令和 6 年度入学試験問題

総 合 問 題 （人文学部）

注 意 事 項

1　この問題冊子は，試験開始の合図があるまで開いてはならない。

2　問題冊子は，全部で 4 ページある。（落丁，乱丁，印刷不鮮明の箇所などがあった場合は申し出ること。）　問題冊子の中に下書き用紙が 2 枚入っている。

3　別に解答用紙が 3 枚ある。

4　**解答は，すべて解答用紙の指定された箇所に記入すること。**

5　受験番号は，各解答用紙の指定された 2 箇所に必ず記入すること。

6　解答時間は，120 分である。

7　問題冊子及び下書き用紙は，持ち帰ること。

問題　次の英文は，大規模言語モデル（large language model）の台頭以後の文章の執筆（writing）について書かれた記事の一部である。これを読んで，後の**問 1** ～**問 5** に日本語で答えなさい。

A lot of uses have been proposed for large language models.[1] Thinking about them as blurry JPEGs[2] offers a way to evaluate what they might or might not be well suited for. Let's consider a few scenarios.

Can large language models take the place of traditional search engines? For us to have confidence in them, we would need to know that they haven't been fed propaganda and conspiracy theories. But, even if a large language model includes only the information we want, there's still the matter of blurriness. There's a type of blurriness that is acceptable, which is the restating of information in different words. Then there's the blurriness of outright fabrication, which we consider unacceptable when we're looking for facts. It's not clear that it's technically possible to retain the acceptable kind of blurriness while eliminating the unacceptable kind, but I expect that we'll find out in the near future.

Even if it is possible to restrict large language models from engaging in fabrication, should we use them to generate Web content? This would make sense only if our goal is (A) to repackage information that's already available on the Web. Some companies exist to do just that — we usually call them content mills. Perhaps the blurriness of large language models will be useful to them, as a way of avoiding copyright infringement. Generally speaking, though, I'd say that anything that's good for content mills is not good for people searching for information. The rise of this type of repackaging is what makes it harder for us to find what we're looking for online right now; the more that text generated by large language models gets published on the Web, the more the Web becomes a blurrier version of itself.

(B) Can large language models help humans with the creation of original

writing? To answer that, we need to be specific about what we mean by that question. There is a genre of art known as Xerox[3] art, or photocopy art, in which artists use the distinctive properties of photocopiers as creative tools. Something along those lines is surely possible with the photocopier that is ChatGPT, so, in that sense, the answer is yes. But I don't think that anyone would claim that photocopiers have become an essential tool in the creation of art; the vast majority of artists don't use them in their creative process, and no one argues that they're putting themselves at a disadvantage with that choice.

So let's assume that we're not talking about a new genre of writing that's analogous to Xerox art. Given that stipulation, (C) can the text generated by large language models be a useful starting point for writers to build off when writing something original, whether it's fiction or nonfiction? Will letting a large language model handle the boilerplate[4] allow writers to focus their attention on the really creative parts?

Obviously, no one can speak for all writers, but let me make the argument that starting with a blurry copy of unoriginal work isn't a good way to create original work. If you're a writer, you will write a lot of unoriginal work before you write something original. And the time and effort expended on that unoriginal work isn't wasted; on the contrary, I would suggest that it is precisely what enables you to eventually create something original. The hours spent choosing the right word and rearranging sentences to better follow one another are what teach you how meaning is conveyed by prose. (D) Having students write essays isn't merely a way to test their grasp of the material; it gives them experience in articulating their thoughts. If students never have to write essays that we have all read before, they will never gain the skills needed to write something that we have never read.

And it's not the case that, once you have ceased to be a student, you can safely use the template that a large language model provides. The struggle to express your thoughts doesn't disappear once you graduate — it can take place

every time you start drafting a new piece. Sometimes it's only in the process of writing that you discover your original ideas. Some might say that the output of large language models doesn't look all that different from a human writer's first draft, but, again, I think this is a superficial resemblance(E). Your first draft isn't an unoriginal idea expressed clearly; it's an original idea expressed poorly, and it is accompanied by your amorphous dissatisfaction, your awareness of the distance between what it says and what you want it to say. That's what directs you during rewriting, and that's one of the things lacking when you start with text generated by an A.I.

There's nothing magical or mystical about writing, but it involves more than placing an existing document on an unreliable photocopier and pressing the Print button. It's possible that, in the future, we will build an A.I. that is capable of writing good prose based on nothing but its own experience of the world. The day we achieve that will be momentous indeed — but that day lies far beyond our prediction horizon. In the meantime, it's reasonable to ask, What use is there in having something that rephrases the Web? If we were losing our access to the Internet forever and had to store a copy on a private server with limited space, a large language model like ChatGPT might be a good solution, assuming that it could be kept from fabricating. But we aren't losing our access to the Internet. So just how much use is a blurry JPEG, when you still have the original?

［注］

1　large language model　ChatGPT などで用いられる，大量のテキストデータとディープラーニング技術を用いて構築された言語モデル。

2　JPEG　画像データの圧縮方式。圧縮された画像ファイルそのものを指す言葉としても用いられる。blurry JPEG（ぼやけた JPEG）という表現は，圧縮によるデータ量の削減によってぼやけた JPEG 画像のことを意味している。本文では明示的に述べられていないが，ウェブ上の文章と大規模言語モデルの出力する文章との関係は，元の画像とぼやけた JPEG 画像との関係に相当するものだというのが著者の見解である。

3　Xerox　コピー機の名称。それを用いて作成されたものを指す場合もある。

4　boilerplate　定型文や標準的な表現のこと。

Ted Chiang, “ChatGPT Is a Blurry JPEG of the Web,” *The New Yorker*, February 9, 2023, https://www.newyorker.com/tech/annals-of-technology/chatgpt-is-a-blurry-jpeg-of-the-web　より，抜粋。（なお，一部を省略している）

問 1　下線部Ⓐについて，大規模言語モデルが「情報を再パッケージ化すること（to repackage information）」を繰り返すと，ウェブはどのように変化するのか。コンテンツミル（content mills）について説明しつつ，本文に即して述べなさい。

問 2　下線部Ⓒを和訳しなさい。

問 3　下線部Ⓓを和訳しなさい。

問 4　下線部Ⓔについて，なぜ類似は表面的と言えるのか，本文に即して説明しなさい。

問 5　第四段落冒頭の下線部Ⓑにある著者の問いについて，本文の内容を踏まえつつ，あなたの議論を展開させなさい。