Poetry Machines: Eliciting Designs for Interactive Writing Tools from Poets

Kyle Booten*
University of Connecticut, Storrs
USA
kyle.booten@uconn.edu

Katy Ilonka Gero*
Columbia University
New York City, USA
katy@cs.columbia.edu

ABSTRACT

Improvements in natural language processing and generation have made possible new and powerful creativity support tools for creative writers. However, it remains unclear how professional writers themselves might want to integrate technology into their existing writing practices. In this work we ran an elicitation study, asking 14 professional poets to consider how they would make use of computation in the context of a custom, interactive writing interface or "Poetry Machine." We found that the poets desired a wide range of functions, from presenting auditory responses to deleting random words. We also found that many poets did not simply report what their ideal interface would do but rather contextualized their designs by describing why they are artistically meaningful, sometimes with respect to specific literary influences and traditions. We present an initial analysis of the elicitation study and observe some differences between the Poetry Machine designs and existing creativity support tools. This study lays the groundwork for a second phase where we will build a selection of the machines and study how the poets use them over time.

CCS CONCEPTS

• **Human-centered computing** → **Interaction techniques**; **Interactive systems and tools**.

KEYWORDS

creativity support tools, creative writing, poetry, natural language processing, interactive writing tools

ACM Reference Format:
Kyle Booten and Katy Ilonka Gero. 2021. Poetry Machines: Eliciting Designs for Interactive Writing Tools from Poets. In *Creativity and Cognition (C&C '21), June 22–23, 2021, Virtual Event, Italy.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3450741.3466813

1 INTRODUCTION

As the abilities of natural language processing improve, researchers continue to experiment with using text-based algorithms to support creative writing [2, 7, 8, 16]. However, this research typically proceeds from certain assumptions about what kind of software poets in general would want—e.g. "autocomplete"-like functionality. We see a need to learn more about how specific poets, representing diverse literary traditions, might imagine integrating computation into their own practices. We present an elicitation study in which we asked 14 contemporary American poets to imagine and describe their own custom "Poetry Machine," a computational, interactive writing interface.

We annotated the resulting designs across a variety of axes and compared their responses to existing creativity support tools for writing found in the literature. In this poster we first give an overview of some key themes that emerged from these elicited designs, with consideration of both the technical aspects of the Poetry Machines and the ways that poets chose to present and justify them. We then observe some differences between Poetry Machines and the kinds of creativity support tools produced by researchers. By paying attention to the ways that these poets' designs are rooted in their own literary tastes and traditions as well as their personal stylistic and expressive goals, we approach this project in the spirit not just of participatory design but of "humanistic" design [1].

2 ELICITATION STUDY

2.1 Methodology

2.1.1 Recruiting Poets and Eliciting Poetry Machine Designs. We wanted to recruit poets who represented a broad spectrum of contemporary American poetics. Based on our knowledge of American literary magazines, we produced a list of those that we believed to represent an array of styles, from the more traditional or formalist to the more experimental. By browsing these magazines, we produced a list of 35 poets whose work we found compelling.

Each of these poets was sent a recruitment email explaining that the study would require them to take part in a design activity lasting about an hour and that we would compensate them $35 in the form of a gift card to a bookstore of their choice. If the poet wanted to partake, we sent them an initial survey that asked questions about their poetic influences, poetry practice, and experience with computers and computation. The questions of the survey can be found in the Appendix A. After they responded to the survey, they were sent an email with an attached pdf we called the "design guide" that explained how to design a Poetry Machine. In this guide we explained what a Poetry Machine is:

> A Poetry Machine is a computer interface for writing that responds to what you write. This response may be in the form of inspiration – like surfacing relevant materials – or it could literally change the text as you type it. Think of it as a kind of "collaborator" that wants to write a poem with you or help you while you

*Both authors contributed equally to this research.

C&C '21, June 22–23, 2021, Virtual Event, Italy

ACM ISBN 978-1-4503-8376-9/21/06…$15.00
https://doi.org/10.1145/3450741.3466813

are writing a poem—not before or after the writing process but during it.

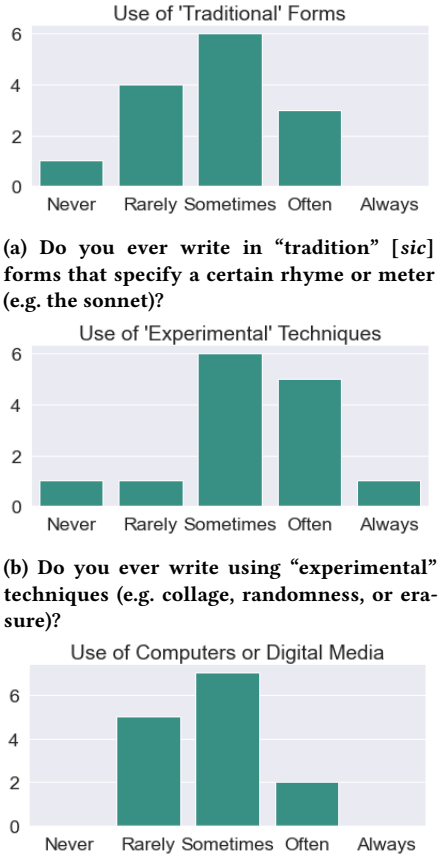
The guide then outlined a way of thinking about a Poetry Machine in terms of a loop: what triggers it, how it can focus on parts of what has already been written, how it might think about what's already been written, and how it may act in some way. Then we gave examples Poetry Machines (see Appendix B). Poets were instructed to present their design either in the body of an email or as a Word document. This study was approved by the IRBs of our institutions. Of the 35 poets selected, 24 indicated interest in participating in the study, 17 completed the survey, and 14 of these actually completed a Poetry Machine design. We report only on the 14 poets who both completed the survey and returned us a Poetry Machine design. 10 of the poets wanted to be referred to by their real name in any publications.

2.1.2 Analyzing the Poetry Machines. The authors of this paper read the Poetry Machine designs together and separately many times, discussing emergent themes among the designs as well as unexpected elements. Through these discussions we developed a set of questions to pose about a Poetry Machine (such as "What triggers the machine to act?") and iterated on these questions as we attempted to answer them about all the machines. We followed an inductive thematic analysis approach, focusing on what was in the designs rather than on our pre-existing ideas about what the designs might do or contain [4]. This led to a set of codes that answered the questions, which we used to annotate the machines. The annotation was done via consensus, where one author annotated all the machines for a particular question and then discussed any edge cases with the other author. The codes and their prevalence are presented in Section 2.2.2, as well as in Table 1.

2.2 Findings

2.2.1 Description of Poets. Overall we found that the poets did reflect a diverse spectrum of contemporary American poetic practices, with some working mostly with traditional forms like the sonnet, others always relying on experimental techniques like collage[1] or erasure[2], and most somewhere in between. When asked questions about the use of traditional forms, experimental techniques, and digital media in their poetic practice, most poets responded that they used all of these every once in a while (see Figure 1). There was a slight bias in our sample when it came to the use of experimental techniques: most poets used experimental techniques "sometimes" or "often," though this may be a reflection of contemporary American poetic tendencies.

2.2.2 Description of Poetry Machines. The 14 poets produced a total of 20 Poetry Machine designs. However, not all designs were valid according to our definition. Designs were disqualified if they were not in some sense interactive—responding to the input of the human writer and thus allowing for a back-and-forth between human and machine. 5 of the 20 designs were non-interactive; these either described an unchanging prompt that was meant to inspire the human poet before writing or a system for pure computer-generation (i.e. no direct human input whatsoever). Thus we were left with 15 designs from 12 different poets. (One poet contributed three valid designs while another contributed two.) The average number of words per these 15 designs was 330. The longest was 504 words, the shortest 101. Two contained images to illustrate their descriptions.



(a) Do you ever write in "tradition" [*sic*] forms that specify a certain rhyme or meter (e.g. the sonnet)?

(b) Do you ever write using "experimental" techniques (e.g. collage, randomness, or erasure)?

(c) Do you ever use computers or digital media in your poetic process (e.g. random word generators)?

Figure 1: Survey responses from the 14 poets when asked about characteristics of their writing practice.

The Poetry Machines display a wide variety of approaches to the task of designing an interface to help one write a poem. Some of the features imagined by the poets were ones that seemed, if not unsurprising, immediately understandable in light of typical poetic techniques. For instance, two different poets designed systems that would furnish them with rhyming words to match the final word of the previous line (not unlike a rhyming dictionary). Several described systems that would rejigger the poet's verse into various traditional poetic forms. However, most Poetry Machines were comprised of features that we would have not predicted and that seemed highly idiosyncratic. One poet's Poetry Machine would adjust the poem's font to match its tone. Another's would try to

[1] Akin to collage in visual art, combining language excerpted from other, often non-poetic texts.
[2] Creating a poem by erasing or otherwise obscuring the words of a pre-existing text.

rearrange one of the poet's lines anagrammatically. The Poetry Machines abound with such unique and imaginative features.

2.2.3 Trends in Poetry Machines. In general, what kind of intervention into the poet's writing process did the Poetry Machine design specify? Our cycles of qualitative coding led us to two categories and several subcategories. First, some machines would present information ("presenting"), while others would directly change the poem ("changing"). There was at times some ambiguity here, as poets did not always specify exactly how they would respond to the action of their Poetry Machine. We considered as examples of "changing" any suggestion that could trivially be integrated by the writer into the poem. On the other hand, the machines that would merely present information would do so for one of two reasons. First, some would "diagnose" the text by describing or visualizing some facts about it, such as whether or not a poem satisfies the demands of a particular poetic form. Others would offer "oblique suggestions," presenting pieces of information that seemed designed to provoke the imaginative thinking of the writer.

The Poetry Machines ranged in their levels of complexity. Most designs specified only one or two actions (tagged either "changing" or "presenting). To be specific, 11 out of 15 valid machines contained one or two such actions, with the average number being 3. Two designs by two different poets were much more complex, consisting of 11 and 12 actions respectively. These designs were "laundry lists" of heterogeneous actions, and their designers made no attempt to explain to what degree they added up to coherent, integrated machines.

Our second question was: what mode would the Poetry Machine produce? The Poetry Machines specified four different types of modes: pure text (i.e. merely words), image, a sound, and a "visual text" (some change to the graphic presentation of words).

Table 1 presents counts of these various codes as well as examples designed to give some sense of the impressive diversity of the poets' designs.

2.2.4 How Poets Contextualize their Designs. We did not specifically request that the poets give any rationale for particular design choices. We were surprised then when several did, connecting their Poetry Machines not to poetry in general but to their very particular artistic interests. For example, Mark Faunlagui started one of his Poetry Machine designs like this:

> The Modernist films of Jacques Tati and the Renaissance canvases of Bruegel present two examples of expansive visual worlds, in which a multitude of participants are—in the same space—engaging in a wide array of actions at the same time. Despite the simultaneity of all the events, seen from a "clear" omnipotent birds-eye view, these works demand repeated watching/ looking…blink, and you miss details!

Faunlagui then described how his machine would scan his poem for elements of the underlying images and present "counterpoints" or "couplings" to those images that pulse and oscillate. While he could have described this machine without this evocative context, he spent time and effort to explain his design in light of an insightful interpretation of the similarities between Tati and Bruegel. This context explains his particular fascination, such that even if another poet had designed a very similar machine, it's unlikely that the other poet would have designed it to achieve the same conceptual purpose.

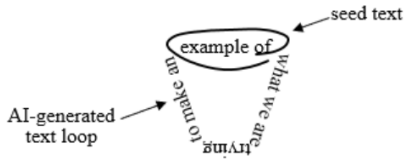


Figure 2: Ava Hofmann's "Ouroboros" design.

Another poet, Melissa Garcia Criscuolo, contextualized her machine by explaining that she wanted it to do something "similar to Ray Bradbury's makings for a story." She went on to quote from Bradbury's book *Zen in the Art of the Writing* and then described her design for a machine that would implement a specific idea Bradbury offers regarding how to write a story from a set of images. Again we see that this context indicates some specific fascination the poet has, this time with a specific process, that explains why she wants her machine to do the specific thing it would do. Likewise, the (anonymous) poet who described a Poetry Machine that would "indicate when a freshly composed line can be anagrammatically rearranged" mentioned via correspondence that this desire reflected an interest in the anagrammatic poetry of James Merrill. Such Poetry Machines designs are not merely in response to some abstract or generalized notion of what poets need or what challenges they face in composing a poem (e.g. "writer's block"). Rather these machines connect to literary and artistic history, just as poets writing poems do so in some relationship to their forebears and favorite writers, not in a vacuum.

While we hope to further investigate the relationship between individual poets' respective styles and their imagined Poetry Machines, already we can point to interesting connections between the two. For instance, Ava Hofmann was one of the two poets whose designs featured images and that had a striking visual component. One of her designs was the "Ouroboros":

> The user would write a single, open-ended line of language/verse, and upon pressing enter, the Ouroboros would generate a piece of text that could incorporate the text included into a loop, and then visually arrange it as such.

Hofmann included a mock-up drawing of the Ouroboros (see Figure 2). A reader familiar with Hofmann's published poetry will note that it frequently experiments with the graphical aspects of text, including the text's direction, in ways that are very reminiscent of this design. This is just one particularly striking example of a Poetry Machine that, while it may on its own seem difficult to explain or justify, connects naturally to its designer's style and mode of composition.

3 COMPARING POETRY MACHINES TO CREATIVITY SUPPORT TOOLS

One of our main research goals is to assess the fit between existing creativity support tools (CSTs) and the kinds of tools that poets themselves would want. Here we compare some existing CSTs with the Poetry Machines imagined by poets in our study. We note that

Code	Count	Description	Example
		Q1: Overall, what is it doing?	
present: oblique suggestion	7	Machine presents artifact/s intended to inspire; however, it does not immediately recommend direct changes to the text.	"bring up a google image search for that phrase"
present: diagnostic	3	Machine presents information about the poem as it currently exists.	"color-codes a poem … according to parts of speech"
change: replacement	3	Machine actively replaces part of poem, or presents potential replacements.	"Change preposition to spatial opposite"
change: restructure	6	Machine actively restructures part of poem, or presents potential new structures.	"take my list of words and ... organize it into some narrative"
change: addition	6	Machines actively adds to poem, or presents potential additions.	"These speculative completions … would be presented "
change: deletion	2	Machine actively deletes part of poem, or presents potential deletions.	"identify and delete random words"
		Q2: What mode does it produce?	
text	12	Machine presents standard text.	"generate an original simile"
visual text	5	Machine presents text in a custom layout.	"that word, once written, will remain, but pulse and oscillate"
image	5	Machine presents images.	"bring up a google image search for that phrase"
auditory	4	Machine presents sound.	"emit an auditory response"

Table 1: List of the annotation codes, along with how many Poetry Machines were tagged with the code, description of the code, and example from a Poetry Machine.

in our study poets were asked to design a custom machine for themselves, whereas CSTs tend to be designed with a larger user group in mind. However, the comparison yields some insight into what areas may be under-explored by creativity support researchers.

3.1 Methodology

To find a set of papers that supported some kind of CST for creative writing, we first searched for "creative writing" in the ACM Digital Library and collected all matching papers published between January 2011 and December 2020. This resulted in 223 papers. We then manually inspected each paper to check if it presented a tool either explicitly developed to support creative writing or was evaluated in a creative writing support setting. This resulted in 35 relevant papers. Finally, we coded each paper to see if it conformed to our definition of a Poetry Machine (as before, via consensus coding.) We used the same definition as outlined in Section 2.2.2. This left us with 14 papers [3, 5–7, 9–18]. These papers were then coded in the same manner as the elicited Poetry Machines were, allowing us to directly compare the elicited Poetry Machines and CSTs.

3.2 Findings

Here we describe some key differences between the CSTs and the Poetry Machines based on the qualitative coding as well as more holistic comparison:

Mode. We confirmed our suspicion that the Poetry Machines experimented with multiple media more so than the CSTs. While only 5 out of 14 (36%) of the CSTs included some kind of multimedia element (visually laying text out in an unexpected manner, or making use of images or sound), 13 out of 15 (87%) of the Poetry Machines did.

Complication. As stated previously, two of the Poetry Machines were "laundry lists" of more than ten heterogeneous features. None of the CSTs approached this kind of multifariousness.

Oddness. Our impression of the Poetry Machines is that they contained more features that were surprising—idiosyncratic, dramatic, or seemingly-illogical, perhaps similar to the way that poetry itself can defy logic. (This may also be related to the way we framed the design task; if CST research implicitly requires researchers to justify their designs, the poets in our study were free to describe what they wanted without any obligation to rationalize it.) We found it noteworthy that none of the poets imagined a straightforward "autocomplete" system that would furnish them with a plausible (i.e. nonrandom) next sentence (as in [16]). In fact, one poet actually requested a mischievous reversal of an autocomplete function: this poet's machine would insert an intentional non-sequitur sentence rather than a well-chosen and coherent one.

4 CONCLUSION AND FUTURE WORK

Our study suggests that researchers' notions of what kind of assistance writers may want—or even what "assistance" means in the context of creative writing—seems to differ from those self-reported by professional poets. In the future we hope to investigate more closely the poets' preexisting views on poetry, including whether there are regularities in the ways that poets of different stylistic persuasions (e.g. "formalist" vs. "experimental") design their Poetry Machines. As part of a more extensive participatory design process, we also plan to build several of these Poetry Machines in further

conversation with the poets who dreamt them up and to study how these poets make use of them over time.

REFERENCES

[1] Jeffrey Bardzell and Shaowen Bardzell. 2016. Humanistic Hci. *Interactions* 23, 2 (2016), 20–29.

[2] Kyle Booten. 2019. Toward digital *progymnasmata*. In *Proceedings of the 10th International Conference on Computational Creativity*. 1–8.

[3] Elizabeth Clark, Anne Spencer Ross, Chenhao Tan, Yangfeng Ji, and Noah A Smith. 2018. Creative writing with a machine in the loop: Case studies on slogans and stories. In *23rd International Conference on Intelligent User Interfaces*. 329–340.

[4] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* (2015), 222–248.

[5] Richard P Gabriel, Jilin Chen, and Jeffrey Nichols. 2015. InkWell: A Creative Writer's Creative Assistant. In *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*. 93–102.

[6] Katy Ilonka Gero and Lydia B Chilton. 2019. How a Stylistic, Machine-Generated Thesaurus Impacts a Writer's Process. In *Proceedings of the 2019 on Creativity and Cognition*. 597–603.

[7] Katy Ilonka Gero and Lydia B Chilton. 2019. Metaphoria: An algorithmic companion for metaphor creation. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. 1–12.

[8] Marjan Ghazvininejad, Xing Shi, Jay Priyadarshi, and Kevin Knight. 2017. Hafez: an interactive poetry generation system. In *Proceedings of ACL 2017, System Demonstrations*. 43–48.

[9] Frederica Gonçalves, Ana Caraban, Evangelos Karapanos, and Pedro Campos. 2017. What shall i write next? Subliminal and supraliminal priming as triggers for creative writing. In *Proceedings of the European Conference on Cognitive Ergonomics 2017*. 77–84.

[10] Philip Heslop, Ahmed Kharrufa, Madeline Balaam, David Leat, Paul Dolan, and Patrick Olivier. 2013. Learning extended writing: designing for children's collaboration. In *Proceedings of the 12th International Conference on Interaction Design and Children*. 36–45.

[11] Chieh-Yang Huang, Shih-Hong Huang, and Ting-Hao Kenneth Huang. 2020. Heteroglossia: In-situ story ideation with the crowd. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[12] Chris Kim, Uta Hinrichs, Saif M Mohammad, and Christopher Collins. 2020. Lexichrome: Text Construction and Lexical Discovery with Word-Color Associations Using Interactive Visualization. In *Proceedings of the 2020 ACM Designing Interactive Systems Conference*. 477–488.

[13] Joy Kim, Justin Cheng, and Michael S Bernstein. 2014. Ensemble: exploring complementary strengths of leaders and crowds in creative collaboration. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 745–755.

[14] Max Kreminski, Melanie Dickinson, Michael Mateas, and Noah Wardrip-Fruin. 2020. Why Are We Like This?: The AI architecture of a co-creative storytelling game. In *International Conference on the Foundations of Digital Games*. 1–4.

[15] Timothy Neate, Abi Roper, Stephanie Wilson, and Jane Marshall. 2019. Empowering expression for users with aphasia through constrained creativity. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–12.

[16] Melissa Roemmele and Andrew S Gordon. 2018. Automated assistance for creative writing with an rnn language model. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces Companion*. 1–2.

[17] Sarah Sterman, Evey Huang, Vivian Liu, and Eric Paulos. 2020. Interacting with Literary Style through Computational Tools. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12.

[18] Gavin Wood, John Vines, Madeline Balaam, Nick Taylor, Thomas Smith, Clara Crivellaro, Juliana Mensah, Helen Limon, John Challis, Linda Anderson, et al. 2014. The department of hidden stories: Playful digital storytelling for children in a public library. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1885–1894.

A SURVEY QUESTIONS

- What is your name?
- Who are some of the poets who have influenced your own writing the most?
- Do you consider yourself to be working in any poetic traditions? If so, which ones?
- What does writing a poem usually look like for you? (Do you write on paper? On a computer? If so, what software?)
- Do you ever write in "tradition" [*sic*] forms that specify a certain rhyme or meter (e.g. the sonnet)?
- Do you ever write using "experimental" techniques (e.g. collage, randomness, or erasure)?
- If your answer to the above was not "never," please give an example of an experimental technique you have used.
- Do you ever used [*sic*] computers or digital media in your poetic process, other than for typing a poem in a program like Microsoft Word? (For example: random word generators for writing prompts, or gathering found language from Twitter.)
- If your answer to the question above was not "never," please describe how you have used digital media in your poetic process.
- For you, what is the point of poetry?
- Please describe any previous experience with programming, user interface design, or other computational fields.
- What do you know about "machine learning"? How does it work?
- Do you think "machine learning" or computer programs in general could help you write a poem? If so, how? If not, why?
- When we report on this study in academic publications, do you want to be referred to using your real name or a pseudonym?

B EXAMPLE POETRY MACHINES PRESENTED TO PARTICIPANTS

B.1

Whenever I don't input words for 2 minutes (*trigger*), the Poetry Machine should examine the last sentence (*focus*), invent a sentence that would make sense to follow it and that sounds like Romantic poetry (*think*) and add it directly to the end of the poem (*act*).

B.2

Every time I write a noun, take that word and have the machine find three other nouns that are blue and might be related and display them for me. For example, if I write "There was a bird" the machine might return "bluejay [*sic*], robin's egg, owl eyes". Whenever I click on a word, take that word and use it to search for paintings whose main color is blue and whose titles contain this word. Display one of these paintings at the end of the poem. If I ever type the word blue, automatically replace it with a more specific blue color word, like turquoise or indigo.