

令和6(2024)年度  
一般選抜公立大学中期日程 試験問題  
情報科学部 情報科学科

# 小論文

## 注意事項

- 1 試験開始の合図があるまで、この問題冊子を開かないでください。
- 2 解答はすべて解答用紙に記入してください。
- 3 解答には鉛筆かシャープペンシルを使用してください。
- 4 問題は全部で4ページ、解答用紙は全部で1枚あります。
- 5 試験時間は60分です。
- 6 試験終了後、問題冊子も回収します。
- 7 何か伝えたいことがあるときは挙手してください。

問題 次の資料を読み、各問に答えなさい。

#### 資料1

私が学生の時、今日のコンピューターには二つの機能があると習いました。一つは「非常に高速に計算すること」。この機能に関する研究を象徴するのが京や富岳のようなスーパーコンピューターで、いかに高速に計算するコンピューターを作るかを目指す研究です。

もう一つは「無限ともいえるほど膨大な情報を記憶し、管理すること」です。私はこの二つ目の機能に注目し、データの管理に関する研究に40年近く取り組んできました。こちらの研究のフォーカスは、データをどう集め、どう管理し、それを柔軟に利用可能にするか。対象はマシンパワーではなく「データ」です。

- 中略 -

新型コロナウイルス（COVID-19）の流行によって大量のデータを瞬時に全国で共有することの重要性はますます明確になってきています。たとえば、感染が疑われる患者さんのCT撮影をする時に、CTを撮ってすぐに新型コロナウイルス肺炎か否かが分かれば、その患者さんがCT検査室を出る前に、一般病棟へ連れて行くか、専用病棟かを判断できます。そのためにはAIによる迅速かつ的確な画像診断が不可欠ですが、CT画像はレントゲン画像と違って一人につき多数の画像がありますから、膨大な量のデータをやりとりしなくてはなりません。しかし国立情報学研究所が日本全国の大学・研究機関等を結ぶ情報通信ネットワークとして構築・運用しているSINETを使えば、全国から大量のCT画像が集まりますし、そのデータを使ってAIによる画像判断を瞬時に行うことができます。国立情報学研究所は放射線学会と連携して医療画像のプラットフォームも構築しました。100名以上の研究者が登録して利用していますが、コロナ肺炎は名古屋大学の森健策先生のチームがいち早く開発され、約83%の精度を達成しています。データが増えればもっと精度は上がるでしょう。

このようなデータ駆動型のシステムが有効なのは医療だけではなくありません。私はどんな研究もデータ駆動に変えていくべきだと思っています。例えばスポーツ科学などもいまや、データがなければ始まりません。体操競技などではAIが採点支援をするようになってきています。練習のときに演技を動画で撮影すればAIがその人に必要な練習を提案してくれる、というシステムも開発されています。あらゆる分野がこのようにデータから始まるようになっていくと見ています。

しかしこの先が難しいところで、個々の研究者がデータを他の人と共有しない限り、データはなかなか増えていきません。データが大量にないと、実用に足るAIを構築することもできません。今のところ、大型研究は別として、個人個人の小さな研究プロジェクトでは、苦労して集めたデータは自分の財産であり、人と共有するものではない、という考え方が支配的です。とくに人文学や社会科学ではデータ共有に抵抗がある研究者が現時点では多いで

しょう。共有を拒絶する習慣が続きますと学問全体がデータ駆動になっていかないわけですが、今後、研究者のマインドが大きく変容していったほしいと期待しています。

とはいえ、いま、メジャーな論文誌では論文提出の際、データの提出も求めるところが増えてきています。研究に対する評価は「集めたデータ」ではなくそのデータで何を見出したかで決まる、データは学問の世界の共有財産である、という考え方に踏み出そうとしています。データが共有されれば、そのデータを今度は別の研究者が別の視点で研究に用い、新たな発見をすることができる。つまりデータには、研究成果以上に「共有される価値」があるともいえるのです。

「ネットワーク基盤とデータベース基盤の融合によってデータがアカデミアで共有され、そこから新たな研究がどんどん生まれてくる世界を作ること、すなわち『データ駆動型学術』が進展すること」が私の夢です。著名なコンピューター科学者であった故ジム・グレイ博士は科学的発見の歴史的変遷を4つのパラダイムの変遷であるとされました。まず観測に基づく科学、次に17世紀ごろからの理論に基づく科学、そして20世紀半ばからの数値計算に基づく科学です。スパコンはまさに第3となる計算科学の象徴でしょう。そしてデータが中核になる科学が第4のパラダイムであると解いておられます。データ駆動型学術を進めていくことこそ21世紀に私たちがなすべきことなのです。

データ駆動型に変わっていくのはアカデミアだけではありません。社会も大きく変わりつつあります。

私の研究室には、マラリアで苦しむ国民を抱えたアフリカの国からリアルタイムでデータが来ます。これを用いて、マラリアの突発的発生や流行の予測をする支援を行っています。これまでは物資や資金の援助が中心だった国際支援のやり方も、データが変えていくのです。経済的に脆弱になりつつある今の日本に、途上国に提供できる潤沢な資金は残念ながらそれほどはありません。しかし、データ駆動型の社会になることによって、日本は「知恵」で途上国を助けることができるのです。

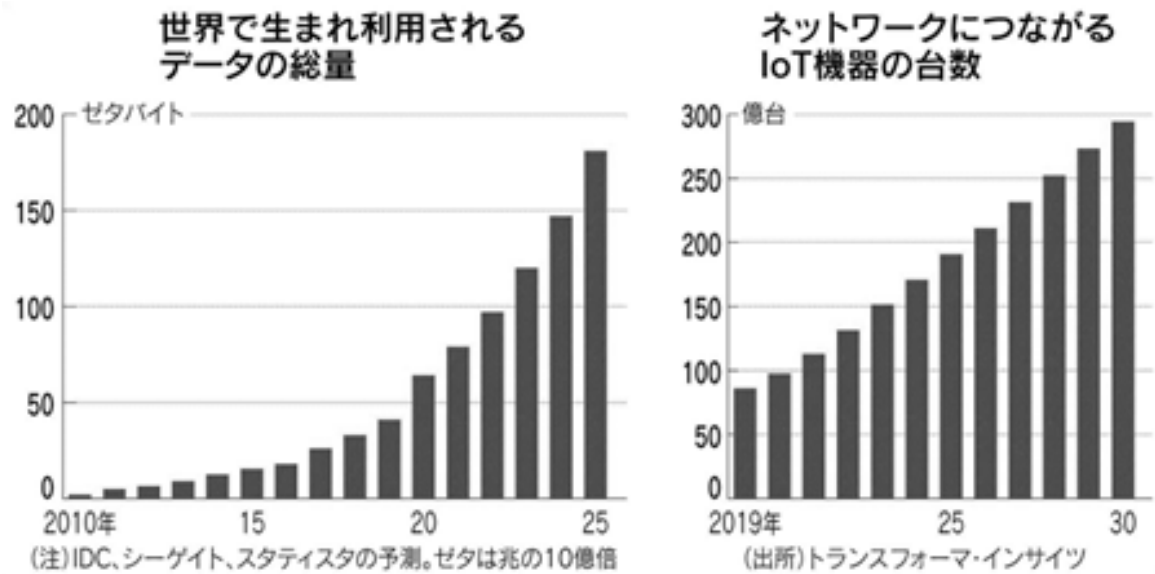
世界がデータ駆動型に変容してゆきますと、結局はマシンパワーをもつ先進国やGAFAsのような超大型IT企業がデータ集約や解析において優位に立ち、持つ者と持たざる者の格差が広がるのではないかという指摘もあります。たしかにそういう将来像もありえるでしょう。しかし私は今のところ、マシンパワーに対するデータの優位性はゆるがないと見ています。なぜなら、マシンは買えますが、たとえば国民全体の長期にわたる健康調査のようなデータは買えるものではないからです。その証拠に、Googleは機械学習などAIのプラットフォームを無料で公開しています。重要なのは解析能力の独占より、そのプラットフォームに集まってくるデータだ、と彼らも考えているのではないのでしょうか。課題があり、それを解くためにどんなデータが必要か、それをどう集めればいいのか、これがポイントです。AIの手法そのものはほとんどがオープンです。学問の世界でも産業の世界でも、そうしたことを考える「データの設計学」のような分野が生まれると考えています。この学問が最もエキサイティ

ングになるでしょう。

出典：東京大学未来ビジョンセンター（編）『未来探究 2050 東大 30 人の知性が読み解く世界』、日経 BP 日本経済新聞出版、2021、P.178-186、「社会をデータ駆動に変えるデータ工学」喜連川優

本文は出題者により一部改変されています。

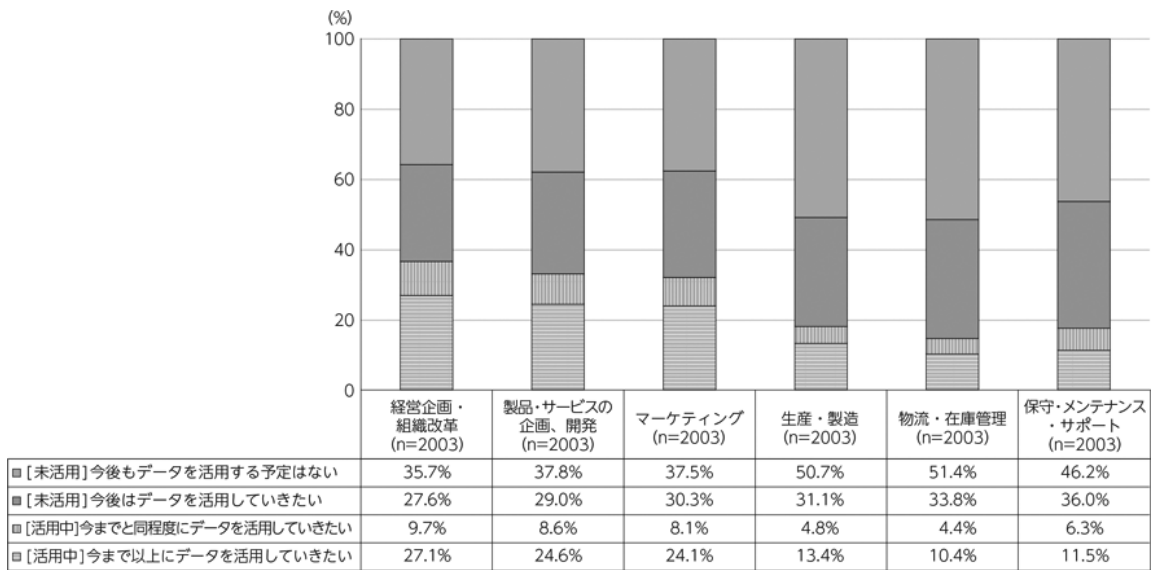
## 資料 2



出典：日本経済新聞電子版 2023 年 1 月 6 日「データ量爆発、25 年に 2 倍」

<https://www.nikkei.com/article/DGKKZO67338770V00C23A1TEC000/>

資料 3



(出典) 総務省 (2020) 「デジタルデータの経済的価値の計測と活用の現状に関する調査研究」

日本の企業における 3 から 5 年先のそれぞれの領域におけるデータ活用予定

出典：総務省『令和 2 年版 情報通信白書』、図表 3-2-1-10 「今後のデータ活用予定」

<https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r02/pdf/index.html>

問 1

多くの研究者のデータ共有に対する認識の変化と、データ駆動型によってどのような新たな価値が生まれると考えられるか。資料 1 を参考に 200 字以内でまとめなさい。

問 2

資料 2 は世界で生まれ利用されるデータの総量及びネットワークにつながる IoT 機器の台数に関して、それぞれの年ごとの推移と今後数年間の見込みを表したものである。資料 3 は日本の企業において 3 から 5 年先の各業務領域におけるデータ活用の意向を示したものである。資料 2 も参考にしながら、資料 3 のグラフを分析したときにわかる問題点は何かを 200 字以内で論じなさい。

問 3

資料 1 の最後で “「データの設計学」のような分野が生まれると考えています” とあるが、これはどのようなものであり、その発展にとって重要なことは何か。すべての資料を参考に 400 字以内で自分の考えを含めて述べなさい。