パソコンなどのキーボード

キーボード
パソコン

インスタンス
獲得

キーボードをはじめとする入
力装置

***をはじめと
する***

なぜ機械学習を用いるのか

パターン:「XなどのY」

上位下位関係の判別
に代表されるパターン

キーボードなどの入力装置

パソコンなどのキーボード

公開公報
データベース

入力装置 — キーボード 上位下位

キーボード — パソコン 上位下位



機械学習による統計的アプローチを
用いることで解決する

14

(ステップ2) 機械学習による クリーニング

Espressoで収集されたパターン

パターン ID	パターン (下位・上位)
1	XなどのY
2	Xに代表されるY
3	XといったY
4	X等のY
5	XのようなY

今, $(X, Y) = (\text{キーボード}, \text{入力装置})$ および $(\text{パソコン}, \text{キーボード})$ を入力とした時, XとYの間に上位, 下位関係が成立するか, Espressoで収集されたパターン集合を用いて調べたい

15

(ステップ2) 機械学習による クリーニング

Espressoで収集されたパターン

パターン ID	パターン (上位・下位)	X=キーボード Y=入力装置	X=パソコン Y=キーボード
1	XなどのY		
2	Xに代表されるY		
3	XといったY		
4	X等のY		
5	XのようなY		

「キーボードなどの入力装置」や
「キーボードに代表される入力装置」
という表現が公開公報データベース
の中に何回出現するか調べる

16

(ステップ2) 機械学習による クリーニング

これらの値の組み合わせでXとYの間に
上位・下位関係が成立するか判断する

Espressoで収集されたパターン

パターン ID	パターン (上位・下位)	X=キーボード Y=入力装置	X=パソコン Y=キーボード
1	XなどのY	1050	12
2	Xに代表されるY	320	0
3	XといったY	12	0
4	X等のY	1300	100
5	XのようなY	85	1

17

(ステップ2) 機械学習による クリーニング

これらの値の組み合わせでXとYの間に
上位・下位関係が成立するか判断する

Espressoで収集されたパターン

パターン ID	パターン (上位・下位)	X=キーボード Y=入力装置	X=パソコン Y=入力装置
1	XなどのY		
2	Xに代表されるY		
3	XといったY		
4	X等のY	1300	100
5	XのようなY	85	1

どのように組み合わせるかは機械学習
(SVM)で決める

18

カテゴリ	パターン	X=入力装置 Y=キーボード	X=入力部 Y=キーボード	X=入力装置 Y=発明
下位<パターン>上位	YなどのX	543	10	1
	YといったX	134	6	0
部分<パターン>全体	Y等からなるX	30	34	0
	Yで構成されたX	5	9	0
属性<パターン>定義域	YにおけるX	1	0	156
	Yにおいて、X	0	0	46
上位<パターン>下位	Xに含まれるY	207	3	2
	X(例えば、Y	58	1	0
全体<パターン>部分	Xが有するY	4	23	3
	Xを構成するY	3	18	0
定義域<パターン>属性	Xの具体的なY	0	0	86
	Xを実現するY	0	0	28

図 5: 機械学習による用語間の関係判別

19

表 1: 本手法により獲得されたインスタンス

カテゴリ	インスタンス
上位下位	(ポインティングデバイス, マウス) (入力装置, マウス)(入力装置, キーボード) (識別情報, ユーザ ID)(識別情報, パスワード) (記録媒体, IC カード)(表示装置, LCD)
部分全体	(プロセッサ, キャッシュメモリ) (記憶装置, 記憶領域)(情報処理装置, 入力装置) (IC カード, メモリ)(IC カード, IC チップ) (ページ, リンク)(メモリセル, トランジスタ)
定義域 属性	(書き込み, 手段)(最適化, 方法) (サービス, 手段)(サービス, 形態) (情報処理装置, 発明)(情報処理装置, 形態) (タッチパネル, 入力手段)(プリンタ, 設定)

20

結論

特許の分類体系のひとつであるFタームを用い、
これに様々な知識体系を獲得することでオント
ロジーを構築する手法を提案

さらに、特許検索に応用

21

第四部

パネル討論 1

計算機によるデータに基づく言い換え

乾 健太郎

東北大学 教授

計算機による データに基づく言い換え

乾健太郎
東北大学 情報科学研究科

言い換えのあれこれ A classification of paraphrases

Announcement

The classification and examples in this page are not continuously maintained (Last update was February 2007). If you need to get access to a newer version of classification, please refer to the one presented at the CBA workshop in December, 2010: [Typology of Paraphrases and Approaches to Compute Them](#).

はじめに

ひとくちに言い換えといっても、人間が実際に生成・認識している言い換えには様々な種類の現象がある。ここでは、そのうち、主に言語学的な知識に基づいてなされていると考えられる言い換え（語彙・構文的言い換え）の例を集め、(i)言い換えのスコープ、(ii)内容表現が機能表現が、(iii)必要な知識の種類の観点で分類・整理する。また、それぞれの言い換えを計算機上で表現するにあたっての課題を考察する。

なお、定義上は、「言い換え」=「同じ意味の言語表現」であるが、表現が異なるからには何らかの違いがあると考えられる。たとえば、能動文と受動文の対を考えてみよう。「誰が何をどうしたか」という命題部分の意味は同じと考えられるが、話者の視点やどの情報を強調するか、あるいはどの情報が新情報かといった情報構造の違いがある。また、[\[Linken and Hirst, 2006\]](#)で詳しく述べられているように、形式性や態度の違いが表現に現れる場合もある。したがって、今後は言い換えによって何がかわるのかという観点からも分類を進める予定である。

このページに掲載している言い換えの例は、今後の事例収集・整理において他の例に置き換えられたり、例番号を変更される可能性がある。引用なさる場合は、あらかじめご了承ください。

謝辞

ここで紹介した言い換えの例は、奈良先端科学技術大学院大学情報科学研究科自然言語処理学講座（松本研究室）において2003年夏に実施された「言い換え百選」プロジェクト（メンバー：乾健太郎氏、高橋哲朗氏、降幡建太郎氏、藤田篤）の成果をベースに整理した。整理にあたっては、[佐藤理史氏](#)、[山本和英氏](#)にコメントを頂戴した。各氏に感謝の意を表す。

もくじ

節間の言い換え

[分裂文の言い換え] [文分割（連体節主節化）] [文分割（連用節・並列節の分割）] [接続表現の言い換え]

節内の言い換え

[否定表現の言い換え] [比較表現の言い換え] [態・使役の交替] [自他の動詞交替] [授受動詞の交替] [壁塗り／場所格交替] [湧出動詞の交替] [相互格の交替] [補文構文と格要素の交替] [可能動詞の言い換え] [授受の構文の言い換え] [修飾要素の交替（係り先の変更）] [数量詞の遊離]

文法カテゴリを変える言い換え

[名詞と動詞（機能動詞構文の言い換え）] [動詞と形容詞] [名詞と形容詞] [ナ形容詞とイ形容詞]

主辞交替

[名詞句] [句と節] [格と副詞句]

内容語の複合表現の言い換え

[複合名詞の分解・構成] [「AのB」⇔連体節] [複合動詞の分解・構成] [名詞接尾辞の着脱]

機能語／モダリティの言い換え

[機能語相当表現] [取り立て助詞の移動] [助詞による特徴づけの削除] [伝達のモダリティ] [敬語表現の言い換え] [文体の変換]

内容語句の言い換え

[名詞の言い換え] [動詞の言い換え] [形容詞の言い換え] [副詞の言い換え]

慣用表現の言い換え

[慣用句] [表記の揺れ／略語] [換喩]

日本語になくて他の言語にある言い換え

[いろいろとりまぜて]

（藤田篤，～2012）<http://paraphrasing.org/paraphrase.html>

多様な言い換え表現を どうやって収集・管理するか？

パラレルコーパスからの言い換えの獲得

統計的機械翻訳と同じ問題

ただし、良質の大規模パラレルコーパスの入手は翻訳ほど容易でない

夜 11時 30分 まで 開い ている レストラン が あり ます
(今村01)

レストラン が 夜 の 11時 半 まで 営業し てい ます

夜11時30分 ⇔ 夜の11時半

(レストラン)が開いている ⇔ (レストラン)が営業している

(パラレルでない)

生コーパスからの言い換えの獲得

Distributional Similarity

出現文脈の分布が似ている語は意味も似ている

局所文脈に出現する単語を数える

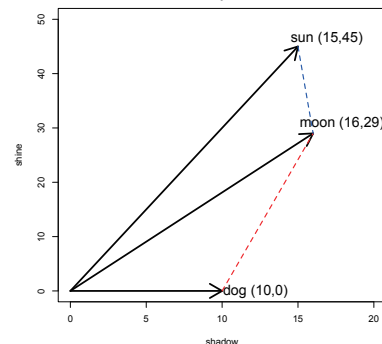
he curtains open and the moon shining in on the barely
ars and the cold, close moon ". And neither of the w
rough the night with the moon shining so brightly, it
made in the light of the moon . It all boils down, wr
surely under a crescent moon , thrilled by ice-white
sun , the seasons of the moon ? Home , alone , Jay pla
m is dazzling snow , the moon has risen full and cold
un and the temple of the moon , driving out of the hug
in the dark and now the moon rises , full and amber a
bird on the shape of the moon over the trees in front
But I could n't see the moon or the stars , only the
rning , with a sliver of moon hanging among the stars
they love the sun , the moon and the stars . None of
the light of an enormous moon . The plash of flowing w
man 's first step on the moon ; various exhibits , aer
the inevitable piece of moon rock . Housing The Airsh
oud obscured part of the moon . The Allied guns behind

(Baroni EACL2012 tutorial)のスライドより抜粋

単語間の共起頻度を表す行列

	planet	night	full	shadow	shine	crescent
moon	10	22	43	16	29	12
sun	14	10	4	15	45	0
dog	0	4	2	10	0	0

似ている



語彙統語パターンの言い換えの獲得も同様

パターン×関係インスタンス行列

	カ フ カ	変 身	カ ミ ユ	異 邦 人	...	...	不 注 意	交 通 事 故	タ バ コ	病 気	...	...
Xが執筆したY	43	24	...	...	...	...	0	0	...	...	...	...
Xの作品Y	34	22	...	...	...	...	0	0	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...
Xが起こすY	0	0	...	...	...	...	23	68	...	...	...	...
Xが原因のY	0	0	...	...	...	...	32	57	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...

各行ベクトル
がパターンの
意味を表す
(パターン意味
ベクトル)

パターンの
クラスター
(意味関係)

著者・著作関係

因果関係

パターン抽出

行列構築

類似度計算

クラスタリング

大規模なデータ
に対して高速・
高効率に動作

大規模コーパス
(60億文, 600GB)

powered by:



関係パターン	意味関係	因果関係
Xが執筆したY	著者関係	イデオロギ
Xの作品Y	著者	因果
...	著者	因果
Xが起こすY	因果関係	因果関係
Xが原因のY	因果関係	因果関係
...	因果関係	因果関係

関係パターン
の言い換え

SVD等の次元削
減でパターンの
意味をよく表現
するベクトルを
獲得

(高瀬, 岡崎ら 2014)

パターン×関係インスタンス行列

(高瀬, 岡崎ら 2014)

543件

Xが->引き起こす->Y

X

Y

並び順

Pattern	X	Y	Freq	PMI
Xが->引き起こす->Y	密室性	犯罪	228	14.8331205257
Xが->引き起こす->Y	ケロロ小隊	騒動	114	14.7290329967
Xが->引き起こす->Y	深部[の]疲労	タンパク質[の]硬化	63	14.5569113848
Xが->引き起こす->Y	ホルモンバランス[の]崩れ	老化	40	14.298825183
Xが->引き起こす->Y	宇宙飛行[の]無重力状態	筋萎縮[の]現象	32	14.1260312992
Xが->引き起こす->Y	発達障害	二次障害	70	14.0405816694
Xが->引き起こす->Y	ユーザー	ウイルス感染	30	13.981354542
Xが->引き起こす->Y	人間模様	本格サスペンス	25	13.8260694086
Xが->引き起こす				13.6444318891
Xが->引き起こす				13.5984088051
Xが->引き起こす				13.5579896451
Xが->引き起こす				13.5579896451
Xが->引き起こす				13.4829801644
Xが->引き起こす->Y	運命	驚愕[の]真実	19	13.4588722281
Xが->引き起こす->Y	体温低下	エネルギー浪費	19	13.4588722281

- 日本語62億文から5300万事例を抽出
(4コア×12台=48CPU分散並列処理で1週間程度)
- Apache Spark上の実装を公開予定

獲得された類義関係パターンの例

(高瀬, 岡崎ら 2014)

Xが->|Yを->|引き起こす

Pattern	Similar
Xや->ウイルスによる->Y	0.311595
Xが->考えられている->Y	0.273014
Xが-> Yを-> 引き起こすと	0.257079
Xが-> Yを-> 引き起こすという	0.228389
Xが-> Yを-> 引き起こします	0.224239
Xは-> Yを-> 引き起こします	0.218061
Xが->なり-> Yに-> なってしまいます	0.205171
Xは-> Yを-> 誘発します	0.204377
Xから->引き起こされる->Y	0.195493
Xが->なって->起こる->Y	0.190255
Xで-> Yに-> なっちゃうよ	0.185453
Xによって-> Yを-> 引き起こす	0.185016
Xが-> Yは-> 影響していると	0.182419
Xが->入って-> Yに-> なる	0.179039
Xは-> Yを-> 引き起こしました	0.159263
Xにより->引き起こされる->Y	0.149517

Xを->取り除く->Y

Pattern	Similar
Xを->除去する->Y	0.476163
Xを->取り除いてくれる->Y	0.449809
Xを->取り去る->Y	0.444928
Xを->除去してくれる->Y	0.386644
Xを->取り除く->ため[の]->Y	0.368256
Xを->取り除くような->Y	0.365391
Xを->とりのぞく->Y	0.358086
Xを->無害化する->Y	0.341775
Xを->無毒化する->Y	0.330335
Xを->剥離する->Y	0.328211
Xを->排出させる->Y	0.32533
Xを->排泄させる->Y	0.317296
Xを->分解除去する->Y	0.309232
Xから->逃れる->ため[の]-> Yを->続けるしか-> ありません	0.307414
Xを->解毒する->Y	0.302393
Xを->和らげる->ため[の]->Y	0.29903

2つの問題

頻度の低い関係表現パターン
(ロングテール)の意味の学習が難しい

文脈の類似性だけでは
類義語と反義語の区別が難しい

パネル討論 2

制限言語と言い換え：
多言語展開の改善へ向けて

影浦 峯

東京大学 教授

制限言語と言い換え：多言語展開の改善へ向けて

影浦峽^ㇿ, 宮田玲^ㇿ, Anthony Hartley[†], Cécile Paris[‡]

ㇿ 東京大学 † 東京外国語大学 ‡ CSIRO, Australia

1 言語とテキスト／文書

最初に、言語表現を前にしたときに、二つの大きく異なる見方があることを確認しておきましょう。それらは、フォーコーの言葉を借りるならば、次のようになります。

〔…〕言語〔ラング〕とは、無限数個の言語運用〔パフォーマンス〕をゆるす有限な規則の集合であるのだ。これに対して、言説とは、じっさいに述べられた言語的まとまり〔シークエンス〕のみからなる、つねに有限で現働的〔アクチャル〕に限定された集合である。〔…〕言語の分析が、ある言説の事実に関して問う問いとはつねに、「どのような規則にもとづいて、このような言表がつくられたのか？ またしたがって、どのような規則によって他の同じような言表をつくることができるのか？」であるのに対して、言説の記述が立てるのはまったく別の問いであって、「このような言表が出現した、しかも、他のいかなる言表もその代わりには出現しなかったのは、どのようなわけなのか？」という問いなのだ [1]。

いわゆる言語学は前者のように言語を見ているとともに、有限の語彙と規則による無限の適格文の生成を考えるというチョムスキーの言葉に代表されるように、文（あるいはそれより小さな単位）を中心的な対象に据えます（もちろん「テキスト言語学」といった領域もありますが、例外的です）。

一方、言説の記述は、一般的に言語現象の単位としてテキストや文書が起点となります。フォーコーの研究では（少し単純化して言うと）テキストや文書の集合が扱われますが、テキストや文書を構成する要素として命題を構成する基本単位である文もまた、言説の記述において対象となる資格を有します。言説を扱う場合、言語学のように背後のメカニズムを探るというよりも存在の形を捉えることが主な課題となることが多

いのですが（そのことは上の引用における「分析」と「記述」という言葉の使い分けにも反映されています）、アプローチの仕方によっては、テキストや文書を生成する規則を考えることもできます。ただしその場合でも、言語の研究のように共時的な体系が許容する潜在的に無限の適格な表現を対象とするのではなく、いわば現実的存在可能性の境界を探ることが主な課題となることは、テキストや文書の性質を考えるならば、明らかです。私たちは、その場で適格な文を作ることができますが、その場で適格な「テキスト」や「文書」を試しに作ることは、その性格上困難である（あるいは定義によっては論理的に不可能である）からです。

2 言い換えと制限言語

私たちが「言い換え」と呼ぶ操作は、どこに位置づけられるものでしょうか。「何か」を「何か」に言い換えるという「言い換え」の基本的な性格を踏まえるならば、言い換えは常に、ある一定の言い換えられるべき「何か」すなわち「言語的まとまり」が「じっさいに述べられた」のち、あるいは「じっさいに述べられ」つつあるときに作動する操作であることが確認できます。その意味で、いわゆる言語学が想定するような、適格文の背後にある有限の語彙と規則と同様のレベルに位置づけられるものではなく、ある具体的な言語表現が実際に出現したとき、あるいは出現するときに作用する、むしろ言説の分析に近いところで捉えられるべき現象であるということが出来ます。

では、「制限言語」についてはどこに位置づけられると考えるべきでしょうか。これについては、二つの可能性があります。制限言語の規則が人工的な言語のように、適格文を生成する語彙と文法を絞り込むかたちで設定され、そのようなかたちで用いられることが想定される場合には、それらは言語学が捉えるような言語の規則に対してさらなる制約を課すものとみなすことができるでしょう。それに対して、制限言語の制限を、具体的な言語表現が作られつつあるとき、あるいは作られたあとで適用するというかたちで定義する

ならば、これは言説の分析に近いレベルで制限言語の位置づけを設定していることになります。

本パネル討論のテーマは「言い換えから見た日本語」であり、その中の一つの話題提起として「制限言語と呼ばれる規則による言い換え」が設定されています。「言い換え」が「じっさいに述べられ」たり述べられつつあるときに作動するものであることを踏まえるならば、パネルのテーマの枠組みで捉えられる「制限言語」もまた、対応したかたちで、言説の分析に近いレベルで位置づけておくことが妥当です。私たちの研究は、幸いにして（というのも最初にお話しがあったときには、パネル全体の枠組み設定をきちんと把握しないままお引き受けすることにしたので）、明示的にそのように位置づけることが可能です。以下、少し私たちの具体的な研究内容に言及しつつ、この点を敷衍していきたいと思います。

3 テキスト／文書と制限言語

「言い換え」が具体的な言語表現が生成されたあとあるいは生成されるときに操作であることを確認し、したがってそれはいわゆる言語学における言語の捉え方よりもむしろ「言説」の研究において採用される視点に近い視点で言語表現を見るものであることを確認し、さらに「制限言語」もそのレベルで捉えることになったので、ただちに、次のような視点が出てきます。すなわち、言い換えや制限言語の規則は、これまで基本的に文あるいはそれよりも小さい単位において定義されてきたが、本来、それらの操作が具体的な言説に即して認識される類のものであるならば、制限言語および言い換えは、具体的な言説の基本的な単位であるテキスト／文書において定義されることが適切である、という視点です。

実際、私たちがテクニカルライティングと言ったりする行為やそこで課される規則、あるいは、他の人が書いた文章を添削すると言った行為は、言い換えの典型的な具体例を伴うとともに言い換えが具体的に発動する典型的な場であり、その中では、例えば段落全体を書き換えたり、順序を入れ替えるといった操作も行われること、また、文や句を対象として言い換えがなされるとしても、その操作は常にテキスト／文書が与える文脈が前提とされていることを観察するならば、ここで述べてきたような理論的検討を経ずとも、言い換えなり制限言語の適用なりといった操作が、基本的に現働的なテキスト／文書をその本質的な対象としていることは明らかです。

とはいえ、言い換えでも制限言語でも、事前に何ら

かのかたちで規則化しておくことを想定するならば、大きくても文の単位までに対象を制限しておかないと難しいという点も、もう一方では明らかです [2]。テキストや文書の言い換え規則や制限規則というのは、どうも文レベルまでで定義される規則のようにきちんと適用できない感があります。テキストの規則が、適格テキストと非テキストを区別できると主張することは非現実的であることと、それは対応しているようです。では、テキスト／文書のレベルを、制限言語／言い換えの問題に技術的にどのように取り入れるか、それが問題となります。

4 制限言語と文書構造

4.1 基本的な枠組み

私たちは現在、自治体文書の多言語化支援システムの研究開発を進めていますが、そこでは、制限言語の考えを、上のような枠組みのもとで、組み込んでいます。私たちがこの研究で採用しているアプローチは上述の視点に立つならば、極めてオーソドックスなものです。基本的には、限定された範囲の現働的なテキスト／文書の生成環境を構築し [3]、その中に、文書構造に依存したかたちで文を基本単位とした制限言語の規則を埋め込む、というものです [4]。

現在の研究では、自治体の手続き文書に対象を限定しています。このように範囲を絞るならば（もしかするともっと一般的なタイプを扱えるかもしれませんが）、実際に存在する文書あるいは存在しうる文書の基本的な構造をかなり具体的に指定することが可能となります。その構造を、私たちは DITA [5] を用いて定義しています。手続き文書は、DITA のトピック型のうち、タスク・トピックが中心となります。DITA タスク本文の各要素を自治体手続きに合わせて具体化したものが、表 1 です [6]。

このような文書構造を指定することで、自治体手続き型文書において「何をどのような順序・構成で書けばよいか」の型が与えられます。

次に、「何をどのような順序・構成で書けばよいか」を受けて「具体的にどのように言語表現を作ればよいか」という言語表現の形を指定していく必要が出てきます。上で述べたように、このレベルでの制限は、基本的に文までの単位に対してしか有効なかたちでは設定しにくいので、ここでいわゆる制限言語の規則を導入します。私たちが以前作成した一般的な制限言語規則 22 個については既に報告していますが [7]、現在は、自治体文書の言語表現を検討し、「法／モダリティ」、「構造」、「語彙」、「テキスト／表記」の 4 つのカテゴリ

表 1: DITA タスク本文の具体化

DITA(初期設定)	詳細機能要素
事前条件 (prereq)	個人条件 イベント条件 アイテム条件
背景情報 (context)	説明 (概要, 目的, 効力, 罰則, 関連概念)
手順 (steps)	必要なものを持参する 申請場所へ行く 様式を提出する (手数料を払う)
期待結果 (result)	得られる結果 (所要期間, 交付物, 連絡)
実行例 (example)	[不要]
タスク完了後の操作 (postreq)	関連手続きへの誘導

りにおいて、合計 55 個のコントロールすべき表現属性を設定しています。例えば、「法／モダリティ」の属性としては、命令形、禁止、可能、時制、態などがコントロールすべき表現として設定されています。

最後に、制限言語の規則の適用を、文書構造の機能要素によって条件付けます。すなわち、どの機能要素について書くときにどの制限言語規則が適用されるかを指定します。

文書構造の指定はいわゆる「言い換え」ではありませんが、仮にある特定の手続き文書のある特定の方法で書くことがある文書作成者の頭の中にあり、その文書作成者が私たちの指定する制限のもとで文書を書くときに頭の中の文書に対して変更を求められるかたちで制限が働いているとするならば、抽象的な意味でこれも「言い換え」と言うことはできるかもしれません。そしてそのレベルでの「言い換え」は、文より下の、制限言語として通常表現されるレベルでの「言い換え」と、人間の文書作成においては自然に繋がっています。ここで説明した枠組みは、従って、そうした人間の文書作成に見られる広い意味での「言い換え」（日常的な言葉では、編集とか推敲と言った方がよいかもしれません）を、文書レベルでの構造の制約と文レベルでの制限言語規則の適用により導くものであると見ることができます。

4.2 可読性と多言語展開の改善

私たちが提案した枠組みは、基本的に、日本語文書の可読性の向上（主に、手続き文書に必要な情報の集合を明示的に定め、論理的な順序で提示する文書レベルでの枠組みにより）と、多言語展開の精度改善に資

すると考えています。少し説明を補足します。

例えば、

身分証明書を持参する。

という文を考えてみます。これを MT にかけてみると、Google translate では

I bring identification.

となります。この表現が、文書中の手続きを実行する手順の一つとして使われているときには、MT が “Please bring your identification”あるいは “Bring your identification”と訳してくれると嬉しいですし、表題であるならば “To bring identification”等と訳すことが望まれます。

ここで、手順の一つとして使われているという DITA が提供する情報をもとに、

身分証明書を持参しなさい。

に書き換えると、Google translate の出力は、

Please bring your identification.

となり、手順の一つとして完全に意味が通じるようになります。一方、表題としては、この表現を「身分証明書の持参」とすると出力は

Bring identification

となり、まあそれなりに意味が通じるようになります。

ところで、手順の一つを表現する日本語としては、「身分証明書を持参する。」が普通です。従って、前者の例では、日本語文書の側は言い換ええないものをそのまま、MT の入力としては言い換えられたものを用いるのが最もよいことになります。

ここでは、二つの異なる場での「制限言語」の適用が問題になっています。

- 第一に、言い換えあるいは制限言語の適用は、その表現がどのような機能を担うかによって異なっていること。とりわけそれは、異なる言語への変換において明示的になること。
- 第二に、起点言語としての日本語の可読性を維持するために言い換えあるいは制限言語の適用が有効である場合と、翻訳のために有効である場合、その両方に有効である場合、一方が他方を阻害する場合がありうること。

実は、上で、私たちが設定した新しい制限言語の属性を「規則」ではなく、「コントロールすべき表現」と

したのは、これらの点とも関係しています。言い換え
ると MT の出力は良くなるけれども日本語では元の表
現を維持する方が望ましい場合には裏で言い換え規則
を自動的に適用することが適切で、私たちは現在、そ
の規則も設計しているところです。その際に、文の機
能要素に応じた自然な範囲での日本語表現が制約のも
とに書かれていれば背後での言い換えも定式化しやす
いというメリットがあります。

ここまで話を進めると、文書執筆における制限言語
の適用と、背後で MT の出力を改善するために適用す
る言い換えは、作業／処理ステップとしては分かれる
ことになります。しかしながら、それは見かけ上のこ
とで、理論的な枠組みとしての、具体的なテキスト／
文書の内容を基盤として「制限言語」あるいは「言い
換え」を考える、ただし、技術的な「制限言語」「言
い換え」の適用は文以下を基本単位とするのでその操
作をテキスト／文書に結びつけるために文書の機能要
素を具体的に絞った文書タイプにおいて定義し活用す
る、という点は背後で一貫しています。

5 テキスト／文書生成の環境と場

話が中途半端に技術的になりました。「言い換え」や
「制限言語」が、「じっさいに述べられた」あるいは「述
べられ」つつある「言語的まとまり」からなる、つね
に有限で現働的に限定された言語表現を対象とする、
という理論的な把握は、もうひとつ、応用においてと
ても重要な視点を含んでいます。すなわち、

「言い換え」や「制限言語」をめぐる処理の
適用は、「なんでもありうる」潜在的に無限の
言語表現を想定して、このような技術がある
から、可能性としてはこのようなときに使える
、誰かが使ってくれるだろう、というオー
プンなかたちで提供されるべきものでは—そ
もその理論的位置づけとして—なく、つね
に有限で限定された言語表現を対象とし、そ
れが「述べられ」つつある場を現実のもの
として設定し、そこに関与／貢献するかたち
で提供されなくてはならない、従って、提供
の環境の中で要素技術は考えられなくてはな
らない、

という点です。

そこで私たちの研究では、第 4 節で述べた技術を
具体的に適用しつつ手続き文書の作成ができるような
オーサリング環境の開発も同時に進めています。現在
は、直接、特定の自治体と共同して作業は行なってい

ませんが、それについてもまもなく開始する予定で技
術側の準備を進めているところです。

謝辞

本研究の一部は、KDDI 財団 2014-2016 年度調査・
研究助成「自治体文書の多言語化支援システムの開発」
の助成を受けて行われました。また、共同研究を進め
るにあたり、日本学術振興会外国人招へい研究者（短
期）2014 年度事業の助成を受けています。

文献等

- [1] Foucault, M. (1968) “Réponse au Cercle d'épistémologie”, *Cahiers pour l'analyse*, 9: *Généalogie des sciences*, pp. 9–40. [石田英敬訳「科学の考古学について〈認識論サークル〉への回答」『ミシェル・フーコー思考集成 III 1968–1970』(東京：筑摩書房)]
- [2] 井佐原均ほか. (2012) 「企業の多言語情報発信を支援する取り組み：国際化をにらんだ産業文書の効率の作成へ向けて」言語処理学会第 18 回年次大会, pp.369-372.
- [3] Hartley, A. F. and Paris, C. (1997) “Multilingual document production: from support for translating to support for authoring,” *Machine Translation - Special Issue on New Tools for Human Translators*, 12, pp. 109–128.
- [4] Hartley, A. (2010) “Enabling multilingual applications of 'controlled language': the DITA framework,” *Asia-Pacific Association for Machine Translation Journal*, 48, pp. 15–18.
- [5] Carey, L., Schlotfeldt, M. and Bellamy, J. (2012) *DITA Best Practices: A Roadmap for Writing, Editing, and Architecting in DITA*, IBM Press.
- [6] 宮田玲ほか. (2015) 「機械翻訳の活用を見据えた文書構造と言語表現の対応づけ—自治体手続き型文書を対象とした予備的報告—」言語処理学会第 21 回年次大会, 予定.
- [7] Tatsumi, M. et al. (2013) “Towards acceptable quality machine translation without post-editing for municipal websites: an evaluation of Japanese controlled language rules,” *MT Summit XIV: QTLaunch-Pad Workshop on Human-Centric Machine Translation and Evaluation*, 2013.

パネル討論 3
日本法令の英訳支援

外山 勝彦
名古屋大学 教授

日本法令の英訳支援

2015. 2.24

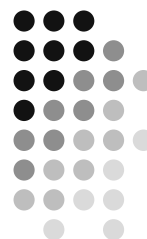
第6回産業日本語研究会・シンポジウム

名古屋大学
情報基盤センター／大学院情報科学研究科

外山 勝彦



名古屋大学大学院法学研究科附属法情報研究センター（JaLII）との
共同研究



日本法情報の国際的発信・共有

社会・経済のグローバル化

- 日本社会（法制度）の透明性確保
 - 国際取引の円滑化
 - 対日投資の促進
- 法制度の国際的協調・国際標準確立
 - 法情報の共有
 - 日本法の反映
- 発展途上国・体制移行国への法整備支援
 - 日本法に関する情報提供・教育
- 日本社会・文化の理解促進
 - 在留外国人への情報提供



日本政府による法令英訳の実施



- 内閣官房 司法制度改革推進室
法令外国語訳・実施推進検討会議
 - 2006年3月23日 「最終報告」(政府が取り組むべき施策)
 - 法令用語標準対訳辞書の策定
 - 重要法令約200本の翻訳実施(3年以内)
 - 翻訳法令を利用しやすい環境の整備(ウェブ公開)
 - 継続的体制(2009年度以降)の整備
 - 2006年4月 暫定公開開始
 - 「法令用語日英標準対訳辞書」Ver.1.0: 日本語見出し語3,315語
 - 英訳法令 6本: 民法(一部), 個人情報保護法など
- 法務省 大臣官房 司法法制部
 - 2009年4月1日 移管
 - 同 日本法令外国語訳データベース(JLT)運用開始

オブザーバ参加

構築支援

設計・開発請負

保守は民間業者が担当

法令用語日英標準対訳辞書(PDF版)



施行期日 (しこうきじつ)
effective date

施行規則 (しこうきそく)
Implementation Ordinance

時効の中断 (じこうのちゅうだん)
interruption of prescription
【注】 刑事には時効の中断と言う概念はなし

注釈

時効の停止 (じこうのていし)
(1) suspension of prescription [一般的な場合]
(2) tolling of statutory limitation [刑事の場合]

訳し分け

施行法 (しこうほう)
Act for Enforcement
【例】 刑法施行法 Act for Enforcement of the Penal Code

例文

自己株式 (じこかぶしき)
(1) own share
(2) treasury share

法令用語日英標準対訳辞書



	Ver. 1.0 (2006. 3.)	Ver. 7.0 (2012. 4.)	Ver. 9.0 (2014. 6.)
日本語見出し語 (異なり数)	3,315	4,482	3,671
英訳語	3,974	5,844	4,624
平均英訳語数	1.19	1.30	1.26

多様な形式で提供: PDF, CSV, XML

Ver.1.0: UTX化 (アジア太平洋機械翻訳協会, AAMT)

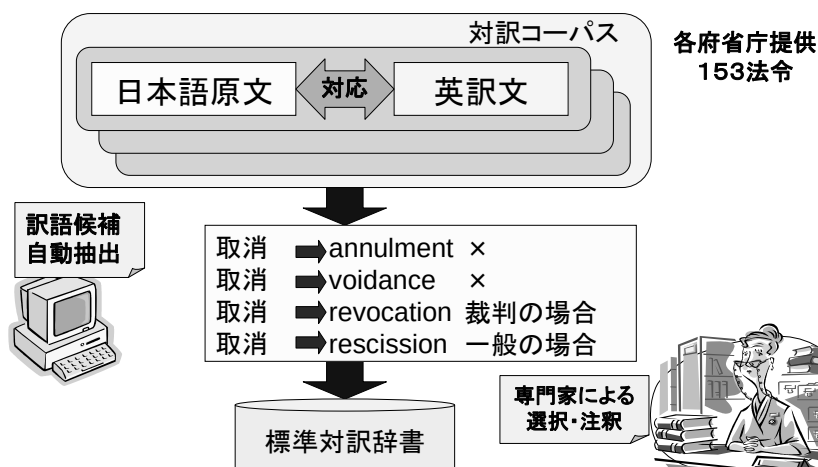
Ver.4.0: 「みんなの翻訳」で利用

5

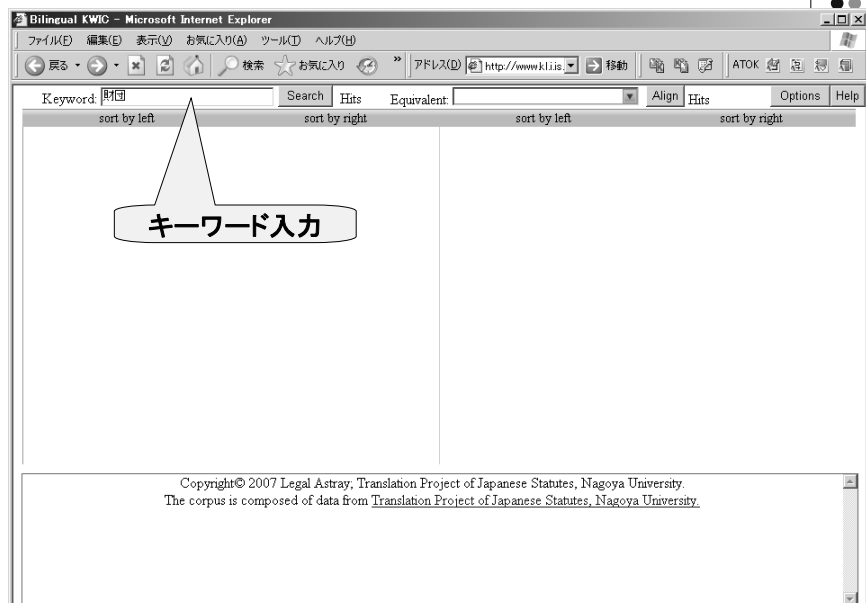
対訳辞書構築の計算機支援



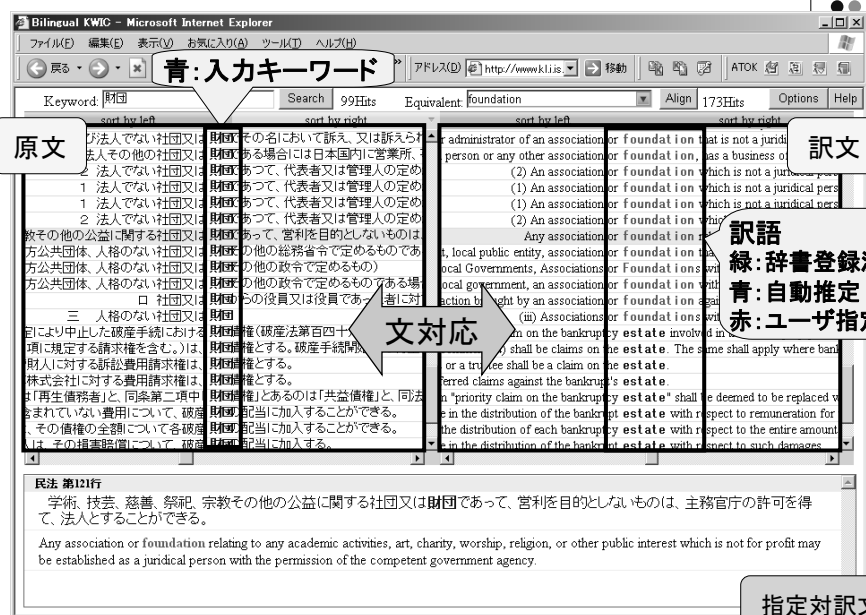
文対応付き対訳コーパスからのワード・アライメント



対訳表現抽出支援ツールBilingual KWIC®



対訳表現抽出支援ツールBilingual KWIC®



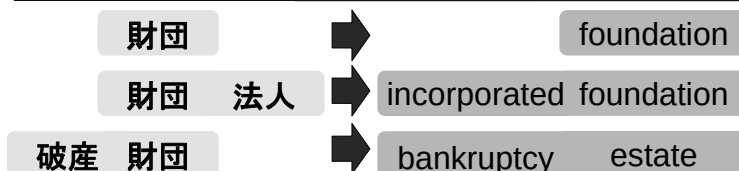
Bilingual KWIC® における文脈の活用



入力キーワード： 財団

検索結果：

団又は	財団	の普通	ther association or	foundation	shall be determine
団又は	財団	からの	y an association or	foundation	against its officer
による	財団	法人へ	An incorporated	foundation	in accordance wit
従い、	財団	法人に	is an incorporated	foundation	, by obtaining appr
、破産	財団	に属す	s to the bankruptcy	estate	and any other matt
各破産	財団	の配当	of each bankruptcy	estate	with respect to the



9

対訳表現抽出支援ツールBilingual KWIC®



- 特長
 - 文脈の表示により訳語獲得を支援
 - 派生語に対する訳語の獲得
 - ワード・アライメントにおける抽出誤りの修正
 - 訳し分け状況の獲得
 - 対訳辞書なしで利用可能
 - 対訳辞書の組込みは可能
 - 言語に非依存
 - 任意の2言語間に適用可能

10

日本法令外国語訳データベースシステム (JLT) (<http://www.japaneselawtranslation.go.jp/>)



11

JLTの目的

- 日本法令の翻訳の一元管理・公開
 - 法令とその翻訳のデータベース構築
 - インターネット上での情報発信
- 翻訳品質(訳語統一性・信頼性)の確保
 - 翻訳法令データと「法令用語日英標準対訳辞書」の公開・無償配布
 - 標準対訳辞書を使用した翻訳を要請・支援

12

法令検索結果(日英交互表示)

日本法令外国語訳データベースシステム - 法令本文表示 - 不当景品類及び不当表示防止法 - Microsoft Internet Explorer

Japanese Law Translation

検索キーワードのハイライト表示: ☐ 強制 (and検索)

不当景品類及び不当表示防止法
法令番号: 昭和三十七年法律第三十四号 改正: 平成十七年法律第三十五号 辞書バージョン: 1.0 翻訳日: 平成21年4月1日
Law number: Act No. 134 of 1962 Amendment: Act No. 35 of 2005 Dictionary Ver: 1.0 Translation date: April 1, 2009

目次

- 第一条 (目的)
- 第二条 (定義)
- 第三条 (景品類の制限及び禁止)
- 第四条 (不当表示の禁止)

不当景品類及び不当表示防止法
Act against Unjustifiable Premiums and Misleading Representations

(目的)
(Purpose)

第一條 この法律は、商品及び役物の取引に關連する不当な景品類及び表示による顧客の誘引を防止するた...
Article 1 The purpose of this Act is, in order to prevent inducement of customers by means of unjustifiable premiums and misleading representations in connection with transactions of goods and services, by establishing special provisions of the Act on Prohibition of Private Monopolization and Maintenance of Fair Trade (Act No. 54 of 1947), to ensure fair competition, and thereby to protect the interests of general consumers.

(定義)
(Definitions)

第二條 この法律で「景品類」とは、顧客を誘引するための手段...
Article 2 The term "premiums" as used in this Act means any...

対訳辞書の内容のポップアップ

他法令へのリンク

表示切替ボタン

法令データのダウンロード PDF, Word, TXT (Shift-JIS, UTF-8), XML

法令検索結果(日英対照表)

日本法令外国語訳データベースシステム - 法令本文表示 - 不当景品類及び不当表示防止法 - Microsoft Internet Explorer

Japanese Law Translation

検索キーワードのハイライト表示: ☐ 強制 (and検索)

不当景品類及び不当表示防止法
法令番号: 昭和三十七年法律第三十四号 改正: 平成十七年法律第三十五号 辞書バージョン: 1.0 翻訳日: 平成21年4月1日
Law number: Act No. 134 of 1962 Amendment: Act No. 35 of 2005 Dictionary Ver: 1.0 Translation date: April 1, 2009

目次

- 第一条 (目的)
- 第二条 (定義)
- 第三条 (景品類の制限及び禁止)
- 第四条 (不当表示の禁止)

不当景品類及び不当表示防止法
Act against Unjustifiable Premiums and Misleading Representations

昭和三十七年五月十五日法律第三十四号	Act No. 134 of May 15, 1962
(目的)	(Purpose)
第一條 この法律は、商品及び役物の取引に關連する不当な景品類及び表示による顧客の誘引を防止するため、私的独占の禁止及び公正取引の確保に關する法律(昭和二十二年法律第五十四号)の特例を定めることにより、公正な競争を確保し、もつて一般消費者の利益を保護することを目的とする。	Article 1 The purpose of this Act is, in order to prevent inducement of customers by means of unjustifiable premiums and misleading representations in connection with transactions of goods and services, by establishing special provisions of the <u>Act on Prohibition of Private Monopolization and Maintenance of Fair Trade</u> (Act No. 54 of 1947), to ensure fair competition, and thereby to protect the interests of general consumers.
(定義)	(Definitions)
第二條 この法律で「景品類」とは、顧客を誘引するための手段...	Article 2 The term "premiums" as used in this Act means any...

対訳辞書の内容のポップアップ

他法令へのリンク

表示切替ボタン

法令データのダウンロード PDF, Word, TXT (Shift-JIS, UTF-8), XML

JLT収録の日英対訳法令データ



提供法令数		現行法令数 (総務省調べ)
2009. 4. 1	2015.1.14	2015.1. 1
166	434* (対訳261,396文)	8,079

*他に、法令改正に伴う修正109

15

JLTの問題点(例1): 翻訳の揺れ



一般社団法人及び一般財団法人に関する法律

Act Concerning General Corporations and General Foundations
Act on General Associations and Foundations
Act on General Associations and Incorporated Foundations
Act on General Incorporated Association and General Incorporated Foundation
Act on General Incorporated Associations and General Incorporated Foundations
General Incorporated Associations/Foundations Act

6種類

この法律は、公布の日から施行する。

This Act shall be enforced from the date of promulgation.
This Act shall come into effect as from the date of promulgation.
This Act shall come into effect as of the date of promulgation;
This Act shall come into effect as of the day of its promulgation.
This Act shall come into effect on the day of promulgation.
This Act shall come into force as from the date of its promulgation.
This Act shall come into force as from the day of its promulgation.
This Act shall come into force as of the date of its promulgation.
This Act shall come into force as of the day of promulgation.
This Act shall come into force from the day of promulgation.
This Act shall enter into force on the day of its promulgation,

30種類

16

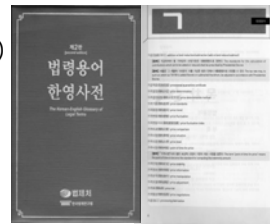
JLTの問題点(1-1)



- 訳語が十分に統一されていない
 - 標準対訳辞書見出し語数(3,671語)は全く不十分
 - 法律学上で関心がある見出し語だけ??
 - 訳し分けなどの注記が十分でない

【参考】

- 「法令用語韓英辞書第2版」(韓国法制処編, 2009)
 - 本編だけで9,306語収録
- IATE (InterActive Terminology for Europe)
<http://iate.europa.eu/>
 - EUの多言語ターミノロジー
 - EU法令等の多言語翻訳のための基盤
 - 約800万語(26言語)収録
 - 用語の定義, 出典, 関連語などの言語情報



17

JLTの問題点(1-2)



- 訳語が十分に統一されていない
 - 対訳文が言語資源として一括提供されていない

【参考】

- EU翻訳総局 DGT-TM
 - 翻訳メモリを無償一括提供
 - 英語対訳だけで590万ユニット
<https://open-data.europa.eu/en/data/dataset/dgt-translation-memory/>



標準対訳辞書の拡充
翻訳メモリのオープンデータ化
翻訳資産としての整備, 共有



18

JLTの問題点(例2): 法令用語の慣習



- 「及び」「並びに」

原油、天然ガス及び石炭並びにこれらから製造される燃料

標準対訳辞書記載の用例

A及びB、C、並びにD ➡ A and B, C, and D

- 「直ちに」「速やかに」「遅滞なく」

...したときは、直ちに、その旨を公告しなければならない。

短	直ちに	<	速やかに	<	遅滞なく	長
	↓		↓		↓	
標準対訳辞書	immediately	< ?	promptly	< ?	without delay	

19

JLTの問題点(2)



- 翻訳者は法律学や法制執務における慣習に詳しくない
 - 法制執務： 法令の起草・改廃に関する実務・文書技術
法令のスタイルガイド
- 【参考】 EUでは “lawyer-linguist” が法令翻訳に責任
- 翻訳法令の読者は日本の法制執務における慣習をほとんど知らない



日本法令のスタイルガイド・翻訳ルールの解説書の作成
 ・ 翻訳者向け
 ・ 海外向け
 法令専門の翻訳者「法令翻訳士」の制度構築・養成

JLTの問題点(3)



- 翻訳の作業体制・フローに問題がある
 - 翻訳の計画立案から公開までに時間が掛かる(平均約3年)
 - 翻訳本数が少ない(現行法令の約5.4%)
 - 法令改正に伴う翻訳の修正は十分に対応できていない
 - 翻訳は所管府省庁の責任でそれぞれ実施
 - 法律専門家による翻訳品質チェックの結果を翻訳者へフィードバックできていない
 - 既翻訳の再点検は皆無
- 日本語原文に誤りがある
 - 正確な日本語法令テキストの入手は容易でない



法令翻訳作業の一元化
日本語法令データ管理の一元化・ワンストップサービス提供

21

翻訳技術・体制以前の問題



そもそも日本の法令は分かりやすいのか？

【参考】

- 米国メイン州など： plain English で書く運動
- 韓国： 分かりやすい法令審査事業(2006～)
 - 基本的な教育を受けた普通の水準の国民が理解できるように
 - 難しい漢字語や日本式用語・外来語、過度の短縮語の置き換え
 - 不要な表現や複雑な表現の整理
 - 「分かりやすい法令審査基準」作成
 - 再整備された法律案1,014件(799件が成立)(～2012)
- 分かりにくい原因は何か
- 分かりやすくするにはどのようにしたらよいか
- 分かりやすくなったかどうかをどのように評価するか
- 分かりやすくすることを国会や政府が受け入れるのか

22

パネル討論 4

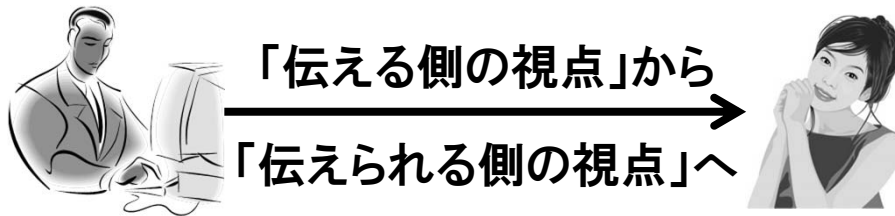
「伝える側の視点」から「伝えられる側の視点」へ

勝田 豊彦

一般財団法人テクニカルコミュニケーター協会

／ハイテクノロジー・コミュニケーションズ株式会社

第6回産業日本語研究会・シンポジウム



2015.2.24
ハイテクノロジー・コミュニケーションズ株式会社
勝田豊彦

High Technology Communications Inc.

1

どんな形状の立体？

ある方向から見ると

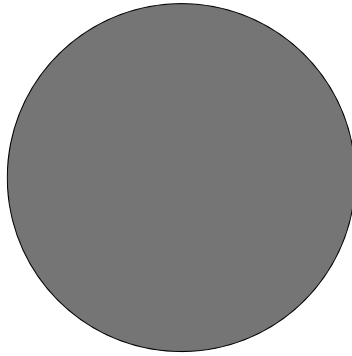


High Technology Communications Inc.

2

どんな形状の立体？

別の方向から見ると

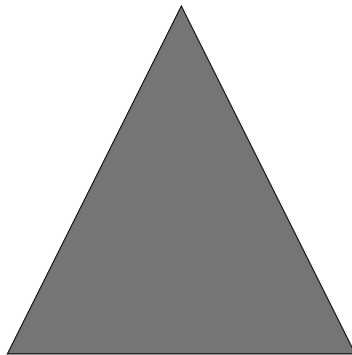


High Technology Communications Inc.

3

どんな形状の立体？

さらに別の方向から見ると



High Technology Communications Inc.

4

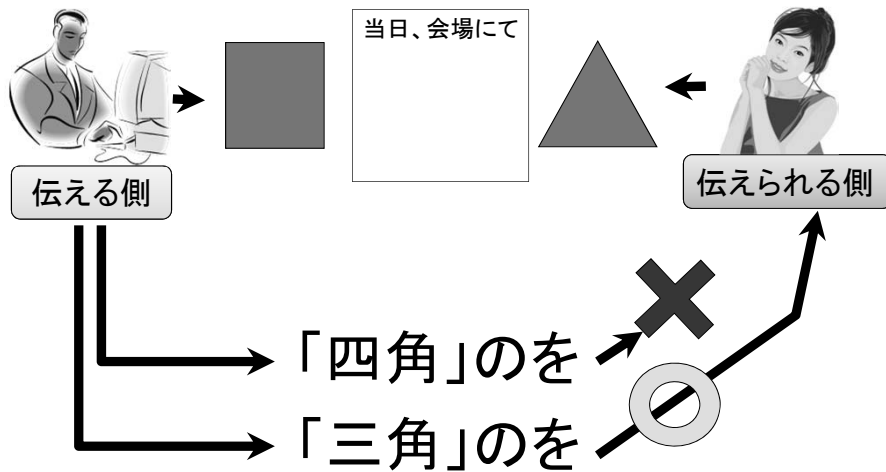
どんな形状の立体？

イメージしてみてください

こんな立体です

当日、会場にて

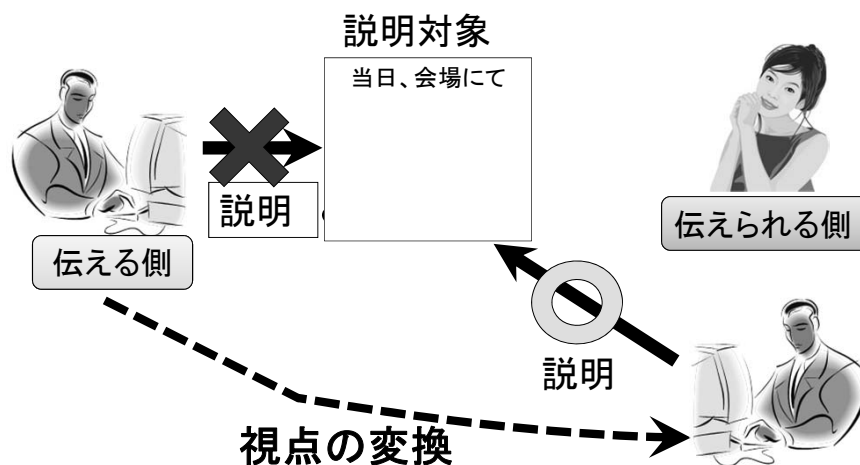
「伝える側の視点」では伝わらない



High Technology Communications Inc.

7

「伝えられる側の視点」への変換が必要

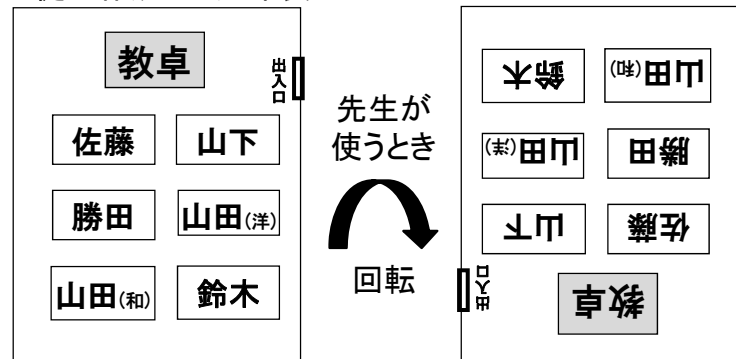


High Technology Communications Inc.

8

視点の変換 → 向きの変換

生徒が作成した座席表

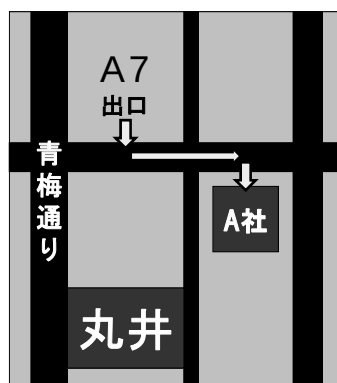


High Technology Communications Inc.

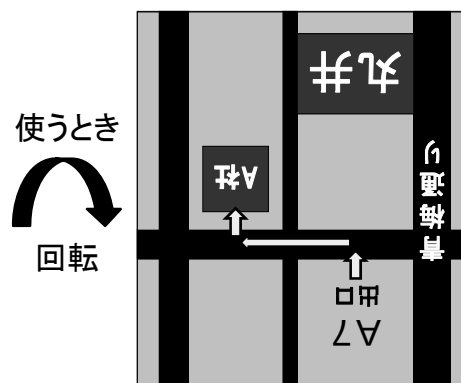
9

視点の変換 → 向きの変換

A



B

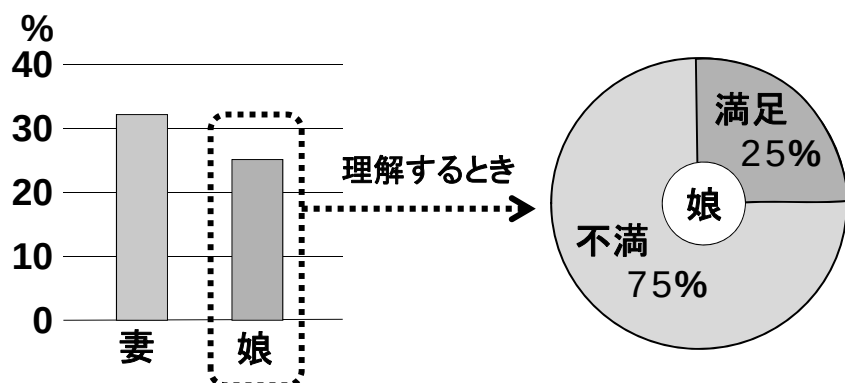


High Technology Communications Inc.

10

視点の変換 → イメージの変換

父親のファッションへの満足度は？



11

視点の変換 → 内容と用語の変換



伝える側(開発者)の視点

本機は、緊急警報放送を受信すると、

ノイズ 0.2秒以内に電源が入り、その0.1秒以内に

用語 緊急アナウンスモード に入ります。



伝えられる側(利用者)の視点

本機は、緊急警報放送を受信すると、

電源が自動的に入り、大音量で警報を発します。

伝えられる側には変換不能



第2章 MS-DOSのコマンド

ADDDRV	16	CLS	36
APPEND	18	COMMAND	37
ASSIGN	21	COPY	39
ATTRIB	24	COPYA	44
BACKUP	26	CTTY	47
BREAK	29	CUSTOM	49
CHDIR(CD)	30	DATE	52
CHGEV	31	DEBUG	54
CHKDSK	32	DEL(ERASE)	55
CHKENV	34	DELDIV	57
CHKFIL	35	DICM	58

High Technology Communications Inc.

13

視点の変換 → 説明目的の変換



キーの説明

[Esc]キー:処理を中断する
[Enter]キー:次の処理に進む



使い方の説明

処理を中断する: [Esc]キー
次の処理に進む: [Enter]キー

High Technology Communications Inc.

14

視点の変換 → 思考(or説明)順序の変換

電卓の例

性能や特長を知りたい場面では

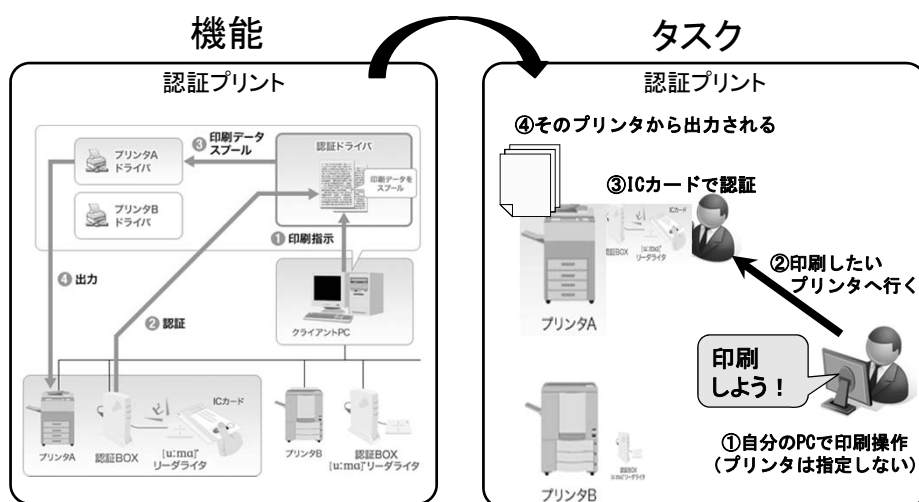
計算結果が16桁までの計算ができる。
結果が16桁を超えるとエラーとなり、「E」と表示される。

発生現象への対処を知りたい場面では

「E」と表示されたときは、計算結果が大きすぎて、
エラーとなっている。

結果が16桁を超える計算はできない。

視点の変換 → 「機能」から「タスク」への変換



視点の変換 → マニュアルの目次構成

①～③に適切な語句を入れてみてください。

【ファックスのマニュアル：3種類の受信方法の設定による違い】

- ・手動受信：呼出音が鳴ったら、受話器を取り、相手がFAXなら受信操作をする。
- ・自動切替受信：相手が電話なら呼出音が鳴り、相手がFAXなら自動受信する。
- ・FAX専用受信：FAX受信だけができる。電話は受けられない。

「機能」での項目立て

- 準備
 - － 回線と電源を接続する
 - － 用紙を入れる
- 電話をかける
- 電話を受ける
- FAXを送る
- FAXを受ける
 - － 受信方法を設定する
 - － 手動受信をする
 - － 自動切替受信をする
 - － FAX専用受信をする

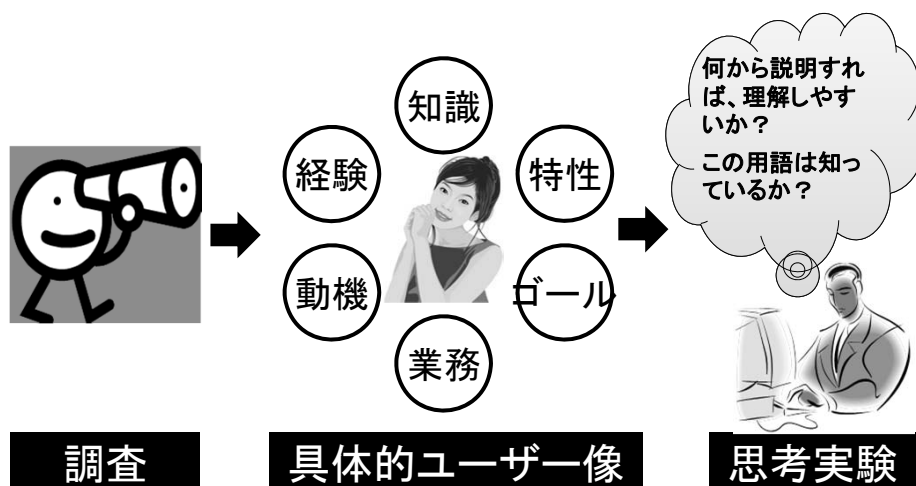
「タスク」での項目立て

- 準備
 - 回線と電源を接続する
 - 用紙を入れる
 - ①
- 電話をかける
- FAXを送る
- ② を受ける
 - ③ たら
 - FAX用紙が出てきたら

High Technology Communications Inc.

17

「伝えられる側」の分析が重要



High Technology Communications Inc.

18

不許複製 禁無断転載

高度言語情報融合フォーラム (ALAGIN)

言語処理学会

一般財団法人日本特許情報機構 (Japio)

発行日 平成27年2月24日

発行者 高度言語融合フォーラム (ALAGIN) 内

産業日本語研究会・シンポジウム事務局

本シンポジウムに関する問合せ先

産業日本語研究会・シンポジウム事務局

email: info@alagin.jp