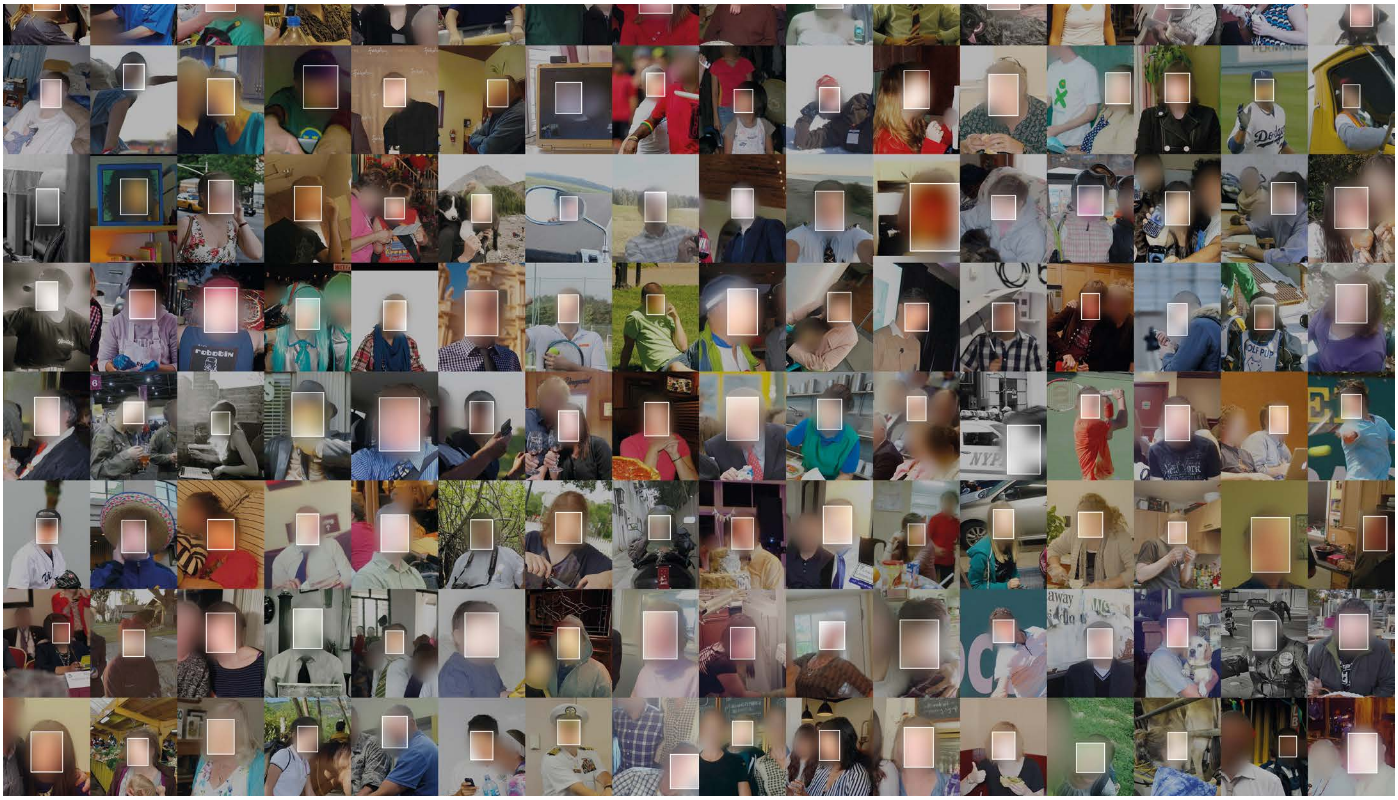


IMAGE VISUALIZATION BY ADAM HARVEY (HTTPS://MEGAPIXELS.CC) BASED ON THE MEGAFACE DATA SET BY IRA KEMELMACHER-SHLIZERMAN *ET AL.* BASED ON THE YAHOO FLICKR CREATIVE COMMONS 100 MILLION DATA SET AND LICENSED UNDER CREATIVE COMMONS ATTRIBUTION (CC BY) LICENCES

A collage of images from the MegaFace data set, which scraped online photos. Images are obscured to protect people's privacy.

THE ETHICAL QUESTIONS THAT HAUNT FACIAL-RECOGNITION RESEARCH

Journals and researchers are under fire for controversial studies using this technology. And a *Nature* survey reveals that many in this field think there is a problem. **By Richard Van Noorden**

In September 2019, four researchers wrote to the publisher Wiley to "respectfully ask" that it immediately retract a scientific paper. The study, published in 2018, had trained algorithms to distinguish faces of Uyghur people, a predominantly Muslim minority ethnic group in China, from those of Korean and Tibetan ethnicity[1].

China had already been internationally condemned for its heavy surveillance and mass detentions of Uyghurs in camps in the northwestern province of Xinjiang – which the government says are re-education centres aimed at quelling a terrorist movement. According to media reports, authorities in Xinjiang have used surveillance cameras equipped with software attuned to Uyghur faces.

As a result, many researchers found it disturbing that academics had tried to build such algorithms – and that a US journal had published a research paper on the topic. And the 2018 study wasn't the only one: journals from publishers including Springer Nature, Elsevier and the Institute of Electrical and Electronics Engineers (IEEE) had also published peer-reviewed papers that describe using facial recognition to identify Uyghurs and members of other Chinese minority groups. (*Nature*'s news team is editorially independent from its publisher, Springer Nature.)

The complaint, which launched an ongoing investigation, was one foray in a growing push by some scientists and human-rights activists to get the scientific community

© 2020 Springer Nature Limited. All rights reserved.

to take a firmer stance against unethical facial-recognition research. It's important to denounce controversial uses of the technology, but that's not enough, ethicists say. Scientists should also acknowledge the morally dubious foundations of much of the academic work in the field – including studies that have collected enormous data sets of images of people's faces without consent, many of which helped hone commercial or military surveillance algorithms. (A feature on page 347 explores concerns over algorithmic bias in facial-recognition systems.)

An increasing number of scientists are urging researchers to avoid working with firms or universities linked to unethical projects, to re-evaluate how they collect and distribute facial-recognition data sets and to rethink the ethics of their own studies. Some institutions are already taking steps in this direction. In the past year, several journals and an academic conference have announced extra ethics checks on studies.

"A lot of people are now questioning why the computer-vision community dedicates so much energy to facial-recognition work when it's so difficult to do it ethically," says Deborah Raji, a researcher in Ottawa who works at the non-profit Internet foundation Mozilla. "I'm seeing a growing coalition that is just against this entire enterprise."

This year, *Nature* asked 480 researchers around the world who work in facial recognition, computer vision and artificial intelligence (AI) for their views on thorny ethical questions about facial-recognition research. The results of this first-of-a-kind survey suggest that some scientists are concerned about the ethics of work in this field – but others still don't see academic studies as problematic.

Data without consent

For facial-recognition algorithms to work well, they must be trained and tested on large data sets of images, ideally captured many times under different lighting conditions and at different angles. In the 1990s and 2000s, scientists generally got volunteers to pose for these photos – but most now collect facial images without asking permission.

For instance, in 2015, scientists at Stanford University in California published a set of 12,000 images from a webcam in a San Francisco café that had been live-streamed online[2]. The following year, researchers at Duke University in Durham, North Carolina, released more than 2 million video frames (85 minutes) of footage of students walking on the university campus[3].

The biggest collections have been gathered online. In 2016, researchers at the University of Washington in Seattle posted a database, called MegaFace, of 3.3 million photos from the image-sharing site Flickr[4]. And scientists at Microsoft Research in Redmond, Washington, issued the world's largest data set, MSCeleb[5], consisting of 10 million images of nearly 100,000 individuals, including journalists, musicians and academics, scraped from the Internet.

In 2019, Berlin-based artist Adam Harvey created a website called MegaPixels that flagged these and other data sets. He and another Berlin-based technologist and programmer, Jules LaPlace, showed that many had been shared openly and used to evaluate and improve commercial surveillance products. Some were cited, for instance, by companies that worked on military projects in China. "I wanted to uncover the uncomfortable truth that many of the photos people posted online have an afterlife as training data," Harvey says. In total, he says he has charted 29 data sets, used in around 900 research projects. Researchers often use public Flickr images that were uploaded under copyright licences that allow liberal reuse.

After *The Financial Times* published an article on Harvey's work in 2019, Microsoft and several universities took their data sets down. Most said at the time – and reiterated to *Nature* this month – that their projects had been completed or that researchers had requested that the data set be removed.

> **"Conferences should avoid sponsors who are accused of enabling abuses of human rights."**

Computer scientist Carlo Tomasi at Duke University was the sole researcher to apologize for a mistake. In a statement two months after the data set had been taken down, he said he had got institutional review board (IRB) approval for his recordings – which his team made to analyse the motion of objects in video, not for facial recognition. But the IRB guidance said he shouldn't have recorded outdoors and shouldn't have made the data available without password protection. Tomasi told *Nature* that he did make efforts to alert students by putting up posters to describe the project.

The removal of the data sets seems to have dampened their usage a little, Harvey says. But big online image collections such as MSCeleb are still distributed among researchers, who continue to cite them, and in some cases have re-uploaded them or data sets derived from them. Scientists sometimes stipulate that data sets should be used only for non-commercial research – but once they have been widely shared, it is impossible to stop companies from obtaining and using them.

In October, computer scientists at Princeton University in New Jersey reported identifying 135 papers that had been published after the Duke data set had come down and which had used it or data derived from it (see go.nature.com/3nlkjre). The authors urged researchers to set more restrictions on the use of data sets and asked journals to stop accepting papers that use data sets that had been taken down.

Legally, it is unclear whether scientists in Europe can collect photos of individuals' faces for biometric research without their consent. The European Union's vaunted General Data Protection Regulation (GDPR) does not provide an obvious legal basis for researchers to do this, reported[6] Catherine Jasserand, a biometrics and privacy-law researcher at the Catholic University of Leuven in Belgium, in 2018. But there has been no official guidance on how to interpret the GDPR on this point, and it hasn't been tested in the courts. In the United States, some states say it is illegal for commercial firms to use a person's biometric data without their consent; Illinois is unique in allowing individuals to sue over this. As a result, several firms have been hit with class-action lawsuits.

The US social-media firm Facebook, for instance, agreed this year to pay US$650 million to resolve an Illinois class-action lawsuit over a collection of photos that was not publicly available, which it used for facial recognition (it now allows users to opt out of facial-recognition tagging). The controversial New York City-based technology company Clearview AI – which says it scraped three billion online photos for a facial-recognition system – has also been sued for violating this law in pending cases. And the US tech firms IBM, Google, Microsoft, Amazon and FaceFirst were also sued in Illinois for using a data set of nearly one million online photos that IBM released in January 2019; IBM removed it at around the time of the lawsuit, which followed a report by NBC News detailing photographers' disquiet that their pictures were in the data set.

Microsoft told *Nature* that it has filed to dismiss the case, and Clearview says it "searches only publicly available information, like Google or any other search engine". Other firms did not respond to requests for comment.

Vulnerable populations

In the study on Uyghur faces published by Wiley[1], the researchers didn't gather photos from online, but said they took pictures of more than 300 Uyghur, Korean and Tibetan 18–22-year-old students at Dalian Minzu University in northeast China, where some of the scientists worked. Months after the study was published, the authors added a note to say that the students had consented to this. But the researchers' assertions don't assuage ethical concerns, says Yves Moreau, a computational biologist at the Catholic University of Leuven. He sent Wiley a request to retract the work last year, together with the Toronto-based advocacy group Tech Inquiry.

It's unlikely that the students were told

© 2020 Springer Nature Limited. All rights reserved.

enough about the purpose of the research to have given truly informed consent, says Moreau. But even if they did freely consent, he argues, human-rights abuses in Xinjiang mean that Wiley ought to retract the study to avoid giving the work academic credence.

Moreau has catalogued dozens of papers on Uyghur populations, including facial-recognition work and studies that gathered Uyghur people's DNA. In December, he wrote an opinion article in *Nature* calling for all unethical work in biometric research to be retracted[7].

His campaign has had some impact, but not quite to the extent he'd hoped. Publishers say the key issue is checking whether participants in studies gave informed consent. Springer Nature, for instance, said in December 2019 that it would investigate papers of concern on vulnerable groups along these lines, and that it had updated its guidance to editors and authors about the need to gain explicit and informed consent in studies that involve clinical, biomedical or biometric data from people. This year, the publisher retracted two papers on DNA sequencing[8,9] because the authors conceded that they hadn't asked Uyghur people for their consent, and it has placed expressions of concern on 28 others.

Wiley has also focused on informed consent. Last November, the publisher told Moreau and Tech Inquiry that it was satisfied that consent forms and university approval for the Dalian study were available, and so it stood by the research, which it felt could be firmly separated from the actions in Xinjiang. "We are aware of the persecution of the Uyghur communities," Wiley said. "However, this article is about a specific technology and not an application of that technology."

In December, however, the publisher opened a formal investigation, after Curtin University in Perth, Australia, where one of the authors is based, also asked for a retraction, saying it agreed that the work was ethically indefensible. This year, Wiley added a publisher's note saying that the article "does appear to be in compliance with acceptable standards for conducting human subject research". In September, after Moreau dug into the authors' previous studies of facial recognition on Uyghurs and pointed out apparent inconsistencies in the year that the data sets had been gathered, Wiley placed an expression of concern on the study, saying that it was not clear when the data collection had taken place.

The publisher told *Nature* that it now considers the matter closed after thorough investigation – but not everyone involved is content. "Curtin University maintains that the paper should be retracted," deputy vice-chancellor Chris Moran told *Nature*. He said the university was still investigating the work.

Wiley says that after its conversations with Moreau, it updated its integrity guidelines to make sure that expected standards for informed consent are met and described in articles. Other publishers say that they have made adjustments, too. The IEEE says that this September, it approved a policy under which authors of articles on research involving human subjects or animals should confirm whether they had approval from an IRB or equivalent local review; editors determine whether research (on biometrics or other areas) involves human subjects.

But Moreau says that publishers' focus on the details of consent is too narrow, and that they should also take a stand on the wider ethics of research. "We are talking about massive human-rights abuses," he says. "At some point, Western publishers should say that there are some baselines above which they don't go." He suggests that publishers should set up independent ethics boards that can give opinions when questions such as these arise. (No publishers asked by *Nature* said that they had taken this step.) Universities and researchers who disapprove of human-rights abuses could also do more to express this by dropping their associations with questionable technology firms, says Kate Crawford, co-director of the AI Now Institute at New York University.

In the past year, there has been growing scrutiny of universities' partnerships with companies or research programmes linked to mass surveillance in Xinjiang. The Massachusetts Institute of Technology (MIT) in Cambridge, for example, said it would review its relationship with the Hong Kong-based tech firm SenseTime after the US government – in the middle of a trade war with China – blacklisted the firm and other Chinese AI companies, such as Megvii in Beijing, over their alleged contributions to human-rights violations in Xinjiang. In 2018, SenseTime and MIT announced they had formed an "alliance on artificial intelligence"; MIT says that SenseTime had provided an undisclosed sum to the university without any restrictions on how it would be used, and that the university will not give it back.

"There are a number of lawful and legitimate applications of face and biometric recognition."

Both Megvii and SenseTime contest the US blacklisting. SenseTime says its technology has "never been applied for any unethical purposes", and Megvii says it requires its clients "not to weaponize our technology or solutions or use them for illegal purposes".

Academic conferences have been contentious, too. The Chinese Conference on Biometrics Recognition (CCBR) was held in Xinjiang's capital, Ürümqi, in 2018. Anil Jain, a computer scientist at Michigan State University in East Lansing, sat on the conference's advisory board and travelled there to give a speech. Some AI researchers, including Toby Walsh at the University of New South Wales in Sydney, Australia, later criticized Jain for this in stories reported by the New York City-based *Coda* magazine.

Coda magazine also noted that Springer Nature sponsored the conference; the company said its role was limited to publishing CCBR proceedings and that it had strengthened its requirements for conference organizers to comply with the publisher's editorial policies after concerns were raised about past content. And Jain challenged the critique, telling *Nature* that attending conferences in China "does not mean that ... international conference participants, like me, condone these atrocities against minorities". Growth in surveillance there shouldn't be a reason to "curtail scientific exchange", he said.

Jain remains on the advisory board for CCBR 2020–21; Springer Nature is still publishing the conference abstracts. And major international computer-vision conferences have continued to accept sponsorship from Chinese firms. Just after the blacklisting, SenseTime and Megvii sponsored the 2019 International Conference on Computer Vision, and Megvii sponsored the 2020 Conference on Computer Vision and Pattern Recognition, although its logo was removed from the conference's website after the meeting occurred. "Conferences should avoid sponsors who are accused of enabling abuses of human rights," reiterates Walsh. However, he notes that last year, the non-governmental organization Human Rights Watch in New York City withdrew initial allegations that Megvii facial-recognition technology was involved in an app used in Xinjiang. Conference organizers did not respond to a request for comment.

Ethical checkpoints

Questionable research projects have popped up in the United States, too. On 5 May, Harrisburg University in Pennsylvania posted a press release declaring that researchers there had developed facial-recognition software "capable of predicting whether someone is likely going to be a criminal", with "80 percent accuracy and no racial bias". The announcement triggered a wave of criticism, as had previous studies that hark back to the discredited work of nineteenth-century physiognomists. One notorious 2016 study reported that a machine-learning algorithm could spot the difference between images of non-criminals and those of convicted criminals that were supplied by a Chinese police department[10].

Harrisburg University removed its press release on 6 May following the outcry, but left a dangling question: the press release had said that the work was to be published by Springer Nature in a book series (which the publisher later denied). On 22 June, more than 2,400 academics signed a letter from a group

© 2020 Springer Nature Limited. All rights reserved.

called the Coalition for Critical Technology (CCT), asking Springer Nature not to publish the work and calling on all publishers to refrain from publishing similar studies.

The letter pointed out that such studies are based on unsound science. It also noted that algorithmic tools that tell police where or who to target tend to provide a scientific veneer for automated methods that only exacerbate existing biases in the criminal justice system. Three days earlier, more than 1,400 US mathematicians had written a letter asking their colleagues to stop collaborating with police on algorithms that claim to help reduce crime, because of concerns about systemic racism in US law-enforcement agencies.

Springer Nature said the work was never accepted for publication: it had been submitted to a conference and rejected after peer review. (The authors, and Harrisburg University, declined to comment.)

Springer Nature was already under fire for a different paper, published in January in the *Journal of Big Data*, on detecting 'criminal tendency' in photos of criminals and non-criminals[11]. After researchers from the IEEE got in touch with ethical concerns, Margeret Hall, the paper's co-author at the University of Nebraska Omaha, asked in June for the paper to be withdrawn. Hall says the now-retracted paper was "indefensible". Springer Nature says the journal reviewed its processes and now requires authors to include statements on ethics approvals and consent when submitting manuscripts.

Nature survey

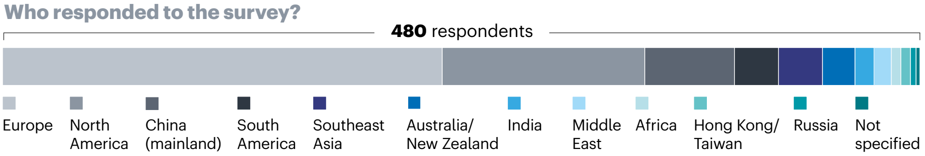
To get a wider sense of academic views on facial-recognition ethics, *Nature* this year surveyed 480 researchers who have published papers on facial recognition, AI and computer science. On some questions, respondents showed a clear preference. When asked for their opinions on studies that apply facial-recognition methods to recognize or predict personal characteristics (such as gender, sexual identity, age or ethnicity) from appearance, around two-thirds said that such studies should be done only with the informed consent of those whose faces were used, or after discussion with representatives of groups that might be affected (see 'Facial recognition: a survey on ethics').

But on other issues, academics were split. Around 40% of the scientists in the survey felt that researchers should get informed consent from individuals before using their faces in a facial-recognition data set, but more than half felt that this wasn't necessary. The researchers' dilemma is that it's hard to see how they can train accurate facial-recognition algorithms without vast data sets of photos, says Sébastien Marcel, who leads a biometrics group at the Idiap Research Institute in Martigny, Switzerland. He thinks that researchers

FACIAL RECOGNITION: A SURVEY ON ETHICS

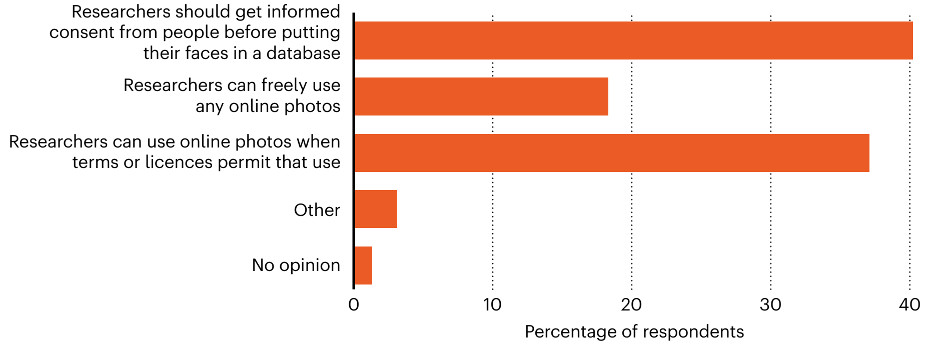
Nature surveyed* nearly 500 researchers who work in facial recognition, computer vision and artificial intelligence about ethical issues relating to facial-recognition research. They are split on whether certain types of this research are ethically problematic and what should be done about concerns.

Who responded to the survey?



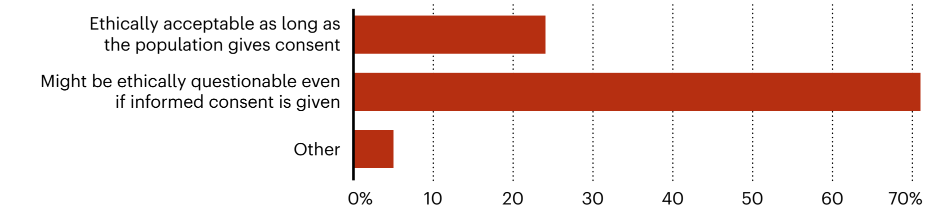
Restrictions on image use

Question: Researchers use large data sets of images of people's faces — often scraped from the Internet — to train and test facial-recognition algorithms. What kind of permissions do researchers need to use such images?



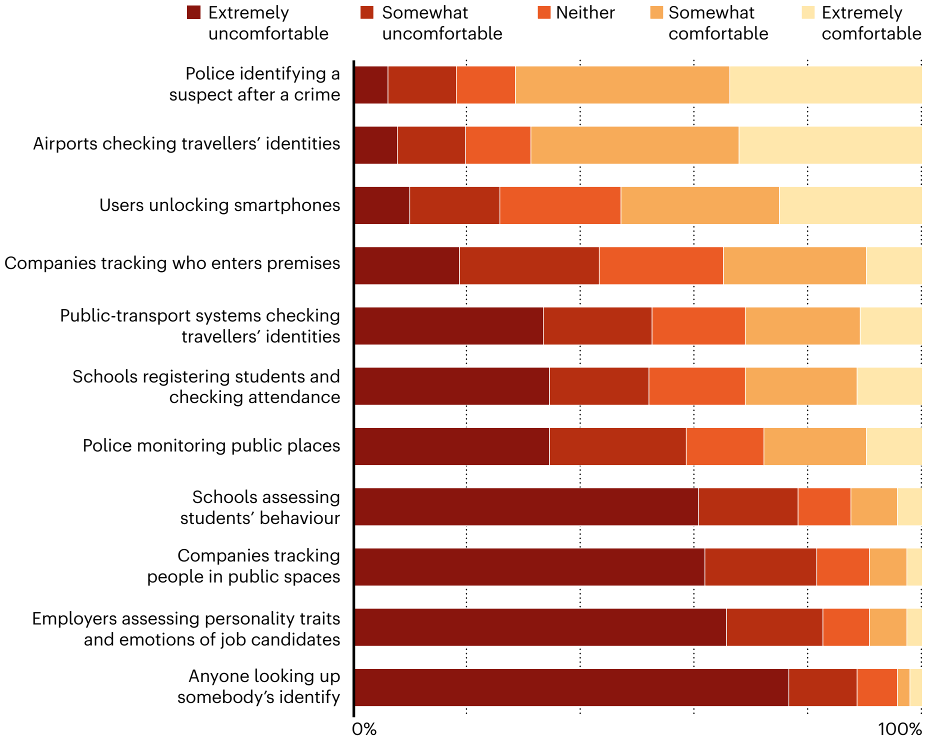
Restrictions related to vulnerable populations

Question: Is it ethical to do facial-recognition research on vulnerable populations that might not be able to freely give informed consent, such as the Muslim population in western China?



Attitudes on different uses

Question: How comfortable are you with facial-recognition technology being used in the following ways?



*Questions and answers have been paraphrased for brevity. The full survey and results are available online at go.nature.com/2uwtzyh

© 2020 Springer Nature Limited. All rights reserved.

Schoolchildren walk beneath surveillance cameras in Xinjiang in western China.

GREG BAKER/AFP/GETTY

should get informed consent – but in practice, they don't. His own group doesn't crawl the web for images, but it does use online image data sets that others have compiled. "A lot of researchers don't want to hear about this: they consider it not their problem," he says.

Ed Gerstner, director of journal policy at Springer Nature, said the publisher was considering what it could do to discourage the "continued use" of image databases that don't have explicit consent for their use in research from the people in the images.

Nature's survey also asked researchers whether they felt that facial-recognition research on vulnerable populations – such as refugees or minority groups that were under heavy surveillance – could be ethically questionable, even if scientists had gained informed consent. Overall, 71% agreed; some noted it might be impossible to determine whether consent from vulnerable populations was informed, making it potentially valueless.

Some of those who disagreed, however, tried to draw a distinction between academic research and how facial recognition is used. The focus should be on condemning and restricting unethical applications of facial recognition, not on restricting research, they said.

Ethicists regard that distinction as naive. "That's the 'I'm just an engineer' mentality – and we're well past that now," says Karen Levy, a sociologist at Cornell University in Ithaca, New York, who works on technology ethics.

Some of the respondents in China said that they were offended by the question. "You should not say that in Xinjiang some groups are detained in camps," wrote one. Just under half of the 47 Chinese respondents felt that studies on vulnerable groups could be ethically questionable even if scientists had gained consent, a lower proportion than respondents from the United States and Europe (both above 73%).

One Chinese American AI researcher who didn't want to be named said that a problem was a cultural split in the field. "The number of Chinese researchers at top conferences who actively support censorship and Xinjiang concentration camp[s] concerns me greatly. These groups have minimal contact with uncensored media and tend to avoid contact with those who don't speak Mandarin, especially about social issues like this. I believe we need to find ways to actively engage with this community," they wrote.

Nature asked researchers what the scientific community should do about ethically questionable studies. The most popular answer was that during peer review, authors of facial-recognition papers should be asked explicitly about the ethics of their studies. The survey also asked whether research that uses facial-recognition software should require prior approval from ethics bodies, such as IRBs, that consider research with human subjects. Almost half felt it should, and another quarter said it depended on the research.

Ethical reflection

Researchers who work on technology that recognizes or analyses faces point out that it has many uses, such as to find lost children, track criminals, access smartphones and cash machines more conveniently, help robots to interact with humans by recognizing their identities and emotions and, in some medical studies, to help diagnose or remotely track consenting participants. "There are a number of lawful and legitimate applications of face and biometric recognition which we need in our society," says Jain.

But researchers must also recognize that a technology that can remotely identify or classify people without their knowledge is fundamentally dangerous – and should try to resist it being used to control or criminalize people, say some scientists. "The AI community suffers from not seeing how its work fits into a long history of science being used to legitimize violence against marginalized people, and to stratify and separate people," says Chelsea Barabas, who studies algorithmic decision-making at MIT and helped to form the CCT this year. "If you design a facial-recognition algorithm for medical research without thinking about how it could be used by law enforcement, for instance, you're being negligent," she says.

Some organizations are starting to demand that researchers be more careful. One of the AI field's premier meetings, the NeurIPS (Neural Information Processing Systems) conference, is requiring such ethical considerations for the first time this year. Scientists submitting papers must add a statement addressing ethical concerns and potential negative outcomes of their work. "It won't solve the problem, but it's a step in the right direction," says David Ha, an AI researcher at Google in Tokyo. The journal *Nature Machine Intelligence* is also trialling an approach in which it asks the authors of some machine-learning papers to include a statement considering broader societal impacts and ethical concerns, Gerstner says.

Levy is hopeful that academics in facial recognition are waking up to the implications of what they work on – and what it might mean for their reputation if they don't root out ethical issues in the field. "It feels like a time of real awakening in the science community," she says. "People are more acutely aware of the ways in which technologies that they work on might be put to use politically – and they feel this viscerally."

Richard Van Noorden is a features editor for *Nature* in London.

1. Wang, C., Zhang, Q., Liu, W., Liu, Y. & Miao, L. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **9**, e1278 (2019).
2. Stewart, R., Andriluka, M. & Ng, A. Y. in *Proc. 2016 IEEE Conf. on Computer Vision and Pattern Recognition* 2325–2333 (IEEE, 2016).
3. Ristani, E., Solera, F., Zou, R. S., Cucchiara, R. & Tomasi, C. Preprint at https://arxiv.org/abs/1609.01775 (2016).
4. Nech, A. & Kemelmacher-Shlizerman, I. in *Proc. 2017 IEEE Conf. on Computer Vision and Pattern Recognition* 3406–3415 (IEEE, 2017).
5. Guo, Y., Zhang, L., Hu., Y., He., X. & Gao, J. in *Computer Vision — ECCV 2016* (eds Leibe, B., Matas, J., Sebe, N. & Welling, M.) https://doi.org/10.1007/978-3-319-46487-9_6 (Springer, 2016).
6. Jasserand, C. in *Data Protection and Privacy: The Internet of Bodies* (eds Leenes, R., van Brakel, R., Gutwirth, S. & de Hert, P.) Ch. 7 (Hart, 2018).
7. Moreau, Y. *Nature* **576**, 36–38 (2019).
8. Zhang, D. *et al. Int. J. Legal Med.* https://doi.org/10.1007/s00414-019-02049-6 (2019).
9. Pan, X. *et al. Int. J. Legal Med.* **134**, 2079 (2020).
10. Wu, X. & Xhang, X. Preprint at https://arxiv.org/abs/1611.04135 (2016).
11. Hashemi, M. & Hall, M. *J. Big Data* **7**, 2 (2020).

© 2020 Springer Nature Limited. All rights reserved.