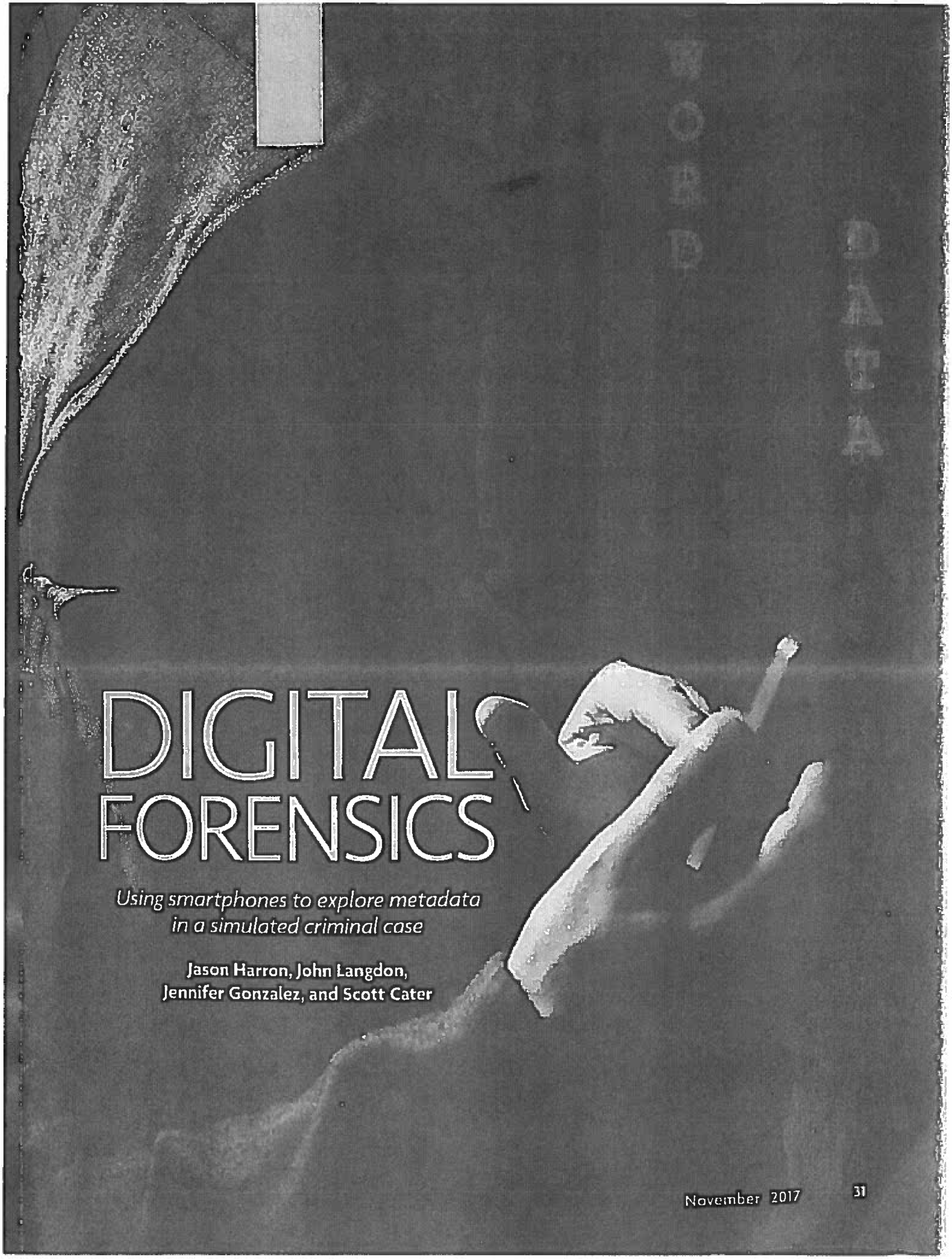


DIGITAL FORENSICS

Using smartphones to explore metadata in a simulated criminal case

Jason Harron, John Langdon, Jennifer Gonzalez, and Scott Cater

The term *forensic science* may evoke thoughts of blood-spatter analysis, DNA testing, and identifying molds, spores, and larvae. A growing part of this field, however, is that of digital forensics (Bertino 2012), involving techniques with clear connections to math and physics (Figure 1).

This article describes a five-part project involving smartphones and the investigation of a hypothetical crime and subsequent mock trial. It was conducted in a forensic science course. Smartphones have become ubiquitous in high schools (Purcell et al. 2013). For our lesson, 31 of 32 students had access to their own smartphones, and, for any who didn't, we made classroom tablet computers available to all. Smartphones can be powerful tools to engage students in a variety of scientific explorations (Kamarainen et al. 2013; Cartwright 2016).

The project

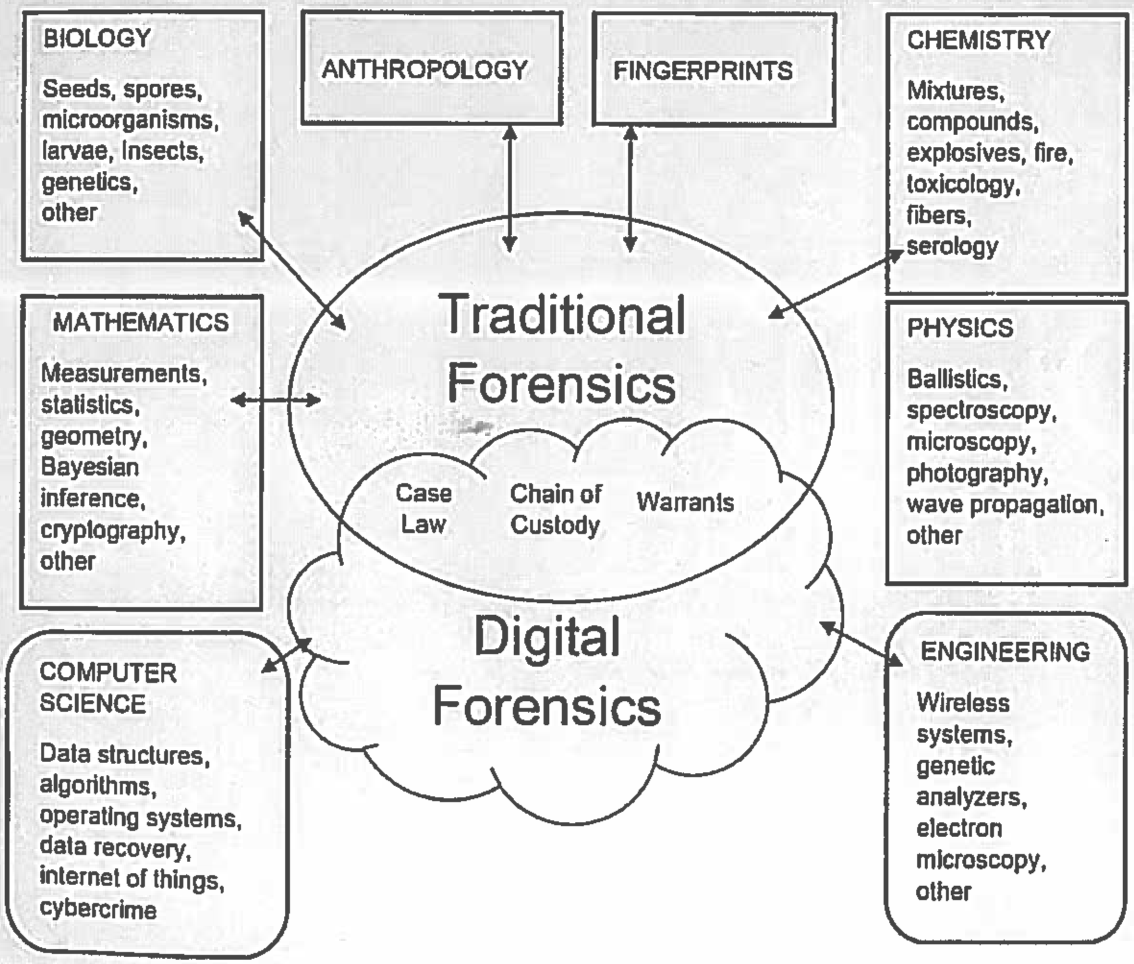
The "big ideas" embedded in the project were password security, cellular metadata, digital forensic examination of a smartphone, interpretation and presentation of evidence, and maintaining evidence through a chain of custody. Resources used, besides the smartphones, included public websites, poster-making materials, a web-based smartphone emulator, an inexpensive "burner" (prepaid) phone, evidence bags, and existing student WordPress blogs.

Part 1: Password strength (60–90 minutes)

The lesson started with a simple exercise about the relative strength of computer passwords. Students entered passwords into a free website (see "On the web") that instantly estimated how long a computer would take to crack their passwords,

FIGURE 1

Forensic science knowledge map.


BIOLOGY
Seeds, spores, microorganisms, larvae, Insects, genetics, other
ANTHROPOLOGY
FINGERPRINTS
CHEMISTRY
Mixtures, compounds, explosives, fire, toxicology, fibers, serology
MATHEMATICS
Measurements, statistics, geometry, Bayesian inference, cryptography, other
Traditional Forensics
PHYSICS
Ballistics, spectroscopy, microscopy, photography, wave propagation, other
Case Law
Chain of Custody
Warrants
Digital Forensics
COMPUTER SCIENCE
Data structures, algorithms, operating systems, data recovery, internet of things, cybercrime
ENGINEERING
Wireless systems, genetic analyzers, electron microscopy, other

ranging from milliseconds to years, either via dictionary or brute force attack. (In a dictionary attack, words from the dictionary are tried as possible passwords. If successful, these attacks only take seconds. In a brute force attack, every combination of characters [e.g., *a, A, b, B,* and so on] and symbols [e.g., *!,* @, #] are tried one at a time until the password is found. This can take days, months, or years, depending on the length of the password.

Instructors could formatively assess how well students met the challenge of seeing who could create the best password, monitoring their progress from an "over the shoulder" perspective. After about 10 minutes, students shared and compared passwords to help compile as a group samples of strong and weak passwords (Figure 2). (*Safety note:* Students should not share actual passwords they use for personal accounts.) After about five minutes of group work, instructors asked the class who came up with the best password.

Students learned that long strings of characters took much longer to crack. For example, the website stated that a computer could figure out the password "glucose," a dictionary word, in two hundred milliseconds. In contrast, "Gr8expect@tions4U!," the winning password, would take seven quadrillion years because of its length, use of uppercase and lowercase letters, alternative spellings, and special characters.

Random strings of characters may make strong passwords but have drawbacks of their own, such as being hard to remember. Students were given 10 minutes to create a password that was (1) short, (2) secure, and (3) easy to remember. They tested these new passwords on the same website. That lead to innovative use of special characters, such as emojis, and passphrases consisting of strings of words or characters mixing upper- and lower-case letters, symbols, and numbers, such as "ForensicScience4Ever!"

Next, 16 groups of two students spent 20 minutes creating posters listing examples of weak and strong passwords. Weak passwords tended to be predictable number patterns (e.g., "1234"), names (e.g., your name, pet's name), or personal information (e.g., birthdate, city where you live), while strong passwords used special characters such as %, @, and ?.

The posters were displayed on walls where classmates affixed sticky-note comments on their peers' work. Students discussed the feedback, then instructors led a 15-minute discussion guided by questions such as:

- What are the ethics of cracking someone's password?
- How has the internet and computing affected privacy and security?
- What are alternatives to text-based passwords?

This last question prompted research and discussion on the science and ethics of fingerprint scanners, facial recognition algorithms, and retina scans as biometric alternatives to passwords. Throughout the discussion students explained their reasoning and supported claims with evidence. The last five minutes of class were dedicated to students writing on their own blogs about the experience (see "On the web").

Part 2: Cracking a real phone (30 minutes)

On day two, the instructors led a 10-minute exploration of passwords or lock patterns used to unlock smartphones, which are typically based on numbers and graphic patterns rather than characters. Using examples from the internet (see "On the web" for Android smartphone lock patterns),

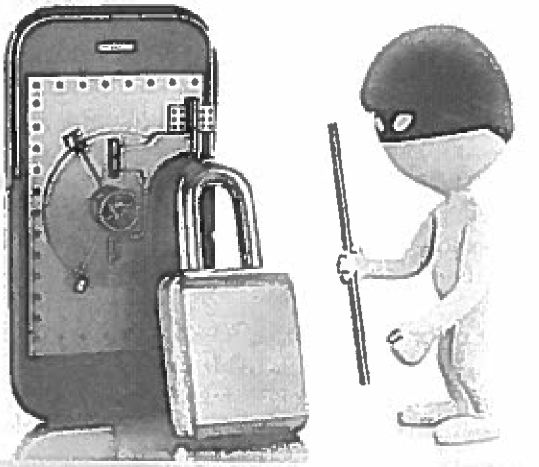


FIGURE 2

Student-derived examples of weak and strong passwords.

Weak passwords or [types of passwords]	Strong passwords (with time needed to crack)
1234	Cats&Hom3 (4 weeks)
5678	#721!?-89@$ (11 months)
[Person's name]	%65@23HN+y (6 years)
[Pet's name]	?8@42Ln (31 minutes)
[Birthday]	900%:Ms38&? (5,000 years)
[Repeating passwords]	PRNBET!72@ (1 month)
[Place name]	[Alternating lower + uppercase]
[Personal information, such as address]	[Symbols + numbers]
[Name of a celebrity]	[A special date only you know]
	[Foreign language words]

students looked at how patterns of lines and shapes could be defeated easily, particularly when formed in the shape of letters, and searched for other online resources to improve this form of security to share with the class.

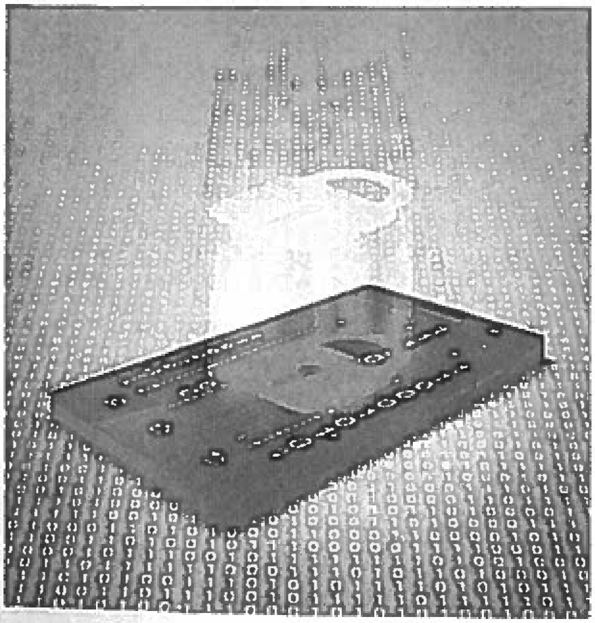
Beforehand, instructors had preloaded an inexpensive prepaid "burner" phone with actual evidence, such as text messages, placed calls, and photos of the campus and city, for the students to investigate. For our purposes, this phone had been found (hypothetically) at a crime scene and needed to be cracked to aid in a criminal investigation. The real phone was supplemented with an online smartphone emulator (see "On the web") that allowed all the student groups to try to crack the smartphone password simultaneously.

For scaffolding, students were provided a list of common graphical patterns used on smartphone lock screens and given about 10 minutes to attempt cracking the password. (*Safety note:* To protect student privacy, don't allow students to try to crack the password on another student's phone.)

A 10-minute class exploration followed, focusing on how using patterns instead of numbers connects to statistics and geometry and how the human visual system processes and recognizes shapes. After the lesson, several students told the instructors they had changed the passwords on their own phones to be more secure.

Part 3: Exploring the data (30–45 minutes)

Search engines such as Google collect data on their users to help them display the most relevant search results and to target advertisements to particular users. The collected data is deep enough to create a profile of the user that includes information about a person's home, school, daycare, work, grocery store, and even favorite restaurant. Search history can also reveal children, pets, major purchases, hobbies, medical conditions, and more.



In this activity, students watched a video (see "On the web") that discusses the breadth of information collected about users. Then, student pairs spent 10 minutes searching the exact same phrase using their smartphones or tablet computers. They discovered that each pair got different results, especially in the "paid for" links appearing at the top of the search results. The class then discussed these questions: Why do they get different results? What kind of data can they infer is being used to target them? Two students who searched for "automobiles" discovered that they received different results apparently based on their previous searches for cars, the zip code they lived in, and whether they had a driver's license.

Student groups then searched the term *cellphone metadata* and spent about 15 minutes compiling lists of what kinds of data are kept by phone carriers and social media websites (see cellphone metadata "On the web"). As an extension, students also searched and discussed some local legal cases in the news for which cellphone metadata were a key part of a trial, thus connecting the lesson back to digital forensics.

Part 4: Constructing the profile of our suspect (45–60 minutes)

Working as digital forensics technicians, students next examined and reported about the evidence they found on our smartphone hypothetically discovered at a crime scene. Students were guided by such questions as: What kind of data will they look for and in what order? How would they identify the phone's owner? Where has the phone been for the last several days? What recent calls and text messages are visible? Does the phone have photos, and if so, do they have date, time, and location data embedded in them? Is any GPS data available?

The class was divided into five groups to search for and analyze: (1) text messages, (2) phone calls and contacts, (3) photos, (4) browser search history, and (5) GPS and map data. Each group was assigned one type of data and given about eight minutes to access the data on the "burner" phone. While one group examined the phone, others were researching the laws related to digital evidence or analyzing the data that they had already found. When the phone was handed off from group to group, the instructors insisted that they use evidence bags and document possession to maintain the chain of custody. The password on the phone was the same as the virtual activity in part 2, and students were instructed not to tamper with the data.

After unlocking the phone, the groups discovered text messages referring to "the bomb" and a meeting scheduled at a local restaurant. Logged calls showed the same unidentified contact as the text messages. Photos stored on the phone were of bus stops and the exterior of a football stadium. One student noticed the photographer's shoes were visible in a photo and resembled those of one of the instructors. The search history included websites for the weather, news, and chemical compounds. Unfortunately, the GPS group couldn't locate any useful data.

FIGURE 3

Lesson rubric.

Rubric components	Point scale			
	4	3	2	1
Group password poster	Bad, better, best with example times to guess and/or trade-off between security and memorability	Weak, strong, with examples and times to guess	Weak, strong examples	Poster incomplete
What is metadata?	+any other usage: GPS track, etc.	+other numbers dialed	+date and time of text/ call	"data about data"
How can metadata be used?	+track locations over time	+timeline of texts/calls	+list of incoming calls	make a list of your friends
Hypothetical mobile phone processed: blog post (Elements examined) • Chemical forensics • Passwords Contacts • GPS data • Metadata • Photos (EXIF) • Text messages Social media Disguised apps • Other	4 or more	3	2	1
Digital ethics discussion (Make a group list of digital ethics issues, then share with class) • Expected issues: • Search warrant required? • User permission? • Chain of custody • Social media posts • Cyberbullying • Searching social media • Phishing	2 or more contribution to class	1 good contribution to class	Minor contribution	No contribution
Questions for next time	More than two good questions	Two good questions	One good question	none

Part 5: Supporting findings with claim-evidence-reasoning model (45–60 minutes)

For the final part, students reformed into two larger groups of prosecutors and defense attorneys to argue a case before the judge (teacher) against the owner of the "burner" phone. Each group was given 15 minutes to prepare its case based on its findings from part 4.

The prosecutors worked together to craft an opening statement that summarized the evidence and to assert that the phone owner was guilty of a crime. In their statement, the defense attorneys promised to show that the evidence was inadmissible. After the opening statements, each side presented its evidence using the claim-evidence-reasoning model, using slides that detailed their claims, evidence for the claims, and reasoning to justify the claims. The defense team challenged the prosecutors' evidence, showing, for example, the chain-of-custody logs were incomplete. Much to the dismay of the prosecutors, the judge ruled that all the gathered evidence was inadmissible.

The lesson concluded with a 15-minute class discussion of digital ethics, including such ethical questions as: How do you balance solving crimes and protecting the public's personal privacy? How do you use data ethically? Other questions addressed the science content: What are the connections between traditional forensics and digital forensics? How has computing affected society? How does this project relate to scientific reasoning and argumentation? Such questions prompted students to make broader connections between forms of data and evidence-based reasoning. Finally, students wrote individual blog posts about the digital evidence from the class investigation and how it related to forensic investigations.

Assessment

Formative assessment included "over the shoulder" observations, multiple instances of small-group and whole-class discussions, poster, presenting evidence at a mock trial, and individual student blogs for self-reflection. Summative assessment followed a rubric (Figure 3, p. 35).

Application to other subjects

In a computer science course, this lesson could help students explore the effects of computing, privacy, and digital citizenship. Physics and math classes could use actual GPS data from the "burner" phone to address GPS technology along with that of digital transmission and storage. Biometric data such as fingerprint and facial recognition could connect biology with machine learning and data sciences.

Accommodations

English language learners (ELLs) in the class were accommodated with the use of visuals, such as the graphical examples of how to "crack" passwords, and opportunities to work with peers. Using special characters in Spanish provided an alternative method of creating strong passwords. Accommodations for students with individual education plans (IEPs) might include alternative forms of assessment, such as allowing a student to audio-record their reflection instead of writing a blog post, or having students work with a partner if a physical disability prevents a student from doing the hands-on graphical password cracking. More advanced students conduct individual research projects to explore metadata automatically embedded in digital images.

Conclusion

Using digital forensics in the classroom addresses the *Next Generation Science Standards* (NGSS Lead States 2013) crosscutting concept of the influence of science, engineering, and technology on society and the natural world. Digital forensics offers exciting opportunities for students interested in STEM careers. The subject can also help students understand how the metadata on their smartphones is shared with others and how they can protect their devices with secure passwords. The authors hope teachers will use this lesson to address these issues while also exploring law-enforcement techniques of the digital age. ■

Jason Harron (jasonharron@utexas.edu) is a doctoral student and teaching assistant, John Langdon is a UTeach preservice teacher, and Jennifer Gonzalez is a UTeach preservice teacher at the University of Texas at Austin; Scott Cater is a science teacher at David Crockett High School in Austin, Texas.

On the web

Android smartphone lock patterns: *www.androidauthority.com/lock-pattern-predictable-636267/*

Cellphone metadata: *http://news.stanford.edu/2016/05/16/stanford-computer-scientists-show-telephone-metadata-can-reveal-surprisingly-sensitive-personal-information*

Class blog with links to individual student blogs: *https://caterforensics.wordpress.com*

How secure is my password? *https://howsecureismypassword.net*

Smartphone password emulator: *http://tinyurl.com/caterforensics*

Video: How to see everything Google knows about you: *www.businessinsider.com/see-what-google-knows-about-you-2016-5*

References

Bertino, A.J. 2012. *Forensic science fundamentals and investigations.* Mason, OH: South-Western Cengage Learning.

Cartwright, J. 2016. Technology: Smartphone science. *Nature* 531 (7596): 669–671.

Kamarainen, A.M., S. Metcalf, T. Grotzer, A. Browne, D. Mazzuca, M.S. Tutwiler, and C. Dede. 2013. EcoMOBILE: Integrating augmented reality and probeware with environmental education field trips. *Computers & Education* 68: 545–556.

NGSS Lead States. 2013. *Next Generation Science Standards: For states, by states.* Washington, DC: National Academies Press.

Purcell, K., A. Heaps, J. Buchanan, and L. Friedrich. 2013. *How teachers are using technology at home and in their classrooms.* Washington, DC: Pew Internet and American Life Project.