

## How AI can identify people even in anonymized datasets

*Weekly social interactions form unique signatures that make people stand out*

How you interact with a crowd may help you stick out from it, at least to artificial intelligence.

When fed information about a target individual's mobile phone interactions, as well as their contacts' interactions, AI can correctly pick the target out of more than 40,000 anonymous mobile phone service subscribers more than half the time, researchers report January 25 in *Nature Communications*. The findings suggest humans socialize in ways that could be used to pick them out of datasets that are supposedly anonymized.

It's no surprise that people tend to remain within established social circles and that these regular interactions form a stable pattern over time, says Jaideep Srivastava, a computer scientist from the University of Minnesota in Minneapolis who was not involved in the study. "But the fact that you can use that pattern to identify the individual, that part is surprising."

According to the European Union's General Data Protection Regulation and the California Consumer Privacy Act, companies that collect information about people's daily interactions can share or sell this data without users' consent. The catch is that the data must be anonymized. Some organizations might assume that they can meet this standard by giving users pseudonyms, says Yves-Alexandre de Montjoye, a computational privacy researcher at Imperial College London. "Our results are showing that this is not true."

de Montjoye and his colleagues hypothesized that people's social behavior could be used to pick them out of datasets containing information on anonymous users' interactions. To test their hypothesis, the researchers taught an artificial neural network — an AI that simulates the neural circuitry of a biological brain — to recognize patterns in users' weekly social interactions.

For one test, the researchers trained the neural network with data from an unidentified mobile phone service that detailed 43,606 subscribers' interactions over 14 weeks. This data included each interaction's date, time, duration, type (call or text), the pseudonyms of the involved parties and who initiated the communication.

Each user's interaction data were organized into web-shaped data structures consisting of nodes representing the user and their contacts. Strings threaded with interaction data connected the nodes. The AI was shown the interaction web of a known person and then set loose to search the anonymized data for the web that bore the closest resemblance.

The neural network linked just 14.7 percent of individuals to their anonymized selves when it was shown interaction webs containing information about a target's phone interactions that occurred one week after the latest records in the anonymous dataset. But it identified 52.4 percent of people when given not just information about the target's interactions but also those of their contacts. When the researchers provided the AI with the target's and contacts' interaction data collected 20 weeks after the anonymous dataset, the AI still correctly identified users 24.3 percent of the time, suggesting social behavior remains identifiable for long periods of time.

To see whether the AI could profile social behavior elsewhere, the researchers tested it on a dataset consisting of four weeks of close-proximity data from the mobile phones of 587 anonymous university students, collected by researchers in Copenhagen. This included interaction data consisting of students' pseudonyms, encounter times and the strength of the received signal, which was indicative of proximity to other students. These metrics are often collected by COVID-19 contact tracing applications. Given a target and their contacts' interaction data, the AI correctly identified students in the dataset 26.4 percent of the time.

The findings, the researchers note, probably don't apply to the contact tracing protocols of Google and Apple's Exposure Notification system, which protects users' privacy by encrypting all Bluetooth metadata and banning the collection of location data.

de Montjoye says he hopes the research will help policy makers improve strategies to protect users' identities. Data protection laws allow the sharing of anonymized data to support useful research, he says. "However, what's essential for this to work is to make sure anonymization actually protects the privacy of individuals."

1) Qual das alternativas abaixo contém a hipótese dos pesquisadores para a pesquisa descrita no texto?

a) A anonimidade dos usuários é algo contestável.

b) O comportamento social de pessoas pode ser utilizado para identificá-las em conjuntos de dados que contenham informações anônimas.

c) Não é possível identificar usuários através de seu comportamento social em conjuntos de dados.

d) Uma rede neural artificial reconhece padrões em interações sociais semanais de usuários.

2) Em um dos testes descritos no texto, os pesquisadores treinaram uma inteligência artificial através de informações geralmente coletadas por aplicativos que rastreiam o contato de usuários. Quais informações foram utilizadas nesse teste?

a) Data, horário e duração das interações.

b) Pseudônimos e tipo de interação (chamada ou mensagem de texto).

c) Pseudônimos, horário dos encontros e a força do sinal recebido.

d) Nome, endereço e número de telefone.

3) Assinale a alternativa que melhor explica, com base nos resultados da pesquisa, o sentido expresso pela sentença introdutória do texto.

a) O leitor pode permanecer anônimo se interagir em um grupo.

b) Quanto maior a interação do leitor em um grupo, mais difícil será sua identificação pela inteligência artificial.

c) Inteligência artificial não é capaz de identificar o leitor se estiver inserido em um grupo.

d) Inteligência artificial é capaz de identificar o leitor através de suas interações em grupo.

4) Os seguintes números foram retirados, em ordem de aparição, do texto: 25 – 14 – 587 – 26.4. Assinale a alternativa que apresenta as informações a que se referem cada número.

a) Data de publicação da pesquisa – intervalo de tempo dos dados – número de participantes – índice de acerto.

b) Data de publicação da pesquisa – número de participantes – intervalo de tempo dos dados - índice de acerto.

c) Data de publicação da pesquisa – índice de acerto – número de participantes – intervalo de tempo dos dados.

d) Data de publicação da pesquisa – número de participantes – número de participantes – índice de acerto.

5) Um dos pesquisadores entrevistados comenta que os resultados da pesquisa descrita no texto são surpreendentes, pois:

a) Padrões de comportamento social podem ser utilizados para identificar o indivíduo.

b) Humanos tendem a permanecer em grupos sociais estabelecidos.

c) Interações regulares formam padrões estáveis com o tempo.

d) A tecnologia é surpreendente.

6) As alternativas abaixo são referentes ao primeiro teste descrito no texto. Assinale a única alternativa que contém uma informação falsa ou não apresentada no texto.

a) Redes neurais foram treinadas com dados de telefones portáteis.

b) Foram utilizados dados de 43,606 participantes referentes a 14 semanas.

c) Os dados incluíram diferentes informações sobre as interações, como data e duração.

d) Pseudônimos foram dados para os participantes das interações.

7) De acordo com o texto, os resultados do primeiro teste revelaram que:

a) A inteligência artificial identificou corretamente a maioria dos usuários quando treinada com as interações do indivíduo.

b) A inteligência artificial identificou corretamente os usuários entre 14.7 e 24.3 por cento dos casos através da interação do indivíduo.

c) A inteligência artificial ainda não é capaz de identificar corretamente o usuário com base em suas interações sociais.

d) A inteligência artificial identificou corretamente mais usuários quando foi treinada com a interação do próprio indivíduo e de seus contatos.

8) O que de Montjoye espera em relação aos achados da pesquisa?

Que a pesquisa incentive autoridades a melhorarem as estratégias de proteção de identidades de usuários.

## **A new origin story for domesticated chickens starts in rice fields 3,500 years ago** *Two studies lay out how the birds went from wild fowl in Southeast Asia to the dinner plate*

It turns out that chicken and rice may have always gone together, from the birds' initial domestication to tonight's dinner.

In two new studies, scientists lay out a potential story of chicken's origins. This poultry tale begins surprisingly recently in rice fields planted by Southeast Asian farmers around 3,500 years ago, zooarchaeologist Joris Peters and colleagues report. From there, the birds were transported westward not as food but as exotic or culturally revered creatures, the team suggests June 6 in the Proceedings of the National Academy of Sciences.

"Cereal cultivation may have acted as a catalyst for chicken domestication," says Peters, of Ludwig Maximilian University of Munich.

The domesticated fowl then arrived in Mediterranean Europe no earlier than around 2,800 years ago, archaeologist Julia Best of Cardiff University in Wales and colleagues report June 6 in *Antiquity*. The birds appeared in northwest Africa between 1,100 and 800 years ago, the team says.

Researchers have debated where and when chickens (*Gallus gallus domesticus*) originated for more than 50 years. India's Indus Valley, northern China and Southeast Asia have all been touted as domestication centers. Proposed dates for chickens' first appearance have mostly ranged from around 4,000 to 10,500 years ago. A 2020 genetic study of modern chickens suggested that domestication occurred among Southeast Asian red jungle fowl. But DNA analyses, increasingly used to study animal domestication, couldn't specify when domesticated chickens first appeared.

Using chicken remains previously excavated at more than 600 sites in 89 countries, Peters' group determined whether the chicken bones had been found where they were originally buried by soil or, instead, had moved downward into older sediment over time and thus were younger than previously assumed.

After establishing the timing of chickens' appearances at various sites, the researchers used historical references to chickens and data on subsistence strategies in each society to develop a scenario of the animals' domestication and spread.

The new story begins in Southeast Asian rice fields. The earliest known chicken remains come from Ban Non Wat, a dry rice-farming site in central Thailand that roughly dates to between 1650 B.C. and 1250 B.C. Dry rice farmers plant the crop on upland soil soaked by seasonal rains rather than in flooded fields or paddies. That would have made rice grains at Ban Non Wat fair game for avian ancestors of chickens.

These fields attracted hungry wild birds called red jungle fowl. Red jungle fowl increasingly fed on rice grains, and probably grains of another cereal crop called millet, grown by regional farmers, Peters' group speculates. A cultivated familiarity with people launched chicken domestication by around 3,500 years ago, the researchers say.

Chickens did not arrive in central China, South Asia or Mesopotamian society in what's now Iran and Iraq until nearly 3,000 years ago, the team estimates.

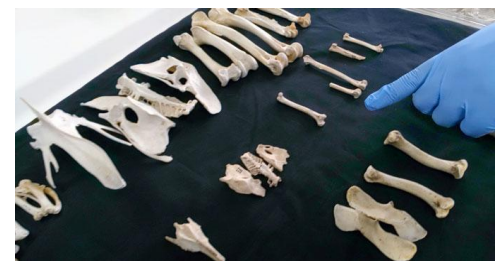
Peters and colleagues have for the first time assembled available evidence "into a fully coherent and plausible explanation of not only where and when, but also how and why chicken domestication happened," says archaeologist Keith Dobney of the University of Sydney who did not participate in the new research.

But the new insights into chickens don't end there. Using radiocarbon dating, Best's group determined that 23 chicken bones from 16 sites in Eurasia and Africa were generally younger, in some cases by several thousand years, than previously thought. These bones had apparently settled into lower sediment layers over time, where they were found with items made by earlier human cultures.

Archaeological evidence indicates that chickens and rice cultivation spread across Asia and Africa in tandem, Peters' group says. But rather than eating early chickens, people may have viewed them as special or sacred creatures. At Ban Non Wat and other early Southeast Asian sites, partial or whole skeletons of adult chickens were placed in human graves. That behavior suggests chickens enjoyed some sort of social or cultural significance, Peters says.

In Europe, several of the earliest chickens were buried alone or in human graves and show no signs of having been butchered.

The expansion of the Roman Empire around 2,000 years ago prompted more widespread consumption of chicken and eggs, Best and colleagues say. In England, chickens were not eaten regularly until around 1,700 years ago, primarily at Roman-influenced urban and military sites. Overall, about 700 to 800 years elapsed between the introduction of chickens in England and their acceptance as food, the researchers conclude. Similar lag times may have occurred at other sites where the birds were introduced.



**Figure 1** - A researcher points to chicken bones from England that are more than 2,000 years old (middle), which lie between bones of larger modern chickens

9) Os parágrafos iniciais do texto apresentam informações resultantes das pesquisas sobre domesticação da galinha. Assinale a única alternativa que contém uma informação falsa ou não apresentada no texto com relação a possível história da origem da galinha.

a) A história da domesticação das aves começou há cerca de 3500 anos.

**b) As aves foram transportadas para o oeste como uma opção de comida exótica.**

c) Há uma relação entre o cultivo de cereais e a domesticação da galinha.

d) A galinha foi primeiro levada para a Europa e depois para a África.

10) Os pesquisadores comentam que a galinha parece ter tido um importante papel social ou cultural em sociedades antigas. Qual achado da pesquisa levou os pesquisadores a essa conclusão?

**a) Ossos de galinha foram encontrados em túmulos humanos.**

b) Ossos de galinha foram encontrados em locais religiosos.

c) A galinha era considerada um animal especial ou sagrado.

d) A galinha era considerada um animal doméstico.

11) A Figura 1 apresenta ossos de galinhas de diferentes períodos históricos. Com base na legenda da figura, assinale a alternativa que contém uma informação que é possível inferir a respeito das galinhas antigas.

a) Elas eram maiores do que a galinha moderna.

**b) Elas eram menores do que a galinha moderna.**

c) Elas foram domesticadas primeiro na Inglaterra.

d) Elas foram domesticadas tardiamente na Inglaterra.

12) De acordo com o texto, os pesquisadores encontraram em Ban Non Wat:

a) Vestígios de uma espécie que precede a galinha doméstica atual.

b) Vestígios de um dos ancestrais da galinha.

c) Os restos mortais da galinha mais bem preservada já encontrada.

**d) Os restos mortais da galinha mais antiga já encontrada.**

13) Qual a limitação apresentada no texto para pesquisas com análises de DNA?

a) Descobriram que a galinha doméstica surgiu na Ásia.

b) Descobriram o ancestral mais recente da galinha doméstica.

**c) Não foram capazes de especificar quando a galinha doméstica surgiu.**

d) Não foram capazes de descobrir o ancestral mais recente da galinha doméstica.

14) Assinale a única alternativa que contém uma informação falsa sobre a metodologia da pesquisa desenvolvida por Peter e colegas.

a) Foram analisados ossos de galinhas provenientes de mais de 600 lugares em 89 países.

b) Os pesquisadores determinaram se os ossos foram encontrados no lugar que foram originalmente enterrados ou se foram movidos ao longo do tempo.

c) Os pesquisadores estabeleceram um intervalo de tempo mínimo de aparição da galinha para ser estudado nos lugares selecionados.

d) Os pesquisadores utilizaram referências históricas para delimitar a domesticação e a propagação da galinha.

15) Com base nas informações apresentadas no texto, o que aumentou a expansão do consumo de galinha e ovos?

a) A virada para o ano 2000 antes de Cristo.

b) A rápida reprodução da galinha doméstica em cativeiro.

c) A escassez de outras fontes de alimentos na idade média.

d) A expansão do império romano.

16) É comentado no texto que alguns ossos de galinhas encontrados na Eurásia e na África foram constatados como mais recentes do que previamente acreditado. Qual a explicação dos pesquisadores para a constatação errônea de que esses ossos aparentavam serem mais velhos do que realmente são?

Esses ossos se assentaram em camadas mais profundas de sedimentos, onde foram encontrados com itens produzidos por culturas humanas mais antigas.