PRIORITY
Criminal Justice
NEEDS INITIATIVE

A project of the **RAND Corporation,** the **Police Executive Research Forum,** **RTI International,** and the **University of Denver**

Digital Evidence and the U.S. Criminal Justice System

Identifying Technology and Other Needs to More Effectively Acquire and Utilize Digital Evidence

Appendix

Sean E. Goodison, Robert C. Davis, and Brian A. Jackson

Sponsored by the National Institute of Justice

For more information on this publication, visit www.rand.org/t/rr890

Published by the RAND Corporation, Santa Monica, Calif.

This project was supported by Award No. 2013-MU-CX-K003, awarded by the National Institute of Justice, Office of Justice Programs, U.S. Department of Justice. The opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect those of the Department of Justice.

This appendix provides additional methodological detail on the various steps of the analysis used in the body of the report. The text in this appendix is based on earlier RAND reports:

- Jackson, B. A., Russo, J., Hollywood, J. S., Woods, D., Silberglitt, R., Drake, G. B., Shaffer, J. S., Zaydman, M., & Chow, B. G. (2015). *Fostering innovation in community and institutional corrections: Identifying high-priority technology and other needs for the U.S. corrections sector,* Santa Monica, Calif., RAND Corporation, RR-820-NIJ. As of March 15, 2015: http://www.rand.org/pubs/research_reports/RR820.html
- Hollywood, J. S., Boon, J. E., Jr., Silberglitt, R., Chow, B. G., & Jackson, B. A. (2015). *High-priority information technology needs for law enforcement*, Santa Monica, Calif., RAND Corporation, RR-737-NIJ. As of March 15, 2015: http://www.rand.org/pubs/research_reports/RR737.html

Rating Digital Evidence Needs by Workshop Participants

As described in the main text, the identification and prioritization of needs was structured around five different objectives related to criminal justice and digital evidence. They were:

- acquiring digital evidence more effectively
- analyzing it more effectively
- searching and organizing it more effectively
- reducing the man-hours required to analyze it and reducing digital forensics backlogs
- facilitating chain of custody and authentication of digital evidence (i.e., bolstering its utility in court)

Each workshop attendee rated the needs in their working group on a 1–9 scale for benefits for each objective (where 9 was intended to correspond to an innovation that would improve performance by 20 or more and 1 was intended to correspond to an innovation having no effect on performance). In addition, participants were also asked to estimate probabilities of success for both technical reasons (i.e., was the proposed innovation easy or hard?) and organizational adoption (if it was produced, would it be broadly picked up by criminal justice agencies?).

The use of a nine-point scale for the benefits judgments in particular was designed to allow participants to make two "high-medium-low"–type judgments: i.e., judge if the benefit of the need was very high (falling in the 7–9 range), medium (4–6), or low (1–3); and then make a second judgment whether they thought it fell in the middle or in one of the extremes of the category (e.g., determining that a need in the 7–9 category was a 7).

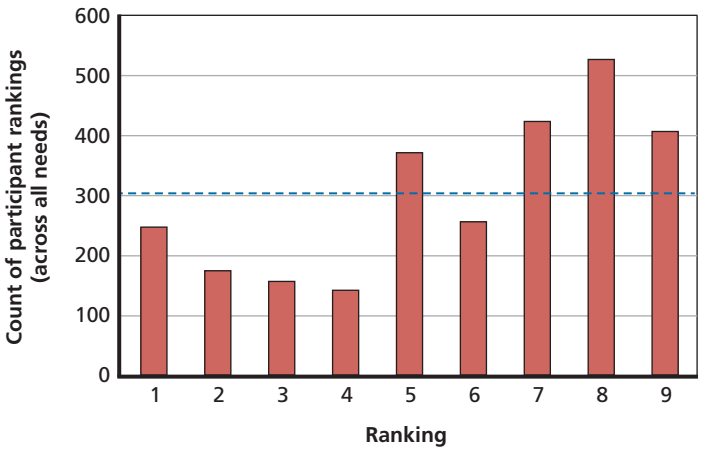
This rating approach meant that, for each digital evidence need that had been identified, participants made up to seven ratings (five if the need related to all of the objectives, plus two ratings of probability of success). Although not every participant believed that all needs contributed to each of the five objectives, every objective was rated for every need by at least a subset of participants.

To provide an aggregate picture of the distribution of the ratings, Figure A.1 shows the total counts of the value ratings on the five objectives across all the needs (i.e., excluding the estimates of probability of success and organizational adoption). If ratings were equally distributed across all of the nine options, each would have a count of 302 (shown by the blue dotted line). The distribution of technical success estimates (the probability that the need could be met successfully) was biased to the high end (Figure A.2), with almost 50 percent of the needs ranked 8 or 9 for technical success (and more than 30 percent ranked 9). The distribution of operational success assessments (reflecting the probability of broad adoption if the need was met) was much more center-weighted (Figure A.3).

Prioritizing Needs

To prioritize the needs, the benefit and probability of success scores were combined mathematically to estimate the likely operational payoff (expected value) of satisfying each need. Here, "expected value" is measured with respect to both the

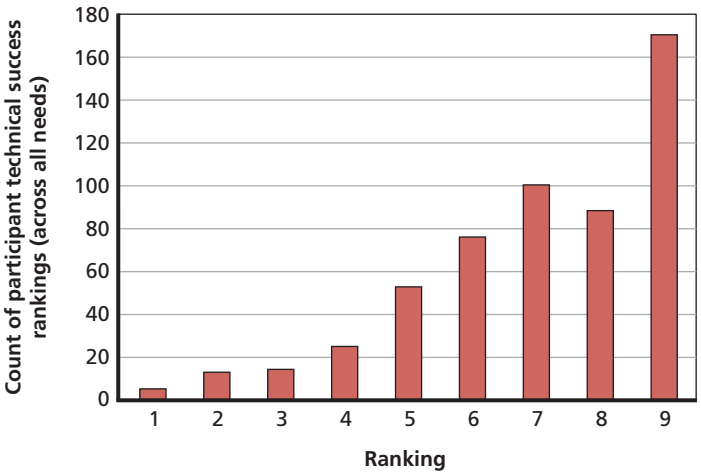
Figure A.1: Distribution of Value Rankings for All Digital Evidence Needs



NOTE: Blue dotted line shows average of total rankings across all nine rankings.

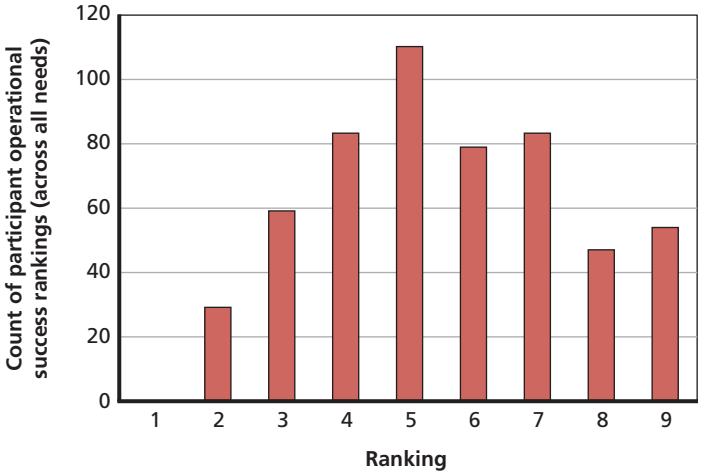
RAND *RR890-A.1*

Figure A.2: Distribution of Technical Success Rankings for All Digital Evidence Needs



RAND *RR890-A.2*

Figure A.3: Distribution of Operational Success Rankings for All Digital Evidence Needs



RAND *RR890-A.3*

operational benefit and probability of successfully fielding a technological breakthrough. Mathematically, the total expected value (EV) for a need is given by

$$EV_i = \sum_j (EV_{ij}) = \frac{\sum_j (w_j I_{ij} v_{ij} P_{1ij} P_{2ij})}{100},$$

where:

- w_j is a weight applying to objective j, which for this analysis was set at 0.2 for each of the five objectives (that is, the value of a need for achieving each objective contributed equally to its overall score). [1]
- I_{ij} is a 0–1 indicator for whether need i supports objective j, and the summation reflects the need's total value across all dimensions.
- v_{ij} is the estimated benefit (measured from 1–9) with respect to objective j if a project to satisfy need i is successful. Here, 9 = a "game changer;" i.e., improvement of 20 percent or more in a performance measure; 1 = no improvement. We defined the top of the scale based on analogies to previous criminal justice innovations that had major impacts (e.g., broad deployment of practical body armor, hot-spots policing in law enforcement) where measured effects were in this range.
- P_{1ij} is the estimated probability that a project will succeed *technically.* High scores occur if there are no major technical risks and the necessary knowledge or science is well understood.
- P_{2ij} is the estimated probability that a project will be *implemented by a large number of agencies.* High scores occur if there are no major operational, political, lifecycle-cost, or cultural barriers to implementation.

In other words, the equation says a need's score is the sum of its expected values toward contributing to individual objectives. Each expected value is the operational benefit if an effort to meet the need is successful, times the probabilities that such efforts will be technically and operationally successful. Put another way, the score for a need is determined by how beneficial it will be in achieving one or more objectives, and how likely the need can be met and deployed into the criminal justice community successfully. High-priority needs will tend to contribute to multiple objectives, make major potential contributions toward those objectives, and be comparatively low risk both technically and operationally. The product was divided by 100 to normalize for the product of the two probability-of-success ratings and convert them back into percentages.

Note that calculating expected values this way assumes linearity in the ranking scales (e.g., that from our top value of 9, associated with 20-percent improvement in performance for the objective, that raters divided the scale below 9 linearly down to a rating of 1, associated with no improvement. This had the effect of truncating the benefit scale at the top (i.e., any need with an expected benefit of greater than 20-percent improve-

[1] In previous RAND work, practitioners were also asked to prioritize the objectives, and those objectives were therefore weighted differently when expected values were calculated. Because of the way the objectives were broken down for this work, such a weighting was viewed as unhelpful and therefore prioritization was done with all objectives weighted equally.

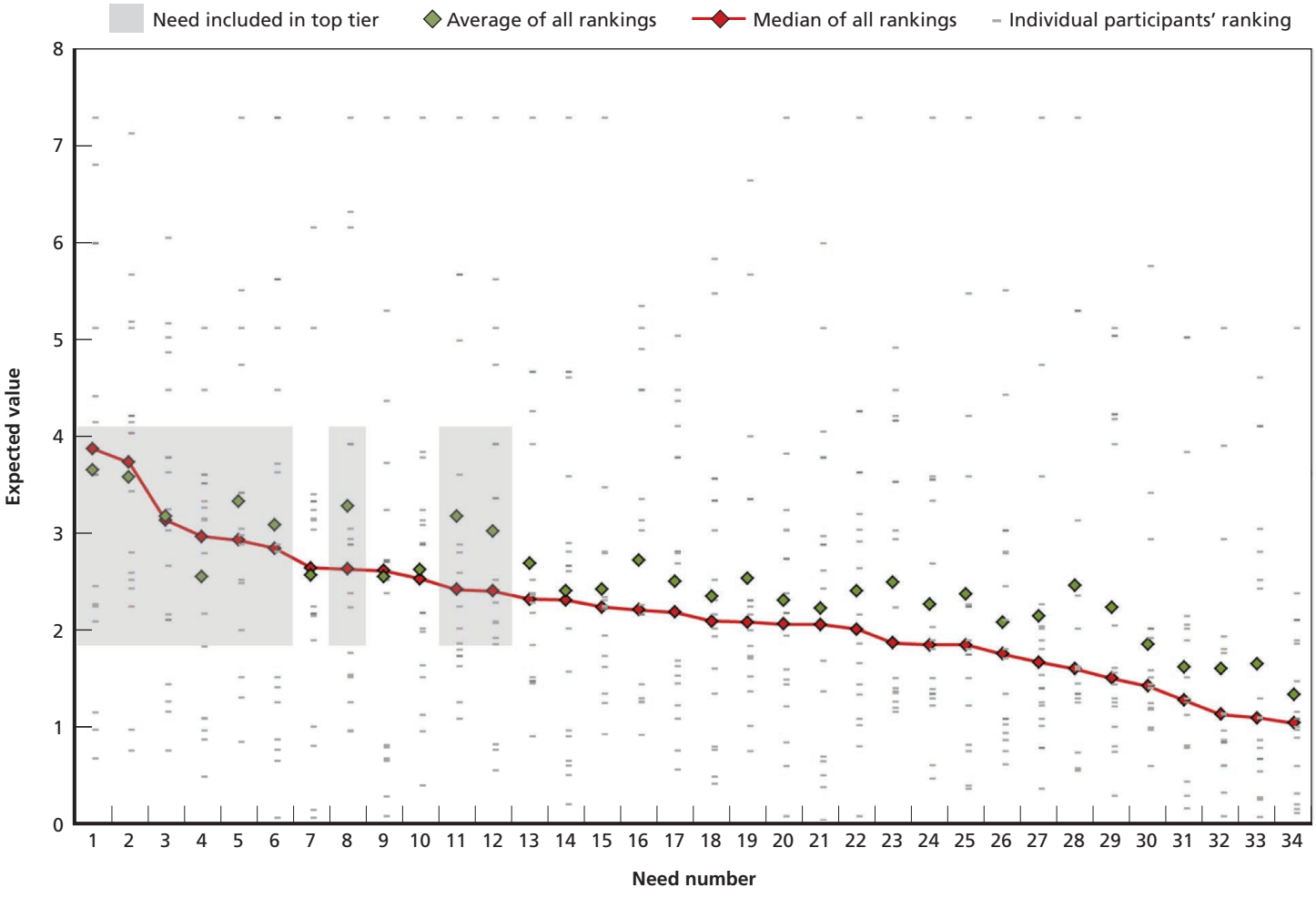
ment would still only be rated a 9). We believed this was an appropriate methodological choice since most innovations in criminal justice—when rigorously evaluated—have produced benefits below the range of 20 to 30 percent, and this made it possible for participants to assign different rankings to and distinguish between more-incremental innovations.

The Delphi method can be implemented a number of different ways, depending on the goals of the prioritization exercise. For problems where identifying an absolute consensus on a set of rankings or judgments, ranking rounds may be performed repeatedly—where the conclusion of the ranking is defined by some quantitative measure of consensus (e.g., the standard deviation around the mean of all participant rankings). The method can also be used more loosely, where the goal is to produce more approximate rankings, and so many fewer rounds of ratings are done. That was the case for this effort: Two rating rounds on the needs were done with one intervening discussion that focused mostly on cases where there was a great deal of spread in the responses across the groups. After discussion, participants were provided the opportunity (though were not obligated) to revisit their initial rankings and change them in response to the discussion.

Identifying Top-Tier Needs

We generated an overall expected value score for each need that combined the individual expected value ratings from the group members. To do so, we first took the *median* of the individual panelists' scores as a need's overall score. The *median* is the score that has the middle rank (50 percent of scores are higher, 50 percent are lower) in the data. Medians were chosen because they are robust—they provide reasonable estimates of the center of the data even given outliers or atypical distributions of the data. They do not require making any assumptions about the underlying statistical distribution of the scores. Median expected values were therefore used as the main ranking criterion for the needs. Figure A.4 plots all of the individual expected value rankings of each need by each of

Figure A.4: Expected Value Ratings for All Digital Evidence Needs, with Calculated Means and Medians



RAND *RR890-A.4*

the participants (gray dashes) and the median expected values (red diamonds). The needs are arranged from left to right from highest median rating to lowest, and the needs corresponding to each of the numbers on the graph are listed in Table A.1.

As is clear in Figure A.4, there is a great deal of dispersion in the rankings of individual needs across the workshop participants, even after they were given the opportunity to revise their ratings. Such dispersion is not unexpected, given the nature of the topic area and participants (who came from different parts of the criminal justice system, private industry, and other groups). As a result, differences in views on the importance of individual needs may come from the differences in perspective of raters based on their own experience.

From a quantitative or statistical perspective, the wide dispersion in rankings means that small differences in rankings between two needs are not particularly meaningful. As a result, we explicitly do not—and would caution against—reading too much meaning into small differences in needs' rank ordering. As a result, in presenting the results of these sorts of needs generation and ranking efforts, we have instead broken needs into tiers based on their overall ranking to provide a reasonable approach to prioritization while not overinterpreting the available data.

In this set of rankings, the most significant separation between needs' median rankings (red diamonds in Figure A.4) is after the two top-rated needs (both of which focus on different training issues related to digital evidence), where a considerable drop in expected value sets them somewhat apart from the rest. Looking at the medians of the rankings, the next obvious breakpoint falls between need 6 (also focusing on training) and need 7 (addressing access to data from GPS devices), suggesting a potential natural boundary in the data for identifying a small set of top priorities. This initial set of "top-tier needs" is identified in Figure A.4 with the gray box capturing the first six needs.

Our previous work (Jackson et al., 2015; Hollywood et al., 2015) has generally relied on median rankings for breaking needs into tiers, but given that workshop participants came from very different backgrounds in the criminal justice system, it could legitimately be argued that outlier values (which could be driven by the assessment of a need by a member of the working group with specific and unique experience) may be meaningful. As a result, to supplement the six needs identified by their highest median rankings, we also calculated simple means (averages) of the participants' ratings as well (graphed in Figure A.4 as green diamonds). In some cases, the median and mean were very close in value, but for most needs, the mean rating exceeded the median. From the remaining needs, we added three additional needs (numbers 8, 11, and 12, shown in smaller gray boxes in Figure A.4) to the top tier because they had medians that were still high, but also had much higher average ratings. This provided a top tier of nine out of the 34 total needs.

Because the remaining needs were separated by relatively small differences in expected value, we did not further differentiate beyond the identification of the nine top-tier needs.

Table A.1: Digital Evidence Needs Numbered as in Figure A.4

Need Number	Problem, Issue, or Technology Area	Associated Need(s)	Top-Tier Need? (M= Median, A=Average)
1	Prosecutors have a tendency to request all information off devices without considering the challenge posed by large volumes of data.	Expand available federal-level training at existing training schools to build knowledge across system.	Yes[M]
2	First-responding officers to an incident or arrest often do not know how to secure and use digital evidence to preserve chain of custody and later admissibility in court; e.g., "a detective searching a computer on his own."	Integrate digital evidence practices into academy training—at least at the awareness/ basic training level.	Yes[M]
3	Departments do not have enough personnel to process volume of digital evidence, no matter what tools are used, resulting in large backlogs.	Develop better prioritization or triaging methods or tools for cases and for what evidence to extract within cases (either for digital evidence examiners or potentially tools usable by officers in the field).	Yes[M]
4	Smaller departments lack capacity to address digital evidence.	Develop regional models for building capability where small departments pay to fund common resources. Incentives could be created through grant mechanisms to facilitate this approach.	Yes[M]
5	The acceptability of results of digital evidence analysis can be challenged in court when extraction and analysis is not performed with the most up-to-date tools.	Routinely update the training and tools available to examiners to ensure they are using the current technology.	Yes[M]
6	Lack of knowledge about digital evidence on the part of judges complicates appropriate use in court.	Expand available federal-level training at existing training schools to build knowledge across system.	Yes[M]
7	Some GPS devices available on the market use proprietary software and access technologies that make it difficult to extract data during investigations.	Utilize alternative approaches to acquire data from the company (e.g., execute search warrants on companies for data that these devices transmit to company servers) rather than focusing on the devices themselves.	
8	Departments lack tools to represent complex data sets in understandable ways for investigation and presentation.	Utilize existing software tools for analysis of data sets like cell tower data. Examples exist that are web-based and can be bought on a case-by-case basis, but knowledge of what is available is limited.	Yes[A]
9	Encryption and passwords on mobile phones prevent access.	Develop alternative access methods to address encryption.	
10	It can be difficult to access on-car digital evidence from systems such as OnStar (and other devices that cannot be removed from the platforms).	Develop tools to allow easier access to that data without disassembling and/or destroying devices, while also maintaining chain of custody.	
11	Volume of data coming from closed-circuit television (CCTV) cameras and video is a challenge—and there are limited tools for evaluating and processing evidence.	Departments must acquire in-house tools to process video evidence.	Yes[A]

Table 1—Continued

Need Number	Problem, Issue, or Technology Area	Associated Need(s)	Top Tier Need? (M= Median A=Average)
12	Collecting digital evidence from victim devices—where broad capture of all data on phone might capture data law enforcement "doesn't want" (e.g., sexting materials)—can be problematic.	Develop tools that allow more narrow collection of data from devices to respect victim privacy while still meeting investigative or protective needs.	Yes[A]
13	Managing multiple video evidence streams (e.g., business CCTV, personal cell phone video) during large incidents poses a data management and analysis challenge.	Develop information systems to better manage data, link with metadata, etc. to allow searchability and analysis.	
14	Having to pay for access to historical data sets of public data (e.g., Craigslist posts) poses a cost challenge for departments.	Build a public access data set for law enforcement for investigative purposes that captures and archives such data.	
15	Departments do not have enough personnel to process volume of digital evidence, no matter what tools are used, resulting in large backlogs.	Increase sworn-in manpower devoted to digital forensics activities.	
16	The performance and acceptability of new evidence collection and analysis techniques for criminal justice use are uncertain.	Provide timely validation/evaluation of technologies and analysis types of different products and techniques against established standards.	
17	Departments face real difficulty in maintaining capability to collect and analyze digital evidence given the pace of technological change.	Develop more standardized certifications for digital forensics personnel, including continuing education requirements.	
18	Need ways to collect "routine" digital evidence in a way that does not require full examiner involvement, and does not always require seizure of the device (e.g., from a crime victim).	Develop deployable tools for detectives to collect evidence in the field, but design in such a way that addresses potential for misuse and appropriately controls information and access.	
19	Departments do not have enough personnel to process volume of digital evidence, no matter what tools are used, resulting in large backlogs.	Define roles for lower-paid "digital evidence assistants" who can perform routine examinations.	
20	Current tools for explicit image detection are not effective at identifying explicit images.	Enhance explicit image detection to narrow how many images need to be reviewed by examiners.	
21	The practice of "promoting out" staff from digital evidence units pose a problem for agencies to maintain technical proficiency.	Create a promotion track within specialist units.	
22	Investigators may have no way to identify that data in suspect or victim cloud storage accounts exists and could provide investigative leads.	Develop tools to identify where accounts exist to trigger follow-up investigation.	
23	Some courts are skeptical of digital evidence due to uncertainties about chain of custody and validity of information obtained from devices.	Need an effort to systematically validate the performance of digital evidence tools to ensure they can withstand *Daubert* challenge.	

Table 1—Continued

Need Number	Problem, Issue, or Technology Area	Associated Need(s)	Top Tier Need? (M= Median A=Average)
24	Law enforcement lacks tools to analyze some types of electronic systems and devices.	Develop digital evidence tools to examine gaming devices.	
25	Law enforcement lacks tools to analyze some types of electronic systems and devices.	Develop digital evidence tools for examining networks.	
26	Departments do not have enough personnel to process volume of digital evidence, no matter what tools are used, resulting in large backlogs.	Address pay scale issues to make it possible to successfully recruit civilian staff for technical roles.	
27	Law enforcement lacks tools to analyze some types of electronic systems and devices.	Develop digital evidence tools for examining routers.	
28	Proprietary codexes for video evidence can create analysis problems.	Though commercially available video conversion tools allow conversion through screen capture, improvements that reduce the time required for such conversion would be valuable.	
29	Technologies developed to address problems have a "whack-a-mole" character trying to catch up with innovation.	Consider prize models to create incentives for many different private-sector actors to work on different digital evidence problems simultaneously.	
30	Cross-international-border issues create significant challenges for issuing and serving warrants for electronic information from entities in other countries.	Improve efficiency of MLAT processes for requesting information from foreign entities.	
31	Agency budget constraints make it difficult to maintain the currency of digital evidence tools and software packages.	Develop low-cost or free digital evidence analysis tools.	
32	Virtual currencies pose challenges for investigations.	Develop tools to identify presence of virtual currency on seized devices.	
33	Within agencies, a lack of leadership commitment to sufficiently funding digital evidence analysis capacity limits the ability to build and maintain expertise.	Develop information to make the case for building and maintaining digital evidence analysis capability outside of federal grant streams, preparing departments for making the transition to funding these capabilities internally.	
34	Quality of video evidence can limit use of other analytic tools (e.g., facial recognition).	Develop information to help persuade entities to adopt better video technologies to broaden technology options for analysis.	

NOTE: Items marked with an "M," for "median," are needs tiered based on their median expected values. Items marked with an "A," for "average," are needs added due to their high average expected values.