

読み書きの支援/診断技術と読み書き

乾 健太郎

東北大学 情報科学研究科
理化学研究所 AIPセンター

ライティング
~~読み書き~~の支援/診断技術と←
~~読み書き~~
ライティングへの影響

乾 健太郎

東北大学 情報科学研究科
理化学研究所 AIPセンター

言語処理技術による執筆の支援

文章生成については過去にも紹介があるので、ここでは主に**校正・添削**、**共同執筆**の話をしていきます

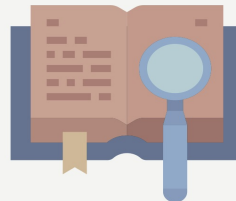
文章生成 (draft)



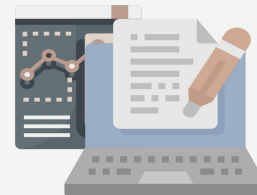
Machine translation



Summarization



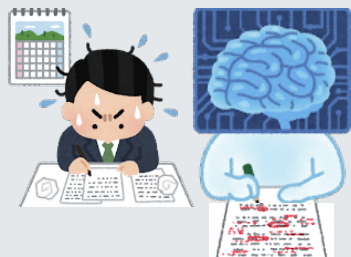
Data-to-text



https://cs.stanford.edu/~minalee/pdf/cornell-writing_with_language_models-slides.pdf

隅田英一郎「2021年の翻訳バンク・ニューラル翻訳」第12回産業日本語シンポジウム、他多数
岡崎直観「自然言語処理によるニュース記事執筆の現状と課題」第11回産業日本語シンポジウム

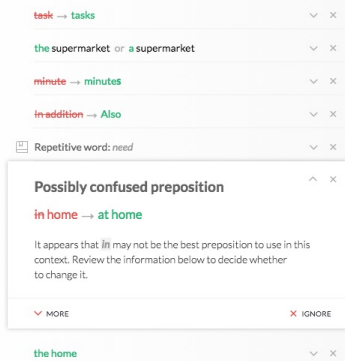
校正・添削
(edit, diagnose)



スペル・文法・スタイルのチェック、言い換え

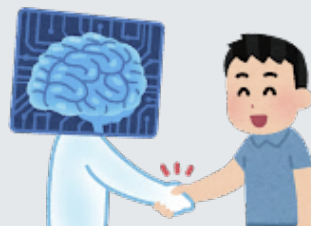
Type your title

To begin with, food that is easier to prepare allows people to save time. If people want to cook food by themselves, they have a lot of task to do. First, people have to go to supermarkets and buy food ingredients. Going to supermarket and choosing preferable food ingredients need at least thirty minute to one hour. In addition, cooking food need more time. What is more, if they cook food in home, people have to wash their plates or pans after enjoy dishes. On the other hand, if people buy prepared food or frozen food, all they have to do is just heat the food in their microwave. I think people need only five minutes to prepare such convenient foods except for the time to buy them.

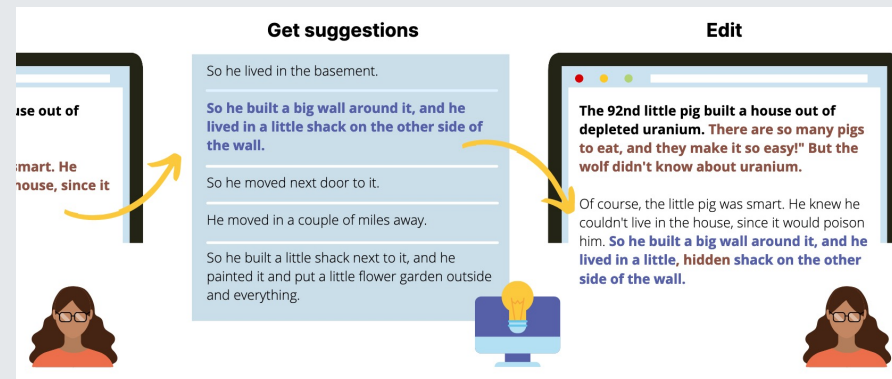


<https://www.path-to-success.net/grammarly>

共同執筆
(co-author)



テキストの自動補完、挿入文の提案



<https://arxiv.org/pdf/2201.06796.pdf>

現在の技術・サービスは**英語**が中心だが、要領は日本語も同じ
日本語の執筆支援サービスも近い将来に出てくると予想される

言語処理技術による執筆の支援

「AIによる知的執筆支援」への関心が高まってきている



In2Writing

The First Workshop on Intelligent and Interactive Writing Assistants

The purpose of this interdisciplinary workshop is to bring together researchers from the natural language processing (NLP) and human-computer interaction (HCI) communities as well as industry practitioners and professional writers to discuss innovations in building, improving, and evaluating intelligent and interactive writing assistants. We plan to alternate our workshop venue between an NLP conference and a HCI conference every year to facilitate collaboration.

The first 100 participants get a free premium subscription to [Grammarly](#) and [Wordtune](#).

This year the workshop will be held at [ACL 2022](#) in Dublin, Ireland on the 26th of May, 2022.

<https://in2writing.glitch.me/>

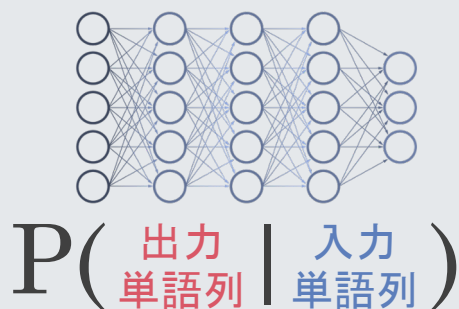
ニューラル翻訳・ニューラル言語モデルによる言語生成

データとモデルの大規模化で言語生成がだいぶ**流暢**になってきた

入力単語列（文脈・条件）

Sendai is the capital city of Miyagi
Prefecture, the largest city in the
Tohoku region.

ニューラル翻訳モデル



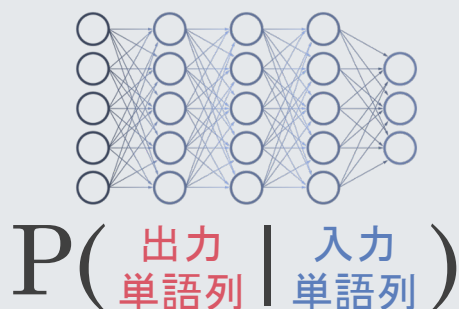
出力単語列

仙台は、宮城県の県庁所在地で、
東北地方最大の都市です。

入力単語列（文脈・条件）

仙台は、宮城県の県庁_____

ニューラル言語モデル



出力単語列

所在地で、東北地方最大の都市です。

ニューラル翻訳・ニューラル言語モデルによる言語生成

データとモデルの大規模化で言語生成がだいぶ**流暢**になってきた

「人間が作成した文章か機械が生成した文章か」
を見分けるのは人間にも難しい

Once upon a time, there lived a pirate. He was the sort of pirate who would rather spend his time chasing away the sharks swimming around his ship than sail to foreign ports in search of booty. He was a good pirate, a noble pirate, an honest pirate. He was a pirate who would rather be at home with his wife and son than out on a ship in the middle of the ocean.

rambles in a way
that make sense.

there were **personal
description[s]** a machine
wouldn't understand, [like]
wanting to be home with
his wife and son.

too **natural** to be AI



A **human** wrote this

seems to have **run on
thoughts**.

no pirate has a home with
his wife and kids unless
they're on the ship with him.
That is **utterly
unbelievable**

repeating itself lots



A **machine** wrote this

- 小説、ニュース、レシピ
- 訓練しても限定的(～55%)
- 判断理由もバラバラ
- (言語生成の人手評価も工夫を)

言語処理技術による執筆の支援

文章生成については過去にも紹介があるので、ここでは主に**校正・添削**、**共同執筆**の話をしていきます

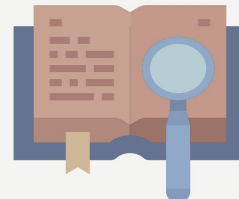
文章生成 (draft)



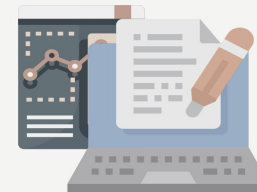
Machine translation



Summarization



Data-to-text



https://cs.stanford.edu/~minalee/pdf/cornell-writing_with_language_models-slides.pdf

隅田英一郎「2021年の翻訳バンク・ニューラル翻訳」第12回産業日本語シンポジウム、他多数
岡崎直観「自然言語処理によるニュース記事執筆の現状と課題」第11回産業日本語シンポジウム

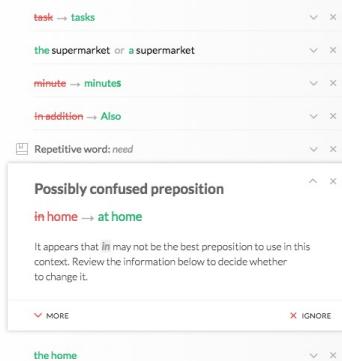
校正・添削 (edit, diagnose)



スペル・文法・スタイルのチェック、言い換え

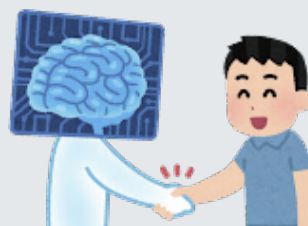
Type your title

To begin with, food that is easier to prepare allows people to save time. If people want to cook food by themselves, they have a lot of task to do. First, people have to go to supermarkets and buy food ingredients. Going to supermarket and choosing preferable food ingredients need at least thirty minute to one hour. In addition, cooking food need more time. What is more, if they cook food in home, people have to wash their plates or pans after enjoy dishes. On the other hand, if people buy prepared food or frozen food, all they have to do is just heat the food in their microwave. I think people need only five minutes to prepare such convenient foods except for the time to buy them.

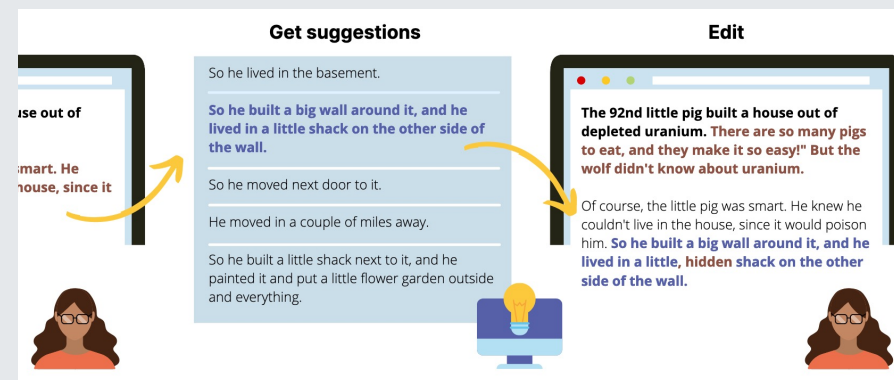


<https://www.path-to-success.net/grammarly>

共同執筆 (co-author)



テキストの自動補完、挿入文の提案



<https://arxiv.org/pdf/2201.06796.pdf>

現在の技術・サービスは**英語**が中心だが、要領は日本語も同じ
日本語の執筆支援サービスも近い将来に出てくると予想される

スペルチェック・文法チェック

Type your title

To begin with, food that is easier to prepare allows people to save time. If people wants to cook food by themselves, they have a lot of task to do. First, people have to go to supermarkets and buy food ingredients. Going to supermarket and choosing preferable food ingredients need at least thirty minute to one hour. In addition, cooking food need more time. What is more, if they cook food in home, people have to wash their plates or pans after enjoy dishes. On the other hand, if people buy prepared food or frozen food, all they have to do is just heat the food in their microwave. I think people need only five minutes to prepare such convenient foods except for the time to buy them.



GENERAL (DEFAULT)

128 WORDS

5 CRITICAL ISSUES 3 ADVANCED ISSUES

SCORE: 45

task → tasks

the supermarket or a supermarket

minute → minutes

In addition → Also

Repetitive word: need

Possibly confused preposition

in home → at home

It appears that in may not be the best preposition to use in this context. Review the information below to decide whether to change it.

MORE

IGNORE

the home

enjoy → enjoying

翻訳としての文法誤り訂正

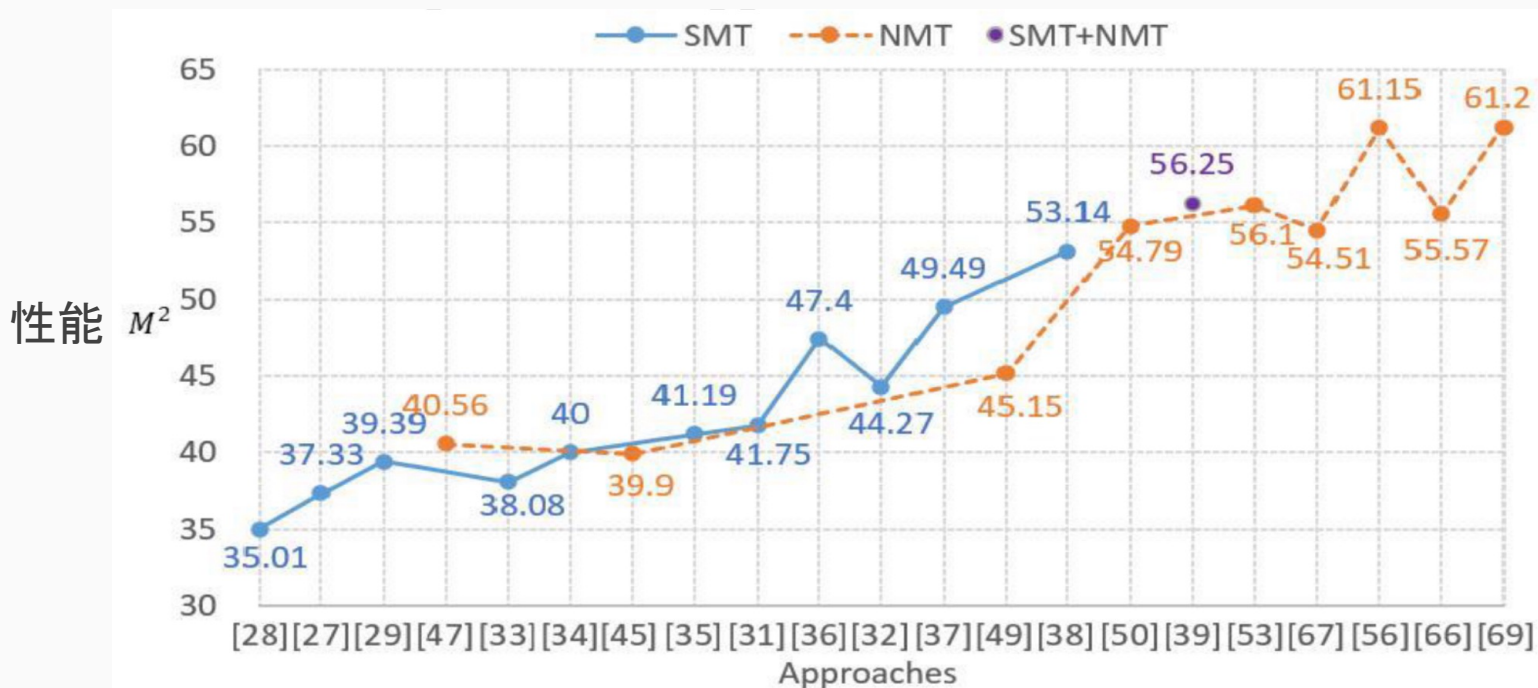
The system is design to help student.

文法誤り訂正

誤りのある文から
誤りのない文への「翻訳」

The system is designed to help students.

ニューラル翻訳 (NMT) でさらに性能向上



Wang et al. A comprehensive survey of grammar error correction. arXiv, 2020.

書き換えによる執筆支援

下手な文からきれいな文への翻訳

(Ito, Kuribayashi+ INLG 2019, EMNLP 2020, NLP 2021)



稚拙・不完全な文

Our model ??? to the SemEval task and ??? good results.

生成

論文で使える綺麗な文

We applied our model to the SemEval task and obtained good results.

We demonstrate that our model is able to achieve good results on the SemEval task and outperforms other baselines.



書き換えによる執筆支援

下手な文からきれいな文への翻訳



Langsmith

<https://editor.langsmith.co.jp/>

2021年5月サービスイン

The screenshot displays the Langsmith Editor web interface. At the top, it says 'Langsmith Editor' and 'Temporarily saved to browser local'. Below this is a menu with 'File', 'Tools', and 'Help'. A toolbar shows 'Field: Natural language processing', 'Section Name: Abstract', and 'Where to Rewrite' with undo, redo, and delete icons. The main text area contains a paragraph: 'Despite the current diversity and inclusion initiatives in the academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.' The word 'barriers' is highlighted in blue. To the right of the text are several sliders for 'typicality', 'readability', 'fluency', 'coherence', and 'style', each with a corresponding icon. Below the main text, several alternative rewrites are shown, each with a 'typicality' slider. The first rewrite is highlighted in yellow: 'Despite the current diversity and inclusion initiatives in the academic community, researchers with a non-native command of English still face significant obstacles when writing papers in English.' A large red text overlay at the bottom reads: 'AIがより適切な表現を提案してくれます'.

Langsmith Editor

Temporarily saved to browser local

File Tools Help

Field: Natural language processing | Section Name: Abstract | Where to Rewrite

Langsmith Editor

1 Despite the current diversity and inclusion initiatives in the academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.

Despite the current diversity and inclusion initiatives in the academic community, researchers with a non-native command of English still face significant obstacles when writing papers in English.

Despite the current diversity and inclusion initiatives in the academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.

Despite the current diversity and inclusion initiatives in an academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.

Despite the current diversity and classification initiatives in an academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.

👍 Despite the c... English still face significant barrier when writing papers in English.

Though current diversity and inclusion initiatives are available in the academic community, researchers with a non-native command of English still face significant barriers when writing papers in English.

AIがより適切な表現を提案してくれます

書き換えによる執筆支援

下手な文からきれいな文への翻訳

(Ito, Kuribayashi+ INLG 2019, EMNLP 2020, NLP 2021)



Langsmith

稚拙・不完全な文

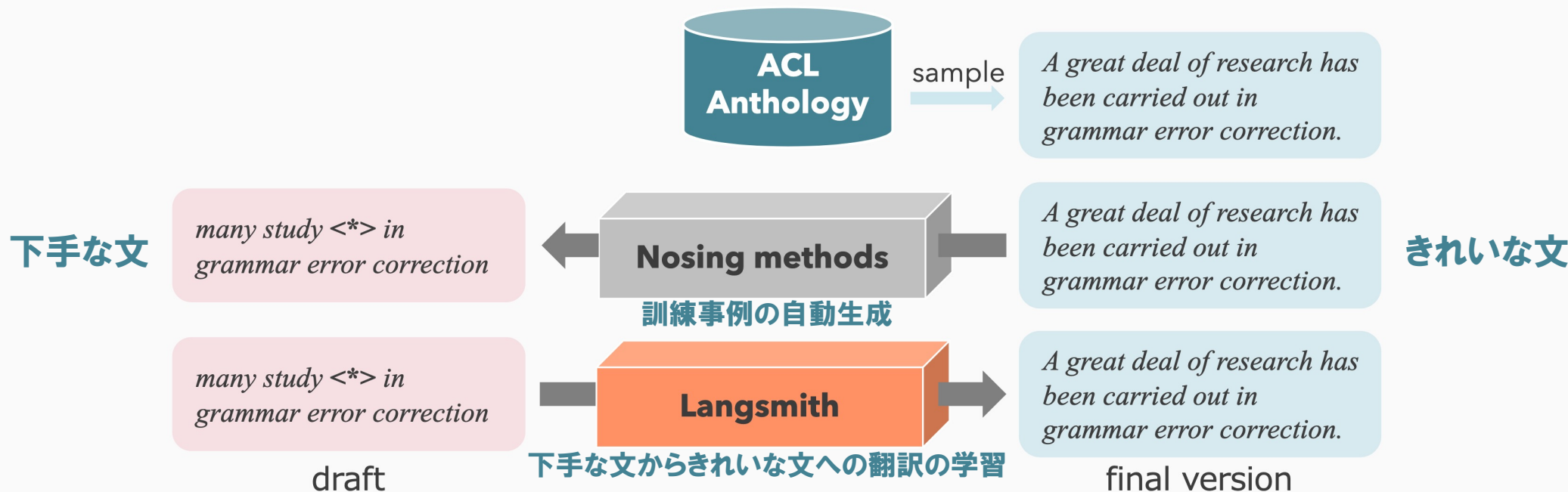
Our model ??? to the SemEval task and ??? good results.

生成

論文で使える綺麗な文

We applied our model to the SemEval task and obtained good results.

We demonstrate that our model is able to achieve good results on the SemEval task and outperforms other baselines.



文法誤り訂正 + 解説文生成

(Nagata+ NLP 2019, EMNLP 2019; Hanawa+ EMNLP 2021)

topic: Gardening

student essay:

*I **had ever been to** gardening.
Becouse my mother's hobby is
gardening. And I **beyond**
gardening club for 11 years ...*

<<have been to>>は「...に行ったことがある」を表します。「ガーデニングをする」という表現で用いられる<動詞>を調べてみましょう。

<<have been to>> is used to express ``an experience in visiting'' Use another verb for ``gardening''.

<過去完了>は、ある過去の時点を示した上で、それまでの行為・状態を表す場合に用いられます。単に「(これまでに) ...したことがある」という経験を表す場合は、<現在完了>で表します。

<Past perfect> is used with a certain point of time to talk about something that happened before that point. Use <present perfect> to talk about experiences.

<<beyond>>は<前置詞>で「...の向こうで; ...を越えて」です。ここは<動詞> belongを用いて「...に所属している」と表してみましょう。

<<beyond>> is a < preposition> meaning ``further away'' or ``outside a limit''. Use the <verb> ``belong'' instead.

初めての国際コンペ開催へ

GenChal 2022: FCG



GenChal 2022: Feedback Comment Generation for Writing Learning

Task overview

Feedback comment generation is a task where, given an input text, a system generates hints or explanatory notes that help language learners improve their writing skills.

For more details of the task, refer to [the task definition page](#) and our [task proposal paper](#).

Please [register from here](#) in order to participate in the task. Since we'll be evaluating the submissions manually, we'll limit the total number of participating systems at around 10.

Important dates

- 28 January 2022: Kick-off and training data release
- 2 May 2022: Test data release
- 9 May 2022: Generation result submission
- 25 June 2022: Evaluation result notification
- Paper submission (following INLG 2023)

Organizers

- Ryo Nagata (Konan University / JST, PRESTO)
- Artem Chernodub (Grammarly)
- Masato Hagiwara (Earth Species Project / Octanove Labs)
- Hanawa Kazuaki (RIKEN / Tohoku University)
- Masato Mita (RIKEN / Tohoku University)
- Olena Nahorna (Grammarly)

<https://t.co/1YdwpOVeZn>

言語処理技術による執筆の支援

文章生成については過去にも紹介があるので、ここでは主に**校正・添削**、**共同執筆**の話をしていきます

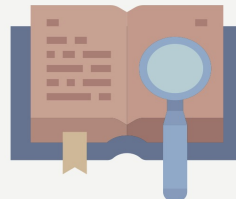
文章生成 (draft)



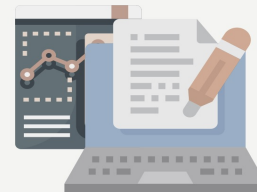
Machine translation



Summarization



Data-to-text



https://cs.stanford.edu/~minalee/pdf/cornell-writing_with_language_models-slides.pdf

隅田英一郎「2021年の翻訳バンク・ニューラル翻訳」第12回産業日本語シンポジウム、他多数
岡崎直観「自然言語処理によるニュース記事執筆の現状と課題」第11回産業日本語シンポジウム

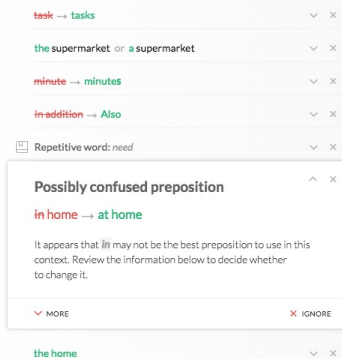
校正・添削 (edit, diagnose)



スペル・文法・スタイルのチェック、言い換え

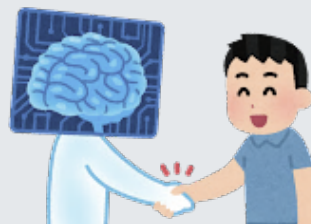
Type your title

To begin with, food that is easier to prepare allows people to save time. If people want to cook food by themselves, they have a lot of task to do. First, people have to go to supermarkets and buy food ingredients. Going to supermarket and choosing preferable food ingredients need at least thirty minute to one hour. In addition, cooking food need more time. What is more, if they cook food **in** home, people have to wash their plates or pans after enjoy dishes. On the other hand, if people buy prepared food or frozen food, all they have to do is just heat the food in their microwave. I think people need only five minutes to prepare such convenient foods except for the time to buy them.

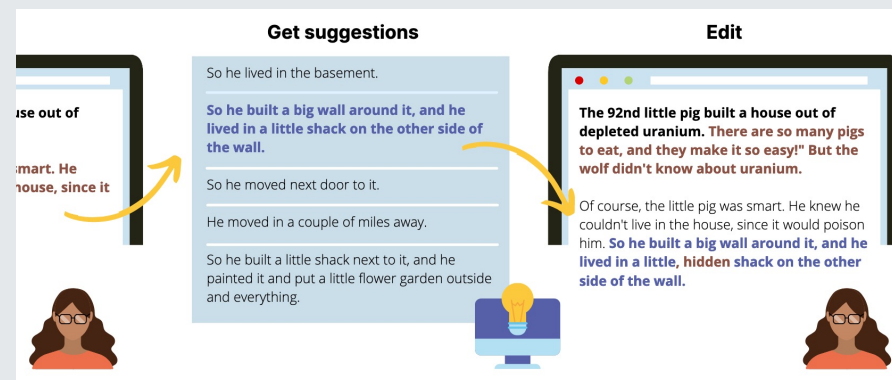


<https://www.path-to-success.net/grammarly>

共同執筆 (co-author)



テキストの自動補完、挿入文の提案



<https://arxiv.org/pdf/2201.06796.pdf>

現在の技術・サービスは**英語**が中心だが、要領は日本語も同じ
日本語の執筆支援サービスも近い将来に出てくると予想される

文の自動補完

Google: Smart Compose (2018)



文の自動補完

IGA: An Intent-Guided Authoring Assistant (Sun+ EMNLP 2021)

CONTEXT: Sarah didn't come to school today,

- + **CAUSE** homework → because she had to get her homework done.
- + **EFFECT** news → therefore there is no more big news!
- + **CONTRAST** exam → but still took the exam and had it done.
- + **DESCRIPTION** <NA> → it was an unplanned break.
- + **BIOGRAPHY** <NA> → David, the principal, believes she will miss tomorrow's class.
- + **IDIOM** <NA> → She was busy making ends meet.
- + **REPHRASE** <NA> → Sarah did not arrive at school today.

文の自動補完

IGA: An Intent-Guided Authoring Assistant (Sun+ EMNLP 2021)

訓練事例を大規模に自動合成

Fine-tuning example construction

(1) Identify **key words** and **target spans**

(2) Replace **target spans** with **tags**

(3) Concatenate **target spans** and **special tokens**

The wind blew over the farm, the rain came down and she heard ominous pings on the roof.

The wind blew over the farm, the rain came down and **she heard ominous pings on the roof**.

The wind blew over the farm, the rain came down and **<descrip>** pings **<descrip>** roof.

The wind blew over the farm, the rain came down and **<descrip>** pings **<descrip>** roof. **<sep>** she heard ominous **<answer>** on the **<answer>** **<eos>**

The wind blew over the farm, the rain came down and **<descrip>** pings **<descrip>** roof. **<sep>**

(3) Post-processing: Infill model output back to spans


(2) Post-processing: Identify tags


(1)  Fine-tuned GPT-2 generates output prefixed with input and **<sep>**


Inference time


言語モデルによるドラフトの自動生成


Rytr: “A better, 10x faster way to write”


 Rytr

 History


 Account








Select language

 English

Select tone

Enthusiastic

Choose use case

 Facebook, Google, LinkedIn Ads

Generate Ad copies for your product or brand

Product name

mug

For best results provide maximum input 3/25

Product description

avocado-shaped mug

For best results provide maximum input 18/200

Please review and comply with [Facebook ad terms](#), [Google ad terms](#), [LinkedIn ad terms](#).

3 variants

Rytr more →

← New document

B

/

U

S

H1

H2

H3

≡

≡

↶

”

✎

≡

≡

≡

📑

✂

↶

↷

Words 135 Characters 754

Avocado mug

Ad Copy: The perfect gift for avocado lovers! This avocado-shaped mug will make a great addition to your kitchen and will be the envy of every avocado lover in the neighborhood!

—

The perfect gift for any avocado lover

Ad Copy: Get ready to have your eyes opened when you see the newest in avocado-inspired kitchenware. Introducing The Mug Company's avocado-shaped mug. The perfect gift for any avocado lover that wants a delicious, healthy breakfast in a mug.

—

Satisfy your avocado cravings

Ad Copy: If you're looking for a fun way to eat your avocado toast, then this mug is for you! You get an avocado-shaped mug that is perfect for holding a slice of toast, and the peel doubles as the handle. For the best breakfast ever, order your mug today!

<https://rytr.me/>

最新情報を反映するテキスト生成

FRUIT: Faithfully Reflecting Updated Information in Text (Logan IV+ arXiv 2021)

Tom Kristensson

Tom Krister Kristensson (born 30 April 1991) is a Swedish rally driver, who drives in the ADAC Opel Rallye Cup.

He is originally from Lund, but now lives in Hörby

In 2016 and 2017, Kristensson competed in the ADAC Opel Rallye Cup, finishing second in 2016 and winning the 2017 championship.

Input: Original Article

Tom Kristensson

Tom Krister Kristensson (born 30 April 1991) is a Swedish rally driver, who drives in the Junior World Championship.

...

+ In the 2019 season of JWRC, Tom finished second behind Jan Solans.

+ The next season he went on to become the 2020 Junior World Rally Champion.

Output: Updated Article

Junior World Rally Championship Results Driver's Championship (1)

Year	Champion	2nd Place	3rd Place
2019	Jan Solans	Tom Kristensson	Dennis Rådström

2020 Junior World Rally Championship (2)

...Tom Kristensson and Joakim Sjöberg won the 2020 junior championships, while Sweden sealed a back-to-back Nation's Trophy...

Input: New Information

Junior World Rally Championship Results Wins by Driver (3)

Driver	Total
Tom Kristensson	5

言語処理技術による執筆の支援：執筆に与える影響は？

作成した**テキスト**の質？ 執筆者・学習者の**態度**への影響？

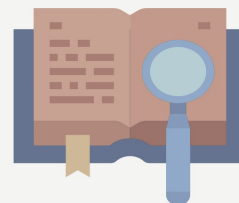
文章生成 (draft)



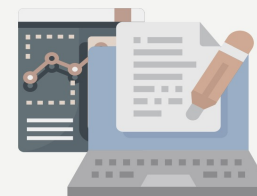
Machine translation



Summarization



Data-to-text



https://cs.stanford.edu/~minalee/pdf/cornell-writing_with_language_models-slides.pdf

隅田英一郎「2021年の翻訳バンク・ニューラル翻訳」第12回産業日本語シンポジウム、他多数
岡崎直観「自然言語処理によるニュース記事執筆の現状と課題」第11回産業日本語シンポジウム

校正・添削
(edit, diagnose)



スペル・文法・スタイルのチェック、言い換え

Type your title

To begin with, food that is easier to prepare allows people to save time. If people want to cook food by themselves, they have a lot of task to do. First, people have to go to supermarkets and buy food ingredients. Going to supermarket and choosing preferable food ingredients need at least thirty minute to one hour. In addition, cooking food need more time. What is more, if they cook food in home, people have to wash their plates or pans after enjoy dishes. On the other hand, if people buy prepared food or frozen food, all they have to do is just heat the food in their microwave. I think people need only five minutes to prepare such convenient foods except for the time to buy them.

task → tasks

the supermarket or a supermarket

minute → minutes

in-addition → Also

Repetitive word: need

Possibly confused preposition

in home → at home

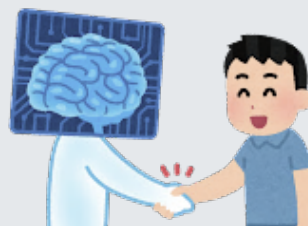
It appears that *lit* may not be the best preposition to use in this context. Review the information below to decide whether to change it.

MORE IGNORE

the home

<https://www.path-to-success.net/grammarly>

共同執筆
(co-author)



テキストの自動補完、挿入文の提案

Get suggestions

So he lived in the basement.

So he built a big wall around it, and he lived in a little shack on the other side of the wall.

So he moved next door to it.

He moved in a couple of miles away.

So he built a little shack next to it, and he painted it and put a little flower garden outside and everything.

Edit

The 92nd little pig built a house out of depleted uranium. There are so many pigs to eat, and they make it so easy!" But the wolf didn't know about uranium.

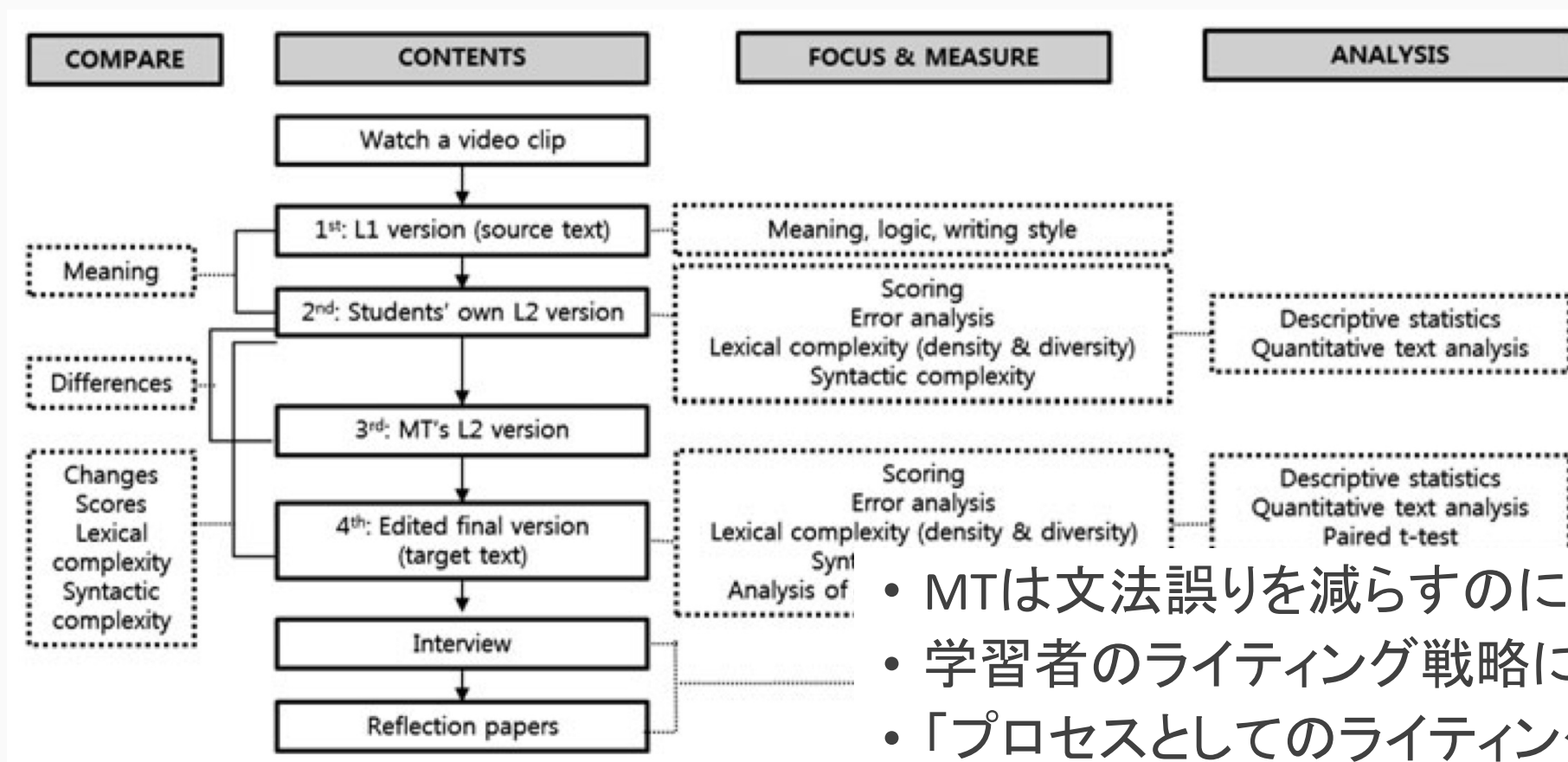
Of course, the little pig was smart. He knew he couldn't live in the house, since it would poison him. So he built a big wall around it, and he lived in a little, hidden shack on the other side of the wall.

<https://arxiv.org/pdf/2201.06796.pdf>

執筆への影響（文章の質、執筆者の態度）を評価する試み

機械翻訳の効果の測定はこの2～3年で増えてきた

例えば、Lee (CALL 2019) The impact of using machine translation on EFL students' writing



- MTは文法誤りを減らすのに役立った
- 学習者のライティング戦略に良い影響を与えた
- 「プロセスとしてのライティング」の見方を促進
- ただし、適切な使い方のガイダンスが必要

執筆への影響（文章の質、執筆者の態度）を評価する試み

執筆支援の文脈での人を入れた評価はまだまだこれから



(前頁のLee氏とは別人)

ライティングの支援/診断技術とライティングへの影響

ニューラル翻訳・言語モデルの発展・大規模化で言語生成が流暢になってきた
流暢になった言語生成技術を執筆支援に応用する試みが広がってきている
執筆支援は英語用の技術やサービスが先行しているが、要領は日本語も同じ
機械による言語生成の理屈を把握した上で人間と機械の共同作業を設計すべき
人間と機械の役割分担の設計とその影響の丁寧な評価・検証はまだまだこれから