language models as necromancy

Allison Parrish

what is a language model?

- A statistical model, trained on a data set of text

- Text is divided into *tokens*, which can be any linguistic unit (typically words or characters)

- A language model assigns a probability to a list of tokens, answering the question, "how likely is this sequence of words?" (or characters, or phonemes, or sentences...)

- A language model can *predict* the next token in a sequence. (Given the current context, what is the most likely token to follow?)

- There are many ways to implement language models! Markov chains, RNNs, Transformers (like GPT-2), etc. Each has benefits and drawbacks.

A simple n-gram Markov chain

- An *n-gram* is a sequence of tokens with a fixed length. (n=1 is a unigram, n=2 is a bigram, n=3 is a trigram, etc.) An *n-gram analysis* of a text consists of finding and counting all of the unique n-grams.

- A Markov chain is a statistical model in which the next step of some sequence can be predicted from the previous step. When there are multiple possible next steps, each step is assigned a probability based on the probability of that step in the data.

- In an n-gram Markov chain, the next token (letter, word) in a sequence is predicted based on the n-gram that precedes it.

An example: condescendences

n-gram	next?
co	n
on	d
nd	e, e
de	s, n
es	c, (end of text)
sc	e
ce	n, s
en	d, c
nc	e

co
con
cond
conde
conden
condend
condendes
condendesc
condendesce
condendesces

From Hartman, *MONOLOGUES OF SOUL & BODY*

N = 4

Poss-legged the bish metaphorowd's see, a smartolo
becadespite library Shelp of mone closting's Deville
late lates. Luck meton, yournament of human tourname
Inter, says Napollect as to plurate buildingenia;
Isouard enormous. Last gament on tournage opedifficians
of perman edifieserves in his unity.
at at two rooking, viole world, and, and Reason shad
to be snow? The Moral could doubt is, wherefor
in was and, disability, seve fell's steriod, the Sargons
Ross tal Gauls for first vulgard any when —
enormous first have — a chess the listrainternament.

N=9

Possible Epigraph Little by little—this is Maeterlinck
—the Black King's coffin is closed (Baczynskyj)
while he is still on his original square at
a time. Candidates Suppose a white male et cetera
at one corner of the royal chapel where Louis heard
Mass each morning, waiting beforehand in the equation;
at the other a woman, you, the President, Christ,
Rossini, Kieseritsky was eliminated in this
very first game, though stronger than many players
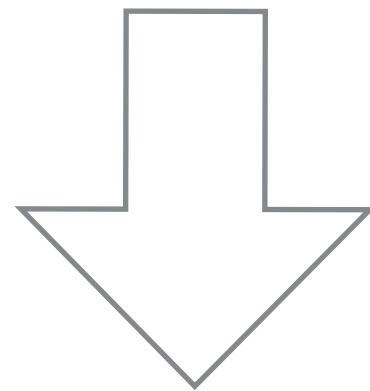who placed ahead of him in the old days, except

automatic writing vs automated writing

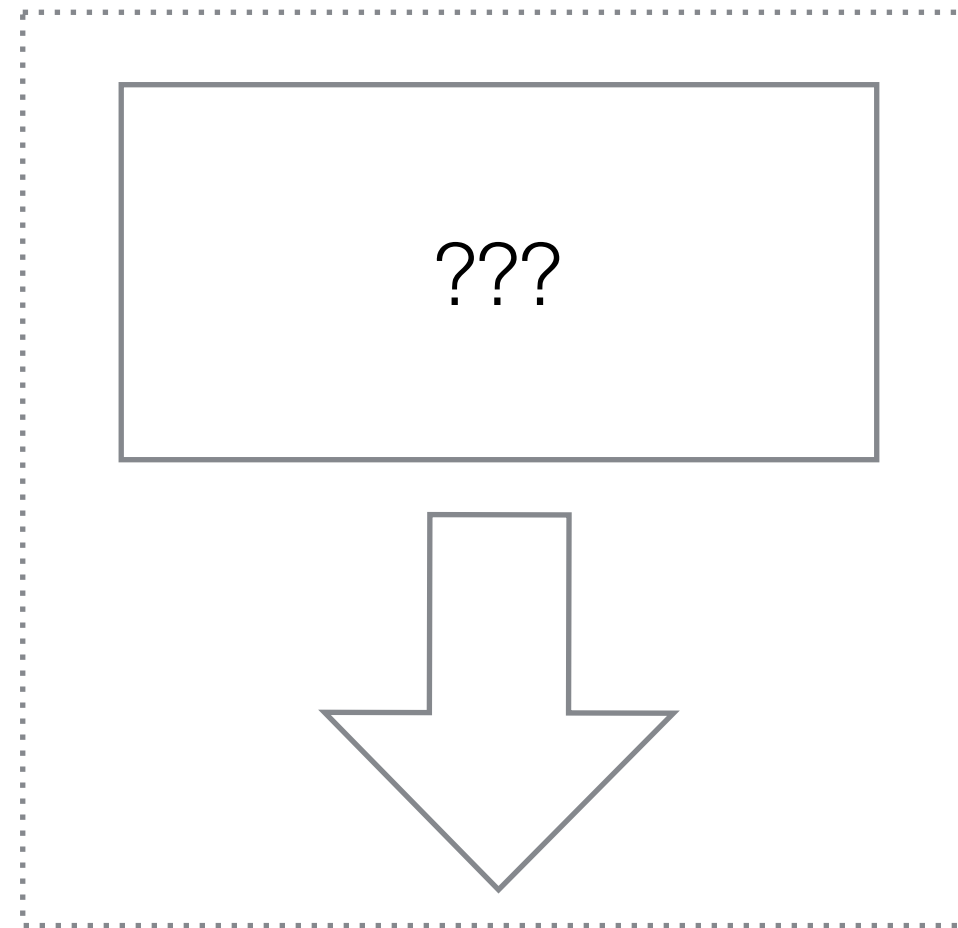## typical writing

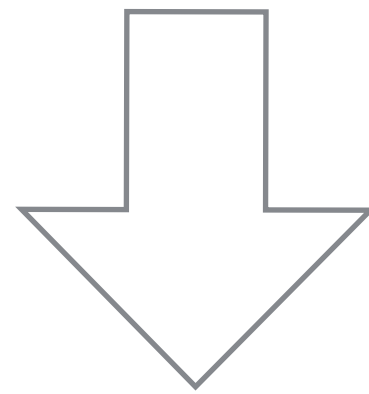| intentional thought |
| :---: |

⬇

hand movement

⬇
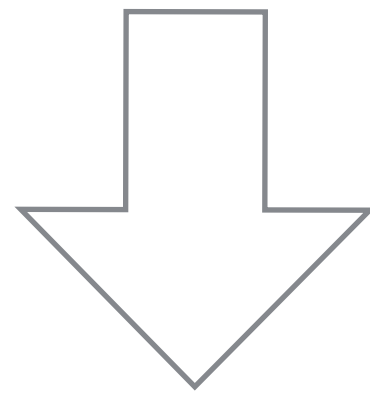
marks applied to paper

automatic writing

???

⬇

hand movement

⬇
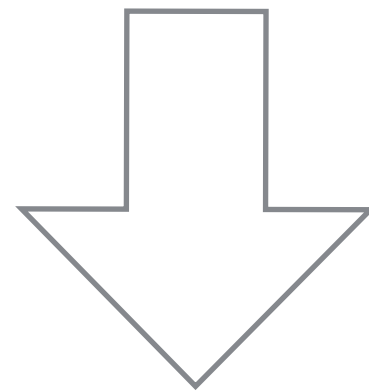
marks applied to paper

automated writing

???

⬇

???

⬇

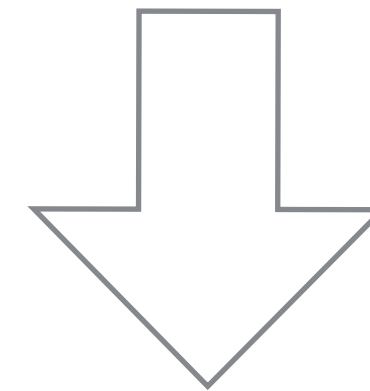| ??? |
| :---: |
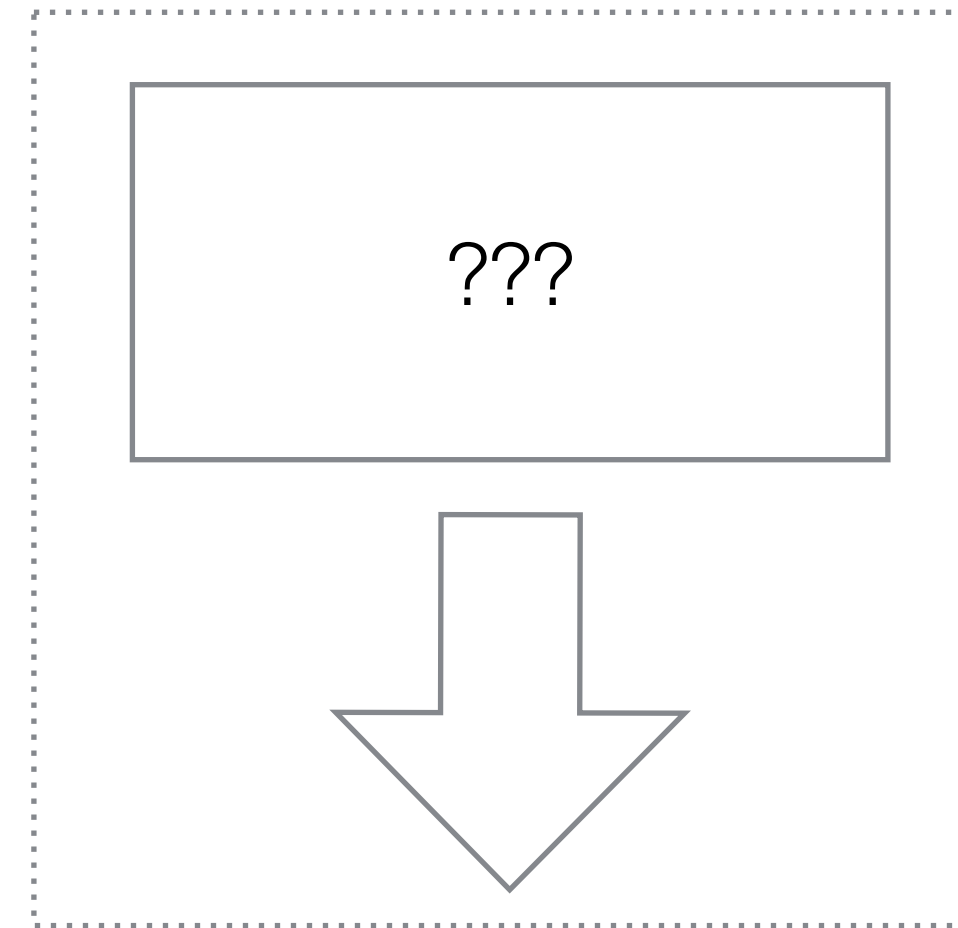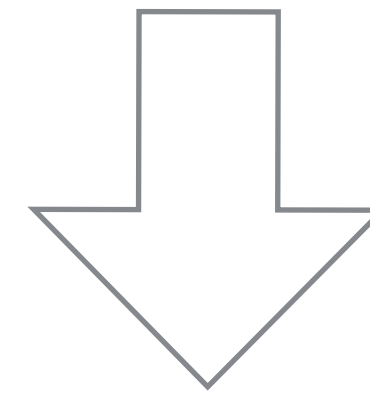
aesthetic similarities between automatic and automated writing (one example)

"A-drip, and drops a-slide 'pon stone walls to pool aneath. A-chill the moon air and mist doth hang 'bout hill like white smock 'bout the shoulders of a wench. Smudge-scant upon the air, and brown fat reeking from crack o' door. A grunt, a shuffle, and door doth ope, and Telka wriggleth bare toes in pool-drip."

Pearl Curran/Patience Worth, quoted in Braude, Stephen E. "Dissociation and Latent Abilities: The Strange Case of Patience Worth." Journal of Trauma & Dissociation, vol. 1, no. 2, June 2000, pp. 13–48.

Its slates, like wings well, it's strange, all is strange, well, that is strange: and flats that (when it rains) as sweet as when in distant years nigh twenty years since it was new — the same strange sweetness which in years 'Tis twenty years and more, since here, twenty years hence.

Allison Parrish, *Articulations: A Fragment*
https://itp.nyu.edu/adjacent/issue-3/articulations-a-fragment-fragment-fragment/

the "control"

- In automatic writing, the writer's intentional thought is not what sets the hand in motion. Instead, the writer's body serves as a medium for some other instigating force.

- Controls in automatic writing: spirits, aliens, the primal subconscious, the right hemisphere of the brain, etc.

- But who or what can be understood as the "control" in computer-generated writing? Who or what is the *medium* of that control, and (construing these terms to include digital media) how does this medium place "marks" on the "page"?

nearly all digitally-mediated writing interfaces are "under control"
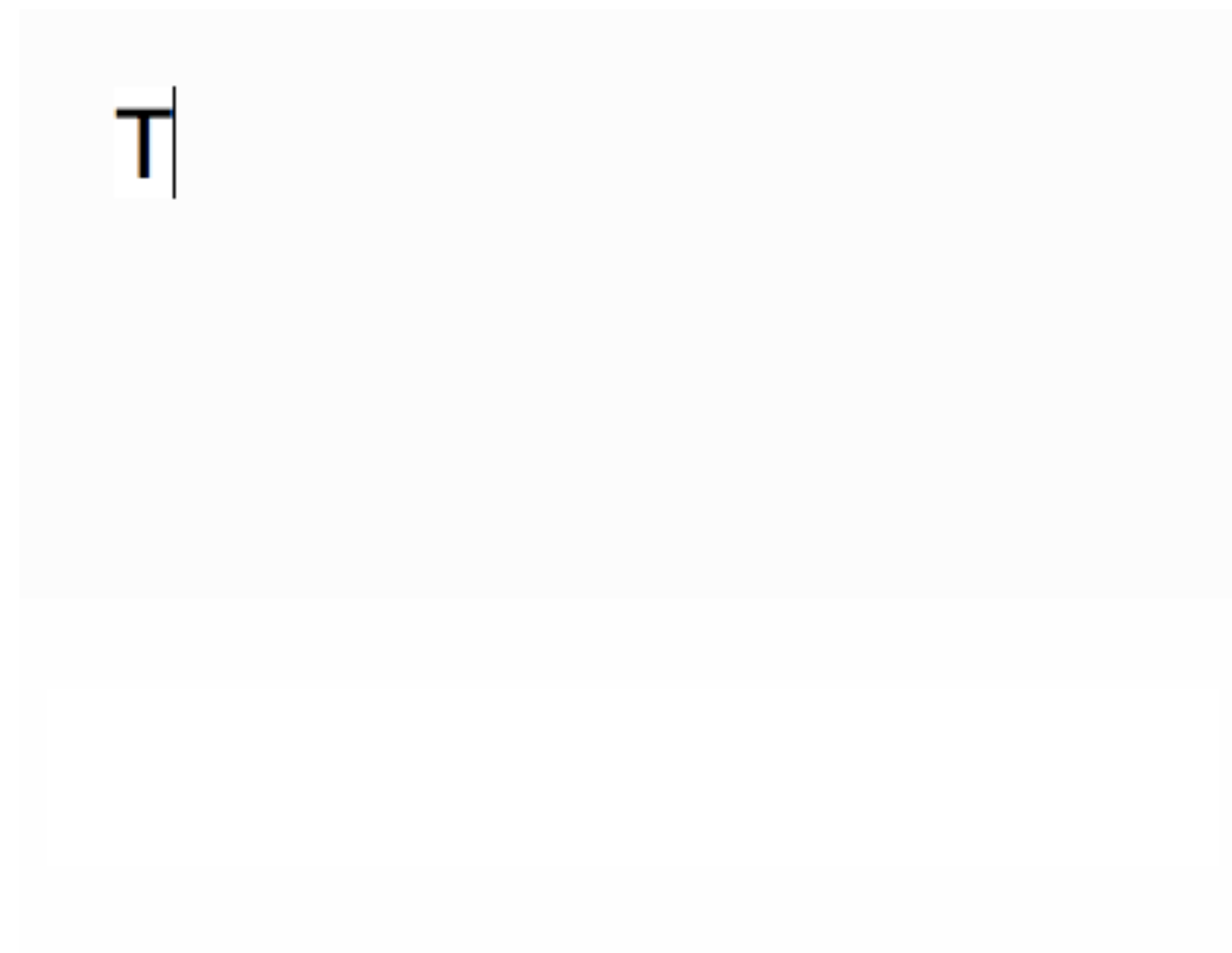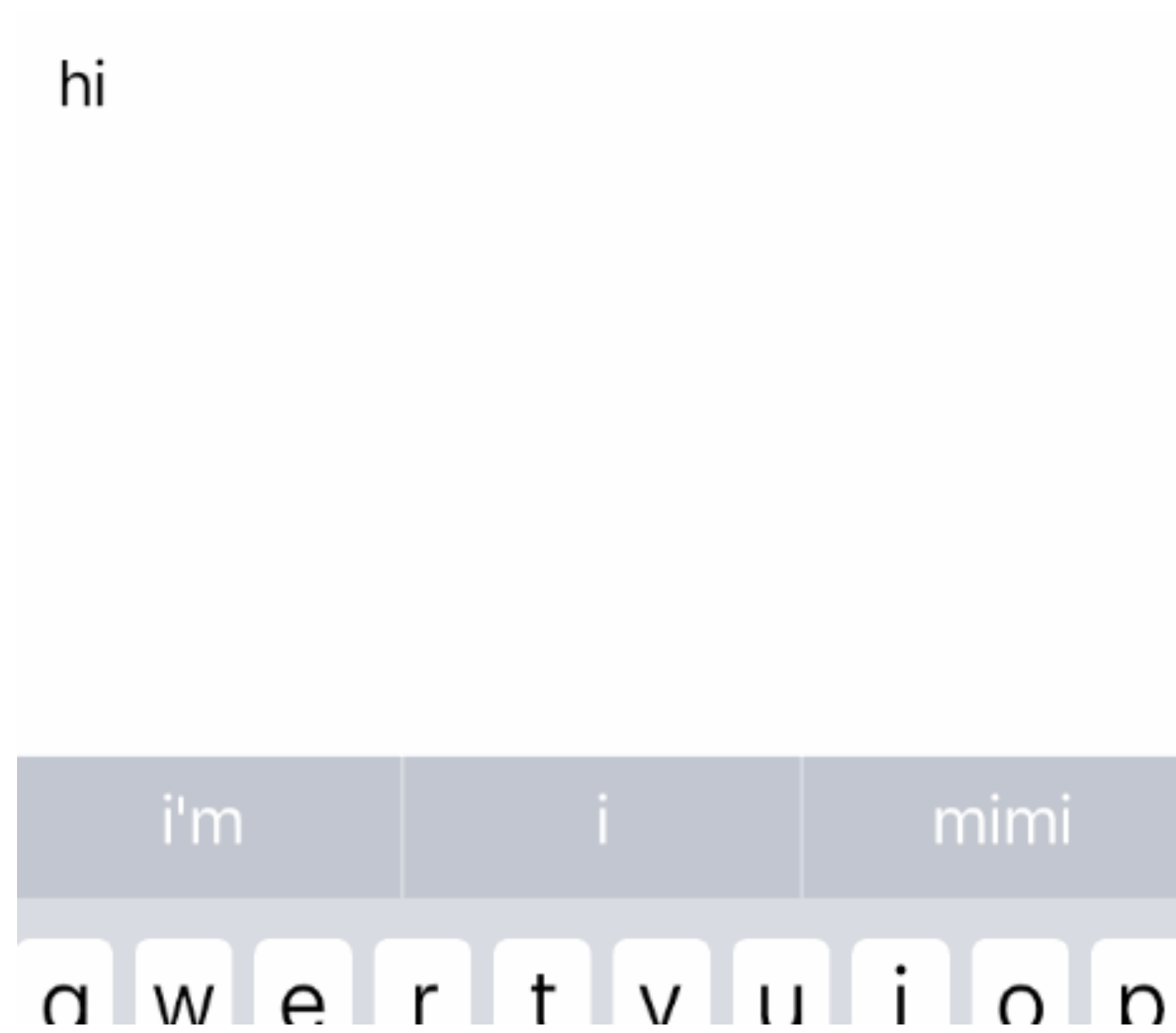
in these interfaces, *data* is the control

data as control

- In automatic writing, the control is understood to be ineffable, transcendent, unknowable, accessible only by side-stepping the logical mind.

- Data is exactly the opposite of that. By definition, data is knowledge that is available to the logical mind; it is exactly the knowledge that *can* be quantified.

data as an artifact of embodiment

- Data is an abstraction of phenomena it records, but still remains an *artifact* of embodiment.

- A text data set is called a *corpus*—literally "body."

- Data directly results from the ideology of those that gather it, and it records the physical and cultural contexts of its collection. (See Onuoha, Mimi. "The Point of Collection." Data and Society: Points, 31 Oct. 2016, https://points.datasociety.net/the-point-of-collection-8ee44ad7c2fa.)

- Texts—like any dataset—are deeply contingent on social, physical and temporal context.

"The text-artifact does indeed have a temporal structure... precisely because it was originally laid down... in the course of a social process, unfolding in real time."

Silverstein, Michael, and Greg Urban, editors. *Natural Histories of Discourse*. University of Chicago Press, 1996.

"[W]riting... is in fact the very stuff, substance, form of inscription that history and the situation of enunciation enter into the linguistic process. [...] [T]he material fact of history is always a part of any written text."

Drucker, Johanna. *The Visible Word: Experimental Typography and Modern Art, 1909-1923*. University of Chicago Press, 1994.

Every text is a transcription. A corpus can be understood as a literal "body" of linguistic actions, bound up with their historical contexts. A predictive model is a literal attempt to re-animate this body.

Narcissus and machine learning

language models and verisimilitude

- Academic and industrial research on machine learning models of language reveals a narrow focus on verisimilitude.

- In these fields, the value of a model lies in how it performs on certain benchmarks, many of which are comparisons to so-called "human" competence.

- The implied goal: a world in which linguistic artifacts deemed "computer generated" will be indistinguishable from linguistic artifacts deemed "written by humans."

The pool paints Narcissus, and the painting represents both the pool and the whole story of Narcissus. [...] Narcissus, it is no painting that has deceived you... you do not realize that the water represents you exactly as you are when you gaze upon it, nor do you see through the artifice of the pool... but acting as though you had met a companion, you wait for some movement on his part. Do you then expect the pool to enter into conversation with you? (1.23)

Philostratus the Elder, et al. Philostratus the Elder, Imagines. Philostratus the Younger, Imagines. Callistratus, Descriptions. Translated by Arthur Fairbanks, Harvard University Press, 1931.

[W]hy, o foolish boy, do you persist? Why try to grip an image? [. . .] The face you discern is but a shadow, your reflected form. That shape has nothing of its own. . . it will retreat when you have gone—if you can ever leave! (III.430–434)

Ovid. The Metamorphoses. Translated by Allen Mandelbaum, Alfred A. Knopf, 2013.

Max Nelson's reading of the Narcissus myth

- Not a fable about the dangers of solipsism, but instead a **description of a scrying ritual** similar to those recorded in various ancient magical papyri.

- In these rituals, a young boy would "stare motionless at his own reflection in calm, clear, resplendent spring water. . . while narcotic and chthonic flowers were burned. . . in order to summon an apparition (and even the voice) of a deceased individual" (383).

- In the original Latin text in Metamorphoses—translated by Mandelbaum (quoted above) as "the face you discern is but a shadow"—is *quam cernis, imaginis umbra est*. "The word umbra... could mean both **shadow or a ghost**, and similarly imago could refer to a likeness or... to the disembodied dead."

- Elsewhere, these words are "specifically used in the context of water divination involving invoking the dead" (378–379)

Nelson, Max. "Narcissus: Myth and Magic." *The Classical Journal, vol. 95, no. 4, 2000, pp. 363–390.*

In Nelson's interpretation, Narcissus never believes his reflection to be indistinguishable from his own person. Instead, **the experience of reflection is used to evoke a voice otherwise hidden from conscious thought**. Narcissus is, in other words, acting as a medium for a kind of automatic writing.

To some extent, all data-driven endeavors function as mirrors that only show us our own faces. Certain works of art and design, however, demonstrate Nelson's more hopeful interpretation of the Narcissus myth.

Technologies intended by their creators to reinforce a solipsistic status quo are détourned, and used instead in elaborate automatic/automated writing rituals, drawing forth the spirits buried in data, in order to find out what they have to say to us.