



# Somebody *IS* Watching Me: Cybersecurity, Privacy, and the Internet of Things

## PANELISTS

**Moderator: The Honorable Lisa S. Walsh**, Circuit Court Judge, 11<sup>th</sup> Judicial Circuit, Miami-Dade, Florida

**Galina Datskovsky, PhD., CRM, FAI**, CEO, Vaporstream

**Ronald J. Hedges**, Senior Counsel, Dentons US LLP

**Jennifer S. Granick**, Surveillance and Cybersecurity Counsel, Speech, Privacy, and Technology Project, ACLU

**Taa R. Grays**, Vice President and Associate General Counsel – Information Governance, MetLife

**Karen Johnson-McKewan**, Intellectual Property Litigation Partner, Orrick, Herrington & Sutcliffe, LLP

March 15, 2018



# LEARNING OBJECTIVES

Upon completing this session, you will be able to:

- Analyze privacy, security and compliance impacts of IoT devices.
- Apply information governance and eDiscovery techniques to real world.
- Properly analyze discovery challenges, privacy concerns and statutory requirements applicable to IoT technology

## DEFINITION

The **Internet of things (IoT)** is the inter-networking of physical devices, (also referred to as "connected devices" and "smart devices"), vehicles, buildings, and other items embedded with electronics, software, sensors, actuators, and network connectivity which enable these objects to collect and exchange data. [https://en.wikipedia.org/wiki/Internet\\_of\\_things](https://en.wikipedia.org/wiki/Internet_of_things)

## SOME RELEVANT STATISTICS

- Gartner forecasts that 20.4 billion connected things will be in use worldwide by 2020.

Source: Gartner press release, Feb 7, 2017

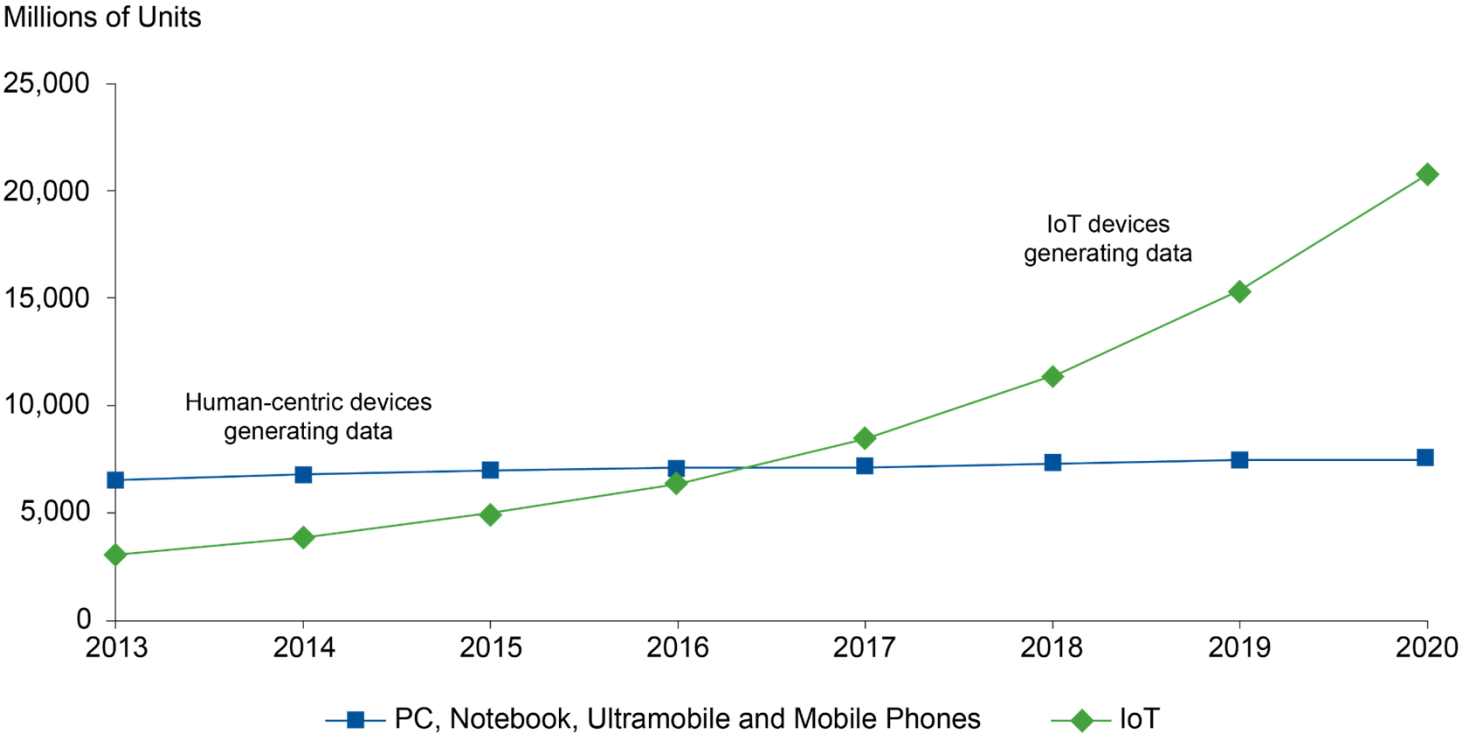
- By 2020, more than half of major new business processes and systems will incorporate some element of the IoT.

Source: Gartner, Why the Internet of Things Will Dwarf Social (Big Data),05 February 2016

- By 2020, the IoT will drive requirements in 25% of new information governance and master data management implementations

Source: Gartner, Data Risks in the Internet of Things Demand Extensive Information Governance,30 June 2016

# DATA GENERATION POTENTIAL OF IoT



# DATA VOLUMES PER DAY

Estimated IoT Data Volume Generation per Day			Estimated Total Data Generation Volume per Day			
	Data Volume (Individual Potential in MB)	Number of IoT-Enabled Things (by 2020)	"Thing" Population Generating Data (%)	MB	GB	TB
<b>Wearables</b>	1	1,294,200,000	30%	388,260,000	388,260	388
<b>Automotive — connected car</b>	20	220,100,000	35%	1,540,700,000	1,540,700	1,541
<b>Commercial aircraft (inclusive of major components such as engines)</b>	300,000	20,000	40%	2,400,000,000	2,400,000	2,400

# CONCERNS: PREPARATION FOR CASE LAW DISCUSSION

- Is the data scrubbed?
- How is it maintained?
- How easy is it to unmask?
- How easy is it to aggregate?
- Where is the data?
- Who owns it?



## HYPOTHETICAL

- Babysitter tells parents child fell down steps and sustained multiple fractures.
- Parents suspect abuse by the babysitter.
- Parents had purchased an interactive toy doll as companion for the toddler. The doll has cameras and microphone.
- Parents ask manufacturer for anything reported for 24 hour period prior to child's injuries.

## QUESTIONS ABOUT COLLECTION

Is it legal to audio record the babysitter without her consent and knowledge?

A: It depends on the state in which they live

Who is liable for that recording, the parents, the manufacturer, or both?

A: Possibly both.

Is it legal to video record the babysitter?

A: Generally, yes.

What data does the doll manufacturer likely have?

A: Likely nothing. These toys, like Amazon Alexa and Google Home require a wake word

## QUESTIONS ABOUT ACCESS

Assume there is a recording of the babysitter and child. Can the parents get that recording?

A: It depends. You have to know (1) how the manufacturer treats the with data and (2) how the manufacturer and information is classified under the Electronic Communications Privacy Act (ECPA).

Could the police get that recording? If so, with what legal process?

A: Yes. You have to answer both a statutory question (ECPA) and a constitutional question (4<sup>th</sup> Am, third party doctrine) to determine how. Search warrant, court order (18 USC 2703(d)), or subpoena.

Does it make a difference if the data is held in another country?

A: Microsoft v. United States (MSFT Ireland), CLOUD Act

# CARPENTER V. UNITED STATES

Question Presented:

Whether the warrantless seizure and search of historical cell phone records revealing the location and movements of a cell phone user over the course of 127 days is permitted by the Fourth Amendment.



# IoT-GENERATED DATA VOLUMES WILL BE AT THE CENTER OF LITIGATION

- Privacy and cybersecurity claims
- Consumer class actions
- Products liability claims (e.g., in autonomous vehicles, medical devices, artificial intelligence applications, etc.)
- Criminal proceedings

March 15, 2018



# THE CENTRAL DATA CHALLENGE IN CIVIL LITIGATION: DISCOVERY

- Data generated and collected through IoT devices will be in massive volumes
- “Burdensome and oppressive” objections will really mean something
- Judges must understand the nature of the work required to collect, cull, and produce responsive “documents”
- Challenges of requesting parties
- Will there be experts for discovery disputes? What degree of proof will be needed?

March 15, 2018



# SPECIAL CHALLENGES IN ESI DISCOVERY?

What does it mean when a lawyer tells you they have X Gigabytes or Terabytes of information?

- **Bit [a binary digit-either 0 or 1]**
- **Byte [8 bits]** 10 bytes = a single word
- **Kilobyte [1,000 bytes]** 2 kilobytes = a typewritten page
- **Megabyte [1,000,000 bytes]** 5 megabytes = the complete Shakespeare
- **Gigabyte [1,000,000,000 bytes]** 50 gigabytes = a floor of books
- **Terabyte [ $10^{12}$  bytes]** 10 terabytes = Library of Congress

# SPECIAL CHALLENGES IN DISCOVERY

- Voluminous and distributed
- Capable of taking many forms
- Contains non-apparent information (“metadata”)
- Created and maintained in complex systems

March 15, 2018





# SPECIAL CHALLENGES IN DISCOVERY – PLACES TO LOOK FOR ESI

- Personal computers at work and/or home
- Laptop computers, phones, and tablets
- IoT devices: where are they sending the data?
- Photocopiers
- Removable media (*i.e.*, flash drives)
- Third parties, including social media companies
- Body cameras
- Drones

March 15, 2018



## DISCOVERY HYPOTHETICAL

- Customer purchases, sets up, and registers a smart TV with manufacturer.
- Express condition of TV registration: TV manufacturer will use information only for software updates and new products.
- Months later, customer learns his data was sold, and the buyer of the data was hacked.
- Class action follows against TV, alleging unauthorized sale of customer data
- In discovery, plaintiffs seek **all data provided to third parties**
- The data collected included: pixels on the screen matched to content databases, viewing and consumer data from other connected IoT devices, IP addresses, physical addresses, and household demographics, including income and education levels.

# DISCOVERY CHALLENGES

Questions for the court and parties:

- The data collected were not “documents” in any traditional sense. How, and in what form, should they be produced?
- How does the requesting party access those data?
  - Special software required?
  - “Clean room” arrangements?
  - Protection of receiving party’s work product in reviewing data?
- What time frames should be imposed on production of these data?
- Does production lead to similar demands on the third parties who received it?

## KEY CASE LAW: *Zubulake v. UBS Warburg LLC* (“*Zubulake I*”), 217 F.R.D. 309 (S.D.N.Y. 2003)”

- Who will pay for restoring email from archival and backup sources?
- Distinction drawn between “accessible” and “inaccessible” sources
- Cost-shifting only available if source is found to be inaccessible







## *Zubulake I, cont'd.*: Cost-Shifting Factors

- Extent to which the request is tailored to discover relevant data
- Availability of the data from other sources
- Total cost of production, relative to the amount in controversy.
- Total cost of production, relative to the resources available to each party
- Relative ability and incentive for each party to control its own costs
- Importance of the issues at stake in the litigation
- Relative benefits to the parties in obtaining those data

# APPENDIX

## ACTIVATION: Manual, Always Ready, or Always On?

By their method of **activation**, consumer devices can be categorized as **manual, always ready, or always on**. In the past, most recording devices could be considered either on or off. Many new voice-based home assistants today can be considered "always ready" because they do not begin transmitting data off-site until they detect a wake phase.

METHOD OF ACTIVATION	 <b>MANUAL</b>	 <b>ALWAYS READY</b>	 <b>ALWAYS ON</b>
<b>BENEFITS</b>	<p>Device begins transmitting audio externally when a button or switch is pressed or held down.</p>  <p>Examples: Smart TVs (some); Mattel's Hello Barbie</p> <ul style="list-style-type: none"> <li>• ability to prevent unauthorized access through a hardware-linked microphone</li> <li>• avoids unintentional activation</li> </ul>	<p>Device processes locally to detect a "wake phrase**" which wakes/triggers the device to begin transmitting data.</p>  <p>Examples: Smart TVs (some); Home Assistant Devices, e.g. Google Home, Amazon Echo *Familiar wake phrases include: "Hey, Siri" "OK, Google," "Hey Cortana," and "Alexa"</p> <ul style="list-style-type: none"> <li>• convenience of verbal activation</li> <li>• accessibility for users with physical limitations, e.g. mobility or visual impairment</li> <li>• contextual responses, e.g. a home security system that alerts the owner and begins transmitting data when it detects a noise</li> </ul>	<p>Device transmits data 100% of the time on a standalone basis, and further processing occurs externally.</p>  <p>Examples: Home security systems; CCTV; Body-worn cameras; Baby monitors</p> <ul style="list-style-type: none"> <li>• enabling physical security for people and property</li> <li>• child safety</li> </ul>

# APPENDIX

## DATA TRANSMITTED

After a device is activated, it may sometimes transmit the full range of audible sounds (including voices), for example to enable cloud-based speech-to-text translation. However, other devices may not send audio at all, but instead may use the microphone to detect patterns and transmit other information about the user's environment.



**AUDIBLE TO HUMANS**

Most microphones only detect sounds within the normal range of human hearing to enable, e.g. voice commands, speech-to-text translation, or music recognition. Depending on the sensitivity, the microphone might also detect unintended background noises (such as dogs barking, or traffic sirens).

**SPECIALIZED RANGES**

Sophisticated microphones might sometimes be designed to capture only certain ranges of audio data, such as very low or very high (even inaudible) sounds. For example, a microphone could be designed to detect a hummingbird's wings or a dog whistle.

## NON-AUDIO

**PATTERN RECOGNITION**

Devices sometimes do not need to transmit audio recordings at all, but might use efficient local processing to detect sound patterns and convey data related to those patterns. For example, a city sensor might alert law enforcement when a "gunshot" pattern is detected.

**METADATA**

Data about when and how a device is used is known as "metadata," and may include, e.g., times and lengths of audio recordings, or where the recording took place. This data may not be as sensitive as a recording's content, but may nonetheless be revealing. For example, the times of day when a device is used may indicate when a person is typically at home.

# APPENDIX

## LEGAL PROTECTIONS

Laws protection audio data, especially voice communications, are sometimes robust — but also in flux as technologies evolve and courts grapple with the limitations of constitutional protection for data sent outside the home. Current applicable laws in the United States include:



- The Wiretap Act, 18 U.S. Code § 2511
- Federal Sectoral Laws for Sensitive Contexts or Populations, such as the Children’s Online Privacy Protection Act (COPPA) or Health Insurance Portability and Accountability Act (HIPPA)
- Federal Trade Commission (FTC)’s Section 5 Enforcement Authority
- State Unfair & Deceptive Practices (UDAP) Laws
- State Anti-Surveillance Statutes
- Civil Tort Remedies for Invasion of Privacy



## APPENDIX

**In a rapidly changing environment, trust is critical for developers seeking to innovate. Key privacy considerations include:**

**Data Security** — regardless of how a device is activated, if the data being transmitted is sensitive (e.g. voices or data from inside the home), strong security is paramount. Product developers should design for technical safeguards, such as limiting microphone sensitivity and range to the purpose of the device; enabling a hardware-linked on/off mute control; and filtering out unnecessary audio data at the point of collection.

**Prominent Visual and Audible Notice** — keeping in mind that users may not be comfortable with uses of their device’s microphone related to detection of acoustic events or ambient noise if they are not aware of those uses or how they work.

**Access to Information** — companies should make it easy for users to access and delete their information, and be transparent about any third-party disclosures, including government requests for access.

**Content vs. Metadata** — although fewer legal protections exist for metadata, companies should be aware of how patterns of use for home devices can be revealing and take steps to mitigate possible privacy risks.