Labeling "Things"

Joanna F. DeFranco and **Phil Laplante**, The Pennsylvania State University

Internet of Things (IoT) devices are in our homes. Many, unbeknownst to us, intrude on our privacy. Labeling IoT products to inform consumers, offering understandable and relevant information about the "things," is discussed.

Labels are everywhere. There are labels on supplement bottles to report percentages of the vitamins provided. Price stickers on cars show their features (for example, leather seats, Wi-Fi, high-end audio system, and so on). Most processed foods have labels for concerns such as sodium, calories, fat, country of origin, and so on. For the most part, for everyday items, there are labels, and they are usually easy to understand—in fact, government regulations standardize the content and look at these labels to make them so.

But "things," which are the main ingredient of the Internet of Things (IoT), might not be so well understood. The IoT is not necessarily an everyday item for everyone. In the IoT, "things" could be a software system, sensor, Wi-Fi connection, device, laptop, and so on.

So therein lies the challenge and also the opportunity. Can we label IoT "things" in the same manner as we label everyday consumer products such that, for example, a system integrator of "things" knows a priori something about what the composite system will do (from a behavioral perspective)?

It is likely that the answer is yes, but only if we can determine the standard measures that offer understandable and relevant information about "things." So, let's jump in and discuss this issue.

Consumer spending on IoT devices is on an upward trend and will reach US$1 trillion if it hasn't already. However, IoT market success should be celebrated cautiously as the security and privacy implications of bringing these trendy smart devices into your home are not insignificant.

Consumers should consider that market competition in this space sometimes causes more focus on product functionality and could shortcut the extremely important nonfunctional requirements: *security* and *privacy*. In addition, using some of these products in your home implies giving up your privacy—which, surprisingly, is not an enormous concern to many.

This article will focus on privacy specifically. Although security and privacy go hand in hand, we focused on security in a previous column and highlighted the agility of hackers. With new IoT devices, bad actors find vulnerabilities and quickly determine how to monetize them.[1] To easily differentiate these two concepts, a security

Digital Object Identifier 10.1109/MC.2023.3236170
Date of current version: 5 April 2023

EDITOR JOANNA F. DeFRANCO
The Pennsylvania State University; jfd104@psu.edu

vulnerability can be the situation some Ring camera owners experienced by using hacked passwords for their networks and devices (that is, hackers watching and talking to them through their cameras). Other times a privacy breach is an unintended feature of an IoT device because of video capture.

It is unfortunate that the average consumer isn't thinking about privacy as much as product functionality or isn't considering the privacy and security of the device he or she just purchased and placed in the home. Smart devices and the IoT introduce privacy vulnerabilities that did not exist in the past, and, more importantly, these vulnerabilities are exploitable by a much larger pool of threat actors. In the past, telephones might be wiretapped or, in rare cases, on-hook audio (obtaining audio from a landline phone even when it is not active) exploited, but these activities would have required a court order or significant skills and physical access. People who thought they might be spied on from another country through their vacuum cleaner or television would have been considered crazy then, but today they are just exercising appropriate caution.

Awareness needs to be heightened as this kind of violation is worse than a scheme to steal credit card or social security numbers. This is like a home invasion. If your location, pictures of your personal possessions, the layout of your home, and so on are all stolen, you can't change them as easily as a credit card number. Consumer awareness starts with education and sometimes legislation and laws. The awareness began with The Privacy Act of 1974, which was written to protect personally identifiable information (PII) collected by federal agencies. Privacy policies are now required to be included with products that collect PII. The policy specifies what information is collected and what will be done with it. However, how many privacy policies have you read? Is it understood that the policy doesn't mean you are protected? These policies are only a way for organizations to explain what they are doing with your information. Here are excerpts from a popular pet nanny camera (product name replaced with X):

> "When you set up the X Camera, we collect any audio, video or pictures you create, upload, save or share through our Services (the 'Content'). We process Content data according to your configurations and settings. We may also collect video and audit information of individuals when they pass in front of the camera or speak when the X Camera is on."...
>
> "We collect your geolocation data when you use our Services."

Those two things together are enough to get your house robbed. This statement could be refuted with the security measures to protect your data; however, your personal video isn't guaranteed security on that company's servers. Even if the company takes measures to keep it safe on their server, most of the time, third parties are the main security issue (that is, where your data are being sent for evaluation). In other words, the third party should be considered a weak link in the security chain.[1] Another argument might be that companies anonymize your data. Still, even if the anonymity is assured, predictive models have a high probability of revealing PII—therefore, anonymization is almost impossible.[2]

Another argument could be "the video in my home is of no value to a hacker." Consumers may not realize the value of these data. The value increases with companies wishing to improve their machine learning (ML) artificial intelligence (AI) algorithms. Much of the AI technology in our homes uses ML (algorithms that assess data to train the AI device). In other words, the device learns from consumers consenting to monitoring and data capturing by these devices—inside their homes.

Because of the cost, many companies also use the AI-as-a-service model so organizations can test ML continuously on the cloud.[3] What may not be realized is that part of the process involves humans to annotate (for example, categorize/label/contextualize) certain types of data. For US$20 an hour, humans are sometimes paid to annotate pictures and videos for ML purposes. These humans could be located anywhere in the world, and so can the ML algorithms. For example, the technology iRobot Roomba J7 images were sent to a third party that further sends the images to contracted workers to categorize the photos/video to train AI systems.[4] Some of these images were "compromising," of people and children inside homes. Some of these private photos ended up on social media. *Note*: The IRobot devices were "labeled," the homeowners agreed to let the Roombas monitor them (for the purpose of AI ML), and the paid contractors also signed agreements to remove sensitive photos and video—or maybe you opted in from a privacy statement.

Here's a scarier scenario. Any vacuuming robot or similar autonomous device, can, at your command, map each floor of your home and the placement of furniture. But it could just as easily identify other possessions, whether you have pets, estimate the number of inhabitants, create a schedule of comings and goings, and so on. In essence, the robot is conducting an ethnographic observation of your life in the home. At a minimum, this information can be used for marketing purposes, unwanted solicitations, and brushing scams (where sellers create fake accounts and order their own products to an address) or more nefariously to plan for home invasions, robberies, blackmail, and more.

LABELS

The United States has recognized the privacy issues concerning consumer-facing devices. An executive order was issued on 12 May 2021 on "Improving the Nation's Cybersecurity," addressing securing software development environments (https://www.federalregister.gov/documents/2021/05/17/2021-10460/improving-the-nations-cybersecurity). Part of the order outlines that the National Institute of Standards and Technology (NIST) will initiate pilot programs, informed by existing consumer product labeling programs, in an effort to educate consumers on the security capabilities of IoT devices and software development practices. The task includes incentivizing manufacturers to participate. In addition, together with the Federal Trade Commission, NIST is identifying IoT cybersecurity criteria for a consumer labeling program as well as secure software development practices or criteria for a consumer software labeling program. Subsequently, there is also forthcoming legislation, called the "Informing Consumers About Smart Devices Act," which will require manufacturers to let consumers know if there is a microphone or camera in an Internet-connected product.

Labeling isn't a new concept. Voas (2000) proposed software warranties or certifications to address software quality due to the differing types of software and target environments.[5] He suggested a framework for a certification to address the software assurance and integrity needs of the organization as well as a way to highlight the peculiarities of that software type. In 2021, Laplante recommended software labels to offer a consistent and coordinated way to assess the level of risk in software to decide if it needs to be labeled, like a food, drug, or hazardous material. It was further suggested that a label could expose important properties of the software to review its safety, security, privacy, and reliability.[6] For example, information should be available to the developers when reusing software components, such as something similar to a food ingredient label: amount of reused (modified and unmodified) code, amount of new code (handwritten and autogenerated by tools), amount of open source code (and type of license), software complexity, testing methodology, and so on. Additionally, the software could be labeled for carbon (power) consumption—perhaps a simple green–yellow–orange–red system for excessive power consumption in typical or exceptional operation profiles.

NIST has provided a white paper in response to the executive order, called "Recommended Criteria for Cybersecurity Labeling for Consumer Internet of Things (IoT) Products" (https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.02042022-2.pdf). This work discussed a binary label, indicating if a product has met a baseline security standard. The label could be in the form of a uniform resource locator or a scannable (QR) code that would lead the consumer to additional information, such as (summarized)

- intent and scope: to address potential misinterpretations
- product criteria: cybersecurity properties included in the baseline and how the criteria address, for example, security risks
- a glossary of technical terms written in simple English
- conformity assessment: evaluation of cybersecurity properties
- declaration of conformity: referring to the baseline criteria, including the date of the last label
- scope: the kinds of products eligible for the label and information to identify labeled products
- changing applicability: the current state of this product's labeling as new cybersecurity threats and vulnerabilities emerge
- security considerations and implications for end-of-life IoT products
- expectations for consumers: consumer responsibility in securing software and how their actions (or inactions) can impact the software's cybersecurity
- contact information for the labeling program.

The European Union's new Cyber Resilience Act will require manufacturers to provide consumer information on the security features of devices and how to securely configure them.[7] Similarly, the Cyber Security Agency of Singapore (CSA) launched the Cybersecurity Labelling Scheme for consumer smart devices (https://www.csa.gov.sg/Programmes/certification-and-labelling-schemes/cybersecurity-labelling-scheme/about-cls). Finland and Germany have signed an agreement indicating they recognize the label issued by the CSA.

Labeling provides awareness and education not only to consumers, but to the developers creating these devices. This entire effort could start small with popular devices, such as electronic doorbells and home security cameras, to include multiple risk levels.

The bottom line is that IoT device buyers need to be aware of the risk involved in utilizing these devices in their homes, and developers need a reminder of what is important to include in these products. In addition to labeling, improving the process of data capture and analysis should be addressed: How can human involvement be made more efficient and safer? C

ACKNOWLEDGMENT

The authors are appreciative of the valuable input for this article from Jeff Voas and Rick Kuhn at NIST.

REFERENCES

1. J. DeFranco and B. Maley, “Closing the security agility gap,” *Computer*, vol. 55, no. 8, pp. 100–102, Aug. 2022, doi: 10.1109/MC.2022.3169400.
2. N. Kshetri and J. DeFranco, “Is privacy dead?” *IEEE IT Professional*, vol. 22, no. 5, pp. 4–12, Oct. 2020, doi: 10.1109/MITP.2020.2992148.
3. V. Dey, “AI-as-a-service making artificial intelligence and data analytics more accessible and cost effective,” VentureBeat, San Francisco, CA, USA, Dec. 2022. Accessed: Jan. 2, 2023. [Online]. Available: https://venturebeat.com/ai/ai-as-a-service-makes-artificial-intelligence-and-data-analytics-more-accessible-and-cost-effective/
4. E. Guo, “A Roomba recorded a woman on the toilet. How did the screenshots end up on Facebook?” *MIT Technol. Rev.*, Dec. 2022. [Online]. Available: https://www.technologyreview.com/2022/12/19/1065306/roomba-irobot-robot-vacuums-artificial-intelligence-training-data-privacy/
5. J. M. Voas, “Limited software warranties,” in *Proc. 7th IEEE Int. Conf. Workshop Eng. Comput.-Based Syst. (ECBS)*, Apr. 2000, pp. 56–61, doi: 10.1109/ECBS.2000.839861.
6. P. Laplante, “Software labels,” *Computer*, vol. 54, no. 11, pp. 82–86, Nov. 2021, doi: 10.1109/MC.2021.3102360.
7. M. Nelson, “EU announces first ever move to legislate cybersecurity for IoT,” *IoT Bus. News*, Oct. 2022. [Online]. Available: https://iotbusinessnews.com/2022/10/12/63479-eu-announces-first-ever-move-to-legislate-cybersecurity-for-iot/

JOANNA F. DEFRANCO is an associate professor of software engineering, associate director of the D.Eng. in Engineering program at The Pennsylvania State University, Malvern, PA 19355 USA, and an associate editor in chief of *Computer*. Contact her at jfd104@psu.edu.

PHIL LAPLANTE is a professor of software and systems engineering at The Pennsylvania State University, State College, PA 16801 USA, a Fellow of IEEE, and an associate editor in chief of *Computer*. Contact him at plaplante@psu.edu.

Digital Object Identifier 10.1109/MC.2023.3254022