Data Fusion Question Set

Understanding surveillance technology in your community

CARNEGIE COUNCIL for Ethics in International Affairs

Automated surveillance fusion technologies are systems that automatically bring together data from different sources for the purposes of investigations. These tools can make surveillance easier, faster, and more scalable. This has implications for safety, security, privacy, and human rights. If a law enforcement agency has acquired, is in the process of acquiring, or is considering acquiring this technology, these are some questions that you can ask them.

What is the technology for?

- What can you do with this technology that you could not do with existing tools or procedures?
- What was/is your process for deciding to acquire this technology and this product specifically?
- What is your concept of operations for using the technology?
- What specific types of crimes will the technology be used to pre-empt or investigate?
- What is the process for deciding what the technology is allowed to be used for?
- What is the likelihood that the technology will be needed?
- Will you present data or findings generated from this technology in court?

What is the technology *not* for?

- What are potential harms that could arise from the use of the technology and what is the plan to prevent them?
- Are there types of crimes or noncriminal behavior that the technology explicitly will not be used to investigate?
- How will these nonpermitted uses be blocked?
- What are specific ways in which the technology could be abused? And how will these abuses be prevented?
- Is there a process that enables citizens or other stakeholders to challenge a decision to use the technology for a particular case or application?

How does the technology work?

- How has the technology's effectiveness been measured?
- What are the credentials of the provider of the technology?
- Has this same product been used and deemed useful and appropriate in another jurisdiction?
- How do you plan to measure and report the technology's effectiveness in your real-world deployment of it?

Who is providing the technology?

- What are the provider's Terms of Service?
- What is the text of the contract you signed with the provider?
- Does the agreement with the provider include a nondisclosure agreement of any kind?
- Does the agreement permit the provider to review, edit, or control elements of the department's communications with the public?
- Does the provider have access to the system data and/or system outputs?
- Do the terms of service permit the provider to share the system data with other parties? If so, what are they?
- Is the system transparent, with the ability for in-house modifications, or is it a closed proprietary system that only the manufacturer can modify?

How much does the technology cost and who is paying?

- Are these taxpayer funds?
- If so, will there be any disclosures on the funding breakdown?
- If the technology is privately funded or funded from a different source outside our jurisdiction, what are the terms of this grant or donation?
- What are the yearly costs of maintaining the technology?

What are the technology's data sources?

- What is the process for deciding which data sources should and should not be integrated with the system?
- How are the data collected?
- How long are the data retained? What use cases or business rules will be used to delete data?
- How accurate are the data?
- How representative are the data?
- What datasets can be fused or correlated? What datasets cannot be fused? What datasets require special authorization to fuse?
- In what cases are data collected by the system seen by a human user? And in what cases are data automatically reviewed by the system? For automatically reviewed data, will different data retention rules apply?
- Are the data sources housed on local services, or are they hosted on the cloud?

What could go wrong with the technology?

- How were the risks of the technology identified?
- Is it possible for the technology to correlate data points that are not, in fact, related? What is the procedure for identifying and correcting these errors?
- Are the system's false positive and false negative error rates known? How and when were these error rates measured?
- How will you measure and report these error rates in your real-world deployment?
- Is the technology more likely to exhibit errors on data related to certain demographic groups?
- If the system is hacked, what type of information could be stolen or compromised?
- What measures have or will be implemented to prevent unauthorized access to the system?
- What measures have or will be implemented to prevent authorized users from using the system in an unauthorized manner?

Who will use this technology?

- How are the users of this technology trained? Will this training be once, or recurrent?
- Are there access controls for different types of users?
- How are the users of this technology monitored?
- Will the logs be systematically audited to detect abuse?
- Will a user's activity on the system be logged? Who will have access to these logs? How long will these logs be stored?

Does the system employ AI such as machine learning, deep learning, or other non-deterministic algorithms?

- Were these components audited? Who conducted the audit?
- How were the components trained? Was the training data representative of the data that the system will encounter here in real use? If there are differences, how will these differences be compensated for?
- What measures or tools are used to ensure that users understand these components and the manner in which they achieved any given output?

Use this space to write in your own questions: