

短歌における言語モデルの実応用 —歌人の視点を通じた生成と作歌支援の実践から—

浦川通¹ 新妻巧朗¹ 田口雄哉¹ 田森秀明¹ 岡崎直観² 乾健太郎^{3,4}
¹ 株式会社朝日新聞社 ² 東京工業大学 ³ 東北大学 ⁴ 理化学研究所
{urakawa-t, niitsuma-t, taguchi-y2, tamori-h}@asahi.com,
okazaki@c.titech.ac.jp, inui@tohoku.ac.jp

概要

ニューラル言語モデルによる生成が多方面で実サービス化され、その利用に関する実際的な検討が始まっている。テキストによる創作活動である文芸分野でも、言語モデルの実応用化が進んでいるが、短詩形文学の一つである短歌においては、その応用可能性について積極的に議論される現状がなく、実応用事例も存在しないため、モデル利用の有用性や課題に不明点が多い。本稿では、短歌を生成する言語モデルの評価や可能性について、実際に歌壇で活動する歌人を取材し、そこで得られた作歌支援への指摘にもとづき構築した、実応用事例の設計法や実施の結果を共有しながら、今後の課題を述べる。

1 はじめに

ニューラル言語モデル（以下、言語モデル）の発展にともない、対話応答 [1]、プログラミング支援 [2]、論文執筆支援 [3] といった幅広い領域で言語モデルの実応用化が進んでいる。これにしたがって、論文投稿の現場において言語モデルの利用に関する方針策定 [4] が行われるなど、実応用における倫理と有用性をふまえた検討がすでに始まっている。この流れはテキストによる創作活動である文芸にも及び、執筆支援での言語モデル応用が検討され [5]、専門家による評価が行われているほか [6]、日本語でも実サービス化されたものが存在する [7]。

短詩形文学の一つに数えられる短歌¹⁾においても、これまでに自動生成手法が提案されてきた [8, 9, 10, 11]。最近では、言語モデルによる生成を一部に含む短歌の連作が新人賞の最終選考を通過するなど、実際の作歌の現場に生成が登場しはじめている [12, 13]。しかし、言語モデルの評価が歌人

1) 本稿でいう短歌は、近世以前に詠まれた和歌を含むものとする。

をはじめとする実作者の立場から行われた例はいまだ少なく [14]、実際的な応用可能性について、積極的に議論されている現状にない。また、広く一般に利用される形で実サービス展開された事例も存在しないため、短歌を対象とした言語モデル応用の有用性や課題については、明らかでない点が多い。

本稿では、実際に歌壇で活動する歌人に対して短歌を生成する言語モデル（短歌生成モデル）に触れる機会を設け、評価また応用可能性について取材し、その結果をふまえて構築した実応用事例の設計法や実施の結果を共有しながら、今後の課題について述べる。

歌人への取材においては、先行研究であるモーラを考慮した短歌生成モデル [11] を用意し、生成に関する自由な意見をもらうとともに、今後考えられる応用可能性について取材した。この結果として、作歌のヒントを得ることを目的とした言語モデル利用の可能性に関する指摘が得られた。

実応用事例においては、ユーザの入力に従う作歌支援システムを構築した。これをウェブサービスとして公開するとともに、短歌の創作を行うワークショップにて展開し、ユーザが実際に触れる機会を提供した。ここで得られた利用データの分析と利用者の声から、実際に言語モデルによる生成がユーザの作歌に援用されうることを確認した。一方で、ユーザの創造性を引き出すためのより効果的な言語モデル応用法の検討や、構築における学習データ利用環境の整備を今後の課題として議論する。

2 歌人からみた短歌生成モデル

短歌生成モデルの評価と応用可能性について、歌人の俵万智氏、永田和宏氏を対象に、以下の方法で取材した。

表1 各取材におけるモデルの生成例, () は入力

俵	(一人称あまり使わぬ日本語に) 君の心を隠しているか
永田	(足裏を砂流れおり) 見る人の心に届く一石となる

2.1 方法

先行研究であるモーラを考慮した短歌生成モデル[11]を, 日本語ウィキペディアから抽出した疑似短歌(5,7,5,7,7 モーラの短歌の定型を満たす文字列)でファインチューニングしたものと, このモデルをさらに俵の短歌データ約2,000件でファインチューニングしたものを用意し, 永田には前者のみの, 俵には両者の出力結果を確認できる場をそれぞれ別個に提供しながら(1)生成過程・結果についての自由な意見(2)今後どんな影響・応用が考えられるかの2点について意見を伺った。なお, 以降の各人の発言内容は, 記事²⁾³⁾⁴⁾からの引用とする。

モデルによる実際の生成例を表1に示す。また, モデル作成に関する詳細を付録Aに記す。

2.2 議論

生成過程と結果について 俵は, 自作を学習したモデルの出力について「私っぽい」と指摘するとともに, ある生成については「やられたな」とその結果に驚く反応を見せた。永田は「まだAIがこのレベルでよかった」としつつも, 「それらしきいい歌ができることは間違いない」と発言し, 両者それぞれにモデルによって生成された短歌の質を評価した。また両者から, 過去に詠まれた短歌を大量に学習する過程は, 歌人が作歌を学ぶときとあまり変わらないかもしれないと, 短歌の学習における人とモデルの類似性が指摘された。

可能性と課題 永田は「たとえば上の句で筆が止まっている時に, AIに提案してもらった言葉を発展させて下の句を付けるという使い方は, 認めてもいいのかもしれない」と, 言語モデルの創作における応用について言及した。「歌をつくる前はこう思っていたけど, 歌をつくるプロセスでこうも思ったんだという自分の発見があって(中略)そうした言葉

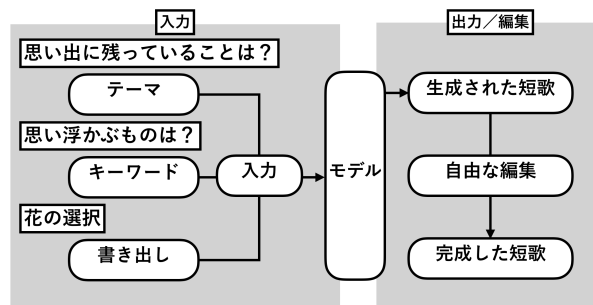


図1 「花と歌」システム図

をAIが見つけてくれようと自分で見つけようと, 本質は変わらないのかもしれない」と指摘する。俵は, 1秒間に100首程度生成できるという, モデルの高い量産性を挙げながら「壁打ち相手としては魅力的」と言及した。一方で「AIに名歌をつくってもらう必要はない」「歌の種は人の心にある」とも発言している。両者ともに, 生成される短歌の質は認めながらも, あくまで人が作歌を行うときにヒントを得ることを目的とした応用可能性を指摘していることがわかる。これらの結果から, 短歌の一部を入力し生成された結果を見ながら, より良い表現を探すといった言語モデルの応用可能性が得られた。

3 短歌生成モデルの実応用

2節での取材の結果, 作歌におけるヒントを得るために言語モデルを用いる可能性について指摘が得られた。これをふまえ, ユーザの入力に従う作歌支援システムを構築した(図1)。

3.1 短歌生成モデルによる作歌支援

システム このシステムは, ふだん言えない気持ちを花と歌(短歌)に込めて届けるウェブサービス「花と歌」⁵⁾のために作られたものである(図2)。ユーザの贈りたい相手に向けて, 花のイラストと短歌を合わせた画像を共有できるというサービスで, 作歌支援に言語モデルを用いている。ここでユーザは(1)贈りたい相手に関する質問に答える(2)質問への回答内容に沿った短歌一首が生成される(3)生成された歌を自由に編集する, という手順を通して短歌を作ることができる。今回の取り組みは, これまでに一度も短歌を作ったことのないユーザの利用を想定している。そのため, 言語モデルの入力として短歌の一部をユーザに考えさせる仕組みにすると, その難しさによってサービスからの離脱を招く

2) <https://digital.asahi.com/articles/ASQ716T8VQ6HUCVL027.html>

3) <https://www.asahi.com/special/tawaramachi-aitanka/>

4) <https://www.asahi.com/articles/ASQ716V8VQ6QUCL006.html>

5) <https://prtimes.jp/main/html/rd/p/000000014.000054763.html>



図2 「花と歌」スクリーンショット

表2 「花と歌」学習データの例. 特殊トークン<sep>, <r>, <k>を用意し, 入力はキーワード<sep>類語 1<r>類語 2...<sep>初句の形で作る. 出力は各句の間を<k>で挟む.

入力	わたし<sep>楽しく<r>ハンバーグ<r>ラーメン<r>カレーライス<r>仲良く<sep>まあいいや
出力	まあいいや<k>わたしは君の<k>ことが好き<k>今日も一緒に<k>カレーを作る
入力	僕<sep>アンビアンス<r>アトモスフィア<r>朝<r>深夜<r>清爽<sep>包まれる
出力	包まれる<k>早朝のこの<k>さわやかな<k>空気を吸うと<k>僕は決まって

表3 「花と歌」モデルによる生成例. sは書き出し, kはキーワード, tは類語リストの入力値.

s=いつまでも, k=ダンス,
t=[心地好い, 楽しい, 仲良し, すこやか, 瞬間]
いつまでも穏やかでいてくれるので
一緒にダンスするときもそう

可能性が考えられた. そこで, ユーザは簡単な質問に答えるのみによって短歌を生成でき, その後一首全体の編集権を得るという構成をとっている.

モデル 今回のシステムにおけるモデルは, 短歌の一部ではなくユーザの質問回答の内容を入力に短歌を生成する. そのため, 2節のようなデコーダのみによって構成された言語モデルではなく, seq2seqの機構を持ち入力をエンコードした結果を用いて生成するモデルを採用した. 具体的には, 事前学習済み言語モデルである BART [15] を CTRL [16] のようにタグコントロールによる条件付き生成モデル学習の枠組みでファインチューニングした⁶⁾. 入力として, 出力される歌に含めたいキーワード, 書き出し(第1句)に加え, 歌のテーマとして出力中に含まれる単語の類語5つを与え, 短歌を生成するものとする(表2). キーワードのみユーザによる自由入

力を受け付け, その他の入力はユーザの質問回答によってシステム内部で自動的に与えられる. 実際の生成例を表3に示す. モデル作成に関する詳細を付録Bに記す.

データセット 今回のように, 広く一般に公開され, 創作支援として利用される言語モデルの作成において, その学習データがどのようなものであるかが, サービスの信頼に関わる大きな問題となりうる. 朝日新聞社の場合, 保有するデータには「朝日歌壇」入選歌などの短歌が含まれるが, これを生成に利用するための許諾は得ていない. 著作権や創作上のモラルといった観点もふまえて社内で議論し, 既存の短歌を学習データとして意図的に利用することは避けるという方針を定めた. そこで, 学習データは[11]と同様にコーパスから疑似短歌を抽出することで作成した. コーパスはCC-100⁷⁾の日本語文書データセットで, 抽出した疑似短歌は65,770件である. 各サンプルにおけるキーワードはTextRank [17]のスコア上位3件よりランダムに1つ抽出した. テーマとして与える類語群は, 朝日新聞単語ベクトル[18]により各サンプルにおける各単語の類似語上位5件ずつを得た中から, 5件をランダムに抽出して作成した.

6) 事前学習済み言語モデルは, 朝日新聞社の保有記事約900万件で事前学習したものを使用した.

7) <https://data.statmt.org/cc-100/>

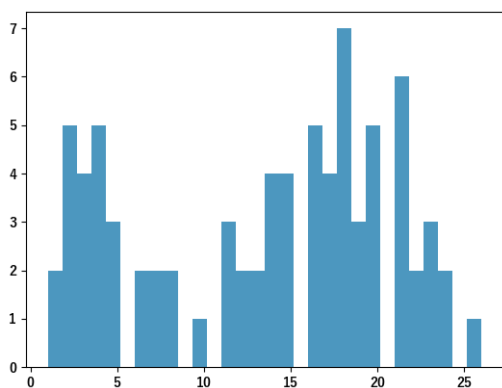


図3 編集が与えられた短歌における、編集前後での編集距離の分布。横軸は編集距離。縦軸は短歌数。

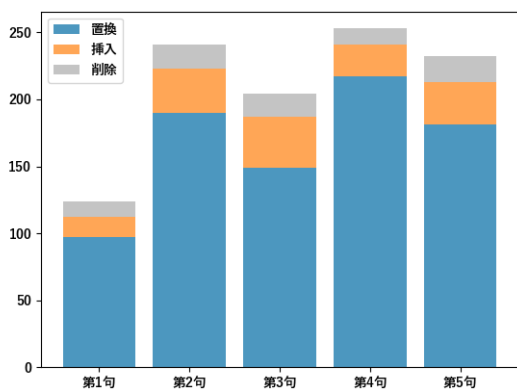


図4 編集が与えられた短歌における、各句ごとの編集操作の分布。縦軸は文字数。

3.2 実施と議論

このシステムは、2022年の9月から1ヶ月間の期間限定でウェブ上で公開されるとともに、短歌創作のワークショップにおいて、参加者が作歌のヒントを得るためのツールとして利用された。

システムが実際にどのように利用されたかを見るため、ウェブ上でユーザが作成し、実際に他者と共有した短歌データから254件を抽出・分析した。図3に、編集が与えられた短歌における編集距離の分布を示す。編集の与えられた短歌の多くが、生成の一部を編集することで作成されたことがわかる。また、図4に示す編集操作の分布を見ると、最も多い操作が置換で、入力と与えられる第1句を除き、各句で偏りなく編集されていることがわかる。これらの結果から、ユーザが生成された短歌の内容を生かしながら、自身の歌を作っていることが窺える。

表4 ユーザによって共有された歌の編集前後の例。波線、下線はそれぞれ編集前後における編集箇所。

編集前	すがすがしい眺められるが観覧車は なんとなく見ているだけなので
編集後	すがすがしいホテルから見る観覧車 なんとなく見ているだけなのに
編集前	希望へと散歩に行つて思い出に 残る時期を迎えることの
編集後	希望へと散歩に行つて思い出に これから先も迎えることの
編集前	誠実さ負けず嫌いの散歩等 楽しむという時もあります
編集後	誠実さ負けず嫌いの散歩道 楽しむという時もあります

ワークショップの参加者からも「出力したフレーズを取り入れて作歌できた」「ヒントとなる言葉に出会えておもしろかった」といった反応が得られ、モデルが作歌時の発想を広げる手助けとなったことが確認できた⁸⁾。表4に、実際の編集例を示す。

一方で、全体の7割近くの生成結果が、ユーザの編集を経ずにそのまま共有されていたこともわかった。これは、生成された短歌の質がユーザに受け入れられた結果とも考えられるが、作歌のヒントを与える応用とはならなかった例が多い結果であるともいえる。よりユーザの創造性を引き出す言語モデルの応用法の検討が、今後の課題である。

また、今回のモデル学習は疑似短歌によって行われている。実際の短歌を学習したモデル生成が歌人による評価を得ている中で、今後の応用に向けたデータセットの整備やデータ利用における合意形成などが、より活発な実応用に向けて必要であると考えられる。

4 おわりに

本稿では、短歌生成モデルの実応用について、歌人への取材を通してその可能性について整理した。また、取材の結果をふまえて構築・展開した作歌支援システムについて説明し、ユーザの反応を分析することで、短歌生成モデルが実際にユーザの作歌を支援する形で利用されることを確認した。今後の課題として、よりユーザの創造性を向上させる言語モデル応用法の検討や、学習のための短歌データセットの整備などに取り組みたい。

8) <https://digital.asahi.com/articles/ASQ9P3Q03Q9LUCVL01L.html>

謝辞

取材に協力頂いた俵万智氏、永田和宏氏に感謝する。データ提供を頂いた株式会社 電通クリエイティブ X に感謝する。

参考文献

- [1] ChatGPT: Optimizing Language Models for Dialogue, (2023-1 閲覧). <https://openai.com/blog/chatgpt/>.
- [2] GitHub Copilot · Your AI pair programmer, (2023-1 閲覧). <https://github.com/features/copilot>.
- [3] Galactica Demo, (2023-1 閲覧). <https://galactica.org/>.
- [4] ACL 2023 Policy on AI Writing Assistance, (2023-1 閲覧). <https://2023.aclweb.org/blog/ACL-2023-policy/>.
- [5] Ann Yuan, Andy Coenen, Emily Reif, and Daphne Ippolito. Wordcraft: Story Writing With Large Language Models. In **27th International Conference on Intelligent User Interfaces**, pp. 841–852, 2022.
- [6] Daphne Ippolito, Ann Yuan, Andy Coenen, and Sehmon Burnam. Creative Writing with an AI-Powered Writing Assistant: Perspectives from Professional Writers. **arXiv preprint arXiv:2211.05030**, 2022.
- [7] Hiroyuki Osone, Jun-Li Lu, and Yoichi Ochiai. BunCho: AI Supported Story Co-Creation via Unsupervised Multi-task Learning to Increase Writers' Creativity in Japanese. In **Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems**, pp. 1–10, 2021.
- [8] Tomonari Masada and Atsuhiko Takasu. LDA-based scoring of sequences generated by RNN for automatic Tanka composition. In **International conference on computational science**, pp. 395–402. Springer, 2018.
- [9] Makoto Nakatsuji and Sohei Okui. Conclusion-Supplement Answer Generation for Non-Factoid Questions. In **Proceedings of the AAIL Conference on Artificial Intelligence**, Vol. 34, pp. 8520–8527, 2020.
- [10] Yuka Takeishi, Mingxuan Niu, Jing Luo, Zhong Jin, and Xinyu Yang. WakaVT: A Sequential Variational Transformer for Waka Generation. **Neural Processing Letters**, Vol. 54, No. 2, pp. 731–750, 2022.
- [11] 浦川通, 新妻巧朗, 田口雄哉, 田森秀明, 岡崎直観, 乾健太郎. モーラを考慮した fine-tuning による口語短歌生成. 言語処理学会第 28 回年次大会, 2022.
- [12] 浦川通. AI が俵万智さんの歌集を学習したら開発者が言語モデルを解説. 朝日新聞デジタル, 2022-07-05. <https://digital.asahi.com/articles/ASQ744WG1Q71UCVL01Y.html>.
- [13] 第六十四回 短歌研究新人賞発表. 短歌研究, Vol. 78, No. 9, 09 2021.
- [14] 中辻真, 奥井颯平, 野口あや子, 加古陽. 特集 作品二十首&修業秘話 歌歴 1 年半でここまで上達! 歌人 AI(人工知能歌人) の歌力: ついに歌壇デビューか! 筆名は「恋する AI 歌人」! 短歌研究, Vol. 76, No. 8, pp. 91–102, 08 2019.
- [15] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Ves Stoyanov, and Luke Zettlemoyer. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. **arXiv preprint arXiv:1910.13461**, 2019.
- [16] Nitish Shirish Keskar, Bryan McCann, Lav R Varshney, Caiming Xiong, and Richard Socher. CTRL: A Conditional Transformer Language Model for Controllable Generation. **arXiv preprint arXiv:1909.05858**, 2019.
- [17] Rada Mihalcea and Paul Tarau. TextRank: Bringing Order into Text. In **Proceedings of the 2004 conference on empirical methods in natural language processing**, pp. 404–411, 2004.
- [18] 田口雄哉, 田森秀明, 人見雄太, 西鳥羽二郎, 菊田洗ほか. 同義語を考慮した日本語の単語分散表現の学習. 研究報告自然言語処理 (NL), Vol. 2017, No. 17, pp. 1–5, 2017.
- [19] Taku Kudo and John Richardson. SentencePiece: A simple and language independent subword tokenizer and detokenizer for Neural Text Processing. In **Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations**, pp. 66–71, Brussels, Belgium, November 2018. Association for Computational Linguistics.

A 取材時に利用した短歌生成モデル

2 節の取材における言語モデルの構成や学習時のハイパーパラメータについて記す。Hugging Face 上で公開されている日本語 GPT-2⁹⁾を事前学習モデルとしてファインチューニングすることで作成した(表 5)。歌人データによるモデルは、疑似短歌によるモデルをさらにファインチューニングすることで作成している(表 6)。モデルの学習に用いたデータ数は、ウィキペディアによる疑似短歌が 10,000 件、俵による短歌データが 2,372 件である。

表 5 疑似短歌データによる短歌生成モデル構成とハイパーパラメータ

アーキテクチャ	GPT2-medium
トークナイザ	Sentencepiece
最適化器	Adam
バッチサイズ	32
エポック数	50

表 6 歌人データによるファインチューニングモデル構成とハイパーパラメータ

アーキテクチャ	GPT2-medium
トークナイザ	Sentencepiece
最適化器	Adam
バッチサイズ	16
エポック数	10

B 「花と歌」短歌生成モデル

3 節のシステムにおける言語モデルの構成や学習時のハイパーパラメータを表 7 に示す。実装には、Hugging Face Transformers¹⁰⁾を用いている。トークナイザは朝日新聞社保有の記事を学習した Sentencepiece [19] で、語彙数は 32,000 である。なお、事前学習済みモデルである BART-base は、朝日新聞社の保有する記事約 900 万件で事前学習したモデルを利用した。

表 7 「花と歌」モデル構成とハイパーパラメータ

アーキテクチャ	BART-base
トークナイザ	Sentencepiece
最適化器	Adam
バッチサイズ	128
エポック数	10

9) <https://huggingface.co/rinna/japanese-gpt2-medium>

10) <https://github.com/huggingface/transformers>