UNIT 12 Audioscript

CONNECT to the topic page 112

Host: Welcome to *Tech Talk*. Today we're talking with data scientist Sonia Chen. I'd like to start by talking about data. We know that in general, *data* means "information." But what exactly is digital data?

Data scientist Sonia Chen: By digital data we mean information stored digitally—that is, stored by computers.

Host: So how do we measure digital data? For example, data is measured in "bytes." But what exactly does that mean?

Chen: A byte is a unit of data that is large enough to store one character—for example, one letter of the alphabet, such as a, b, or c.

Host: OK, that's pretty small!

Chen: Right. Which is why we usually use larger units to measure data. For example, one megabyte is over 1 million bytes. Or about the amount of information in one medium-sized book.

Host: I see.

Chen: And gigabyte is about 1,000 times bigger—1,024 megabytes, to be exact.

Host: Whoa. What are some larger units of measure?

Chen: Next is a terabyte. So if a megabyte is equal to one book, one terabyte is equal to about the amount of information you'd find in ALL the books in a large library.

Host: And is that the biggest?

Chen: No, no. It keeps going up, multiplied each time by 1,024—or 2 to the 10th power. So next is a petabyte and exabyte and zettabyte. All the way up to a yottabyte, which is, well, it's really, really big.

Host: OK, now that's a lot of data!

FOCUS your attention page 115

Speaker: And how much data do we use? It's easy to use a lot with out realizing it. For example, when I woke up this morning, I went to the gym and streamed some music on my favorite music site. I listened for about 30 minutes during my workout. And that used about 900 megabytes of data. After the gym, while I was on my way home, I checked my email on my phone. I had about 25 messages. Too many, right? But downloading those, that was only about 7 megabytes. Not too bad. Another thing I had to do was make several video calls—for a total of about 30 minutes. And that used almost 8,000 megabytes. Eight thousand! So these are some everyday ways that we use data. And clearly, each activity uses a different amount of data.

WATCH the lecture page 116

Professor Colin Edwards: **E01** In Information Technology, we often talk about data. Today, I'd like to talk about "big data." As you know, *data* simply means information or facts about a person or a topic. So, then what is *big data*? *Big data* means groups of data that are very large or very complex. Today I'd like to focus on three aspects of this topic. First, the ways that big data is collected and stored. Second, some everyday uses of big data. And third, some concerns about using big data, especially privacy concerns. **E02** First, the collection process. All kinds of data are collected all the time. But who is collecting this data? Well, banks collect information about their customers' credit card purchases. Insurance companies collect information about their patients' visits to the doctor. Internet providers collect information about what websites you visit. Entertainment companies collect information about what music you listen to and what movies and TV programs you watch. And the list goes on and on. What you do, what you buy, what you like, even who you like—it's all collected by someone. **E03** Information becomes data when we save it—when we store it for future use. How do we store data? We store it through *digitization*, which means converting or changing information, including images and sounds, into byte streams—which are sequences of bytes. **E04** Once the data is stored in byte streams, we can then use special software programs to access and analyze it—to find patterns and create models. Because of the power of computer technology, all of this data can be compared in many, many ways. And guess what? This creates *new* data. More and more data every moment. In fact, the amount of data in the world is increasing all of the time. Some say the amount of data doubles—increases *two times*—every two years. **E05** Now, let's look at some uses of data. Maybe you're thinking, "So, there's a lot of data in the world. So what?" Well, *one way* to look at big data is to consider the potential benefits for individuals. You may not realize it, but our lives are affected by big data every day—probably in some ways you don't even think about. For example, when I woke up this morning, I made coffee and then checked the weather. I wanted to know: Is it going to rain? Should I take my umbrella? And, thanks to big data, I saw that yes, the forecast called for rain. Forecasts come from weather scientists who collect large amounts of weather data—like temperature changes and cloud movement—analyze it with computers and then look for weather patterns, all so they can accurately predict the weather. And yes, help us make decisions. And then this morning I checked the news on my tablet. And guess what? My favorite news website had

topics and stories that interest me. How? Because it analyzes the reading patterns of thousands of other people using their website and finds stories that interest people like me. Oh, and it also tells me what my friends are reading because it knows who my friends are. That's a little weird, right? But, yes, it's also helpful for me—it helps me learn things. Another thing I wanted to do this morning was plan a trip. So I visited my favorite travel website. I entered the information about where and when I want to go. And the website quickly found the best available flights and hotels with the lowest prices. With just a few clicks, and before I'd finished my morning coffee, my trip was planned—all using big data. Another benefit for me—it saves time. After I booked my trip, I checked the map app on my phone to plan my trip to the airport. I entered the day and time I planned to go. Then big data analyzed the traffic patterns and told me the best route to the airport from my house and how much time it should take me to get there. So I won't have to worry about being late. And that helps me too: less stress. **E06** So those are some everyday uses of big data and some ways that big data helps us in our everyday lives. And clearly, big data does have benefits. But we also pay a price for these benefits. Big data is following us. And our right to privacy is in danger. Every time we use our phones, make a credit card purchase, or surf the web, we leave a trail of evidence. Same thing when you apply for a credit card, or insurance, or a loan. When you apply for a job or for a school. You're leaving a trail. And data collectors are following it. Our trail can be captured, stored, and "mined." Data "miners" can look into our personal lives, and they can also *share* this information. But is this fair? Who can collect this data? Who can use this data? And is the information about you accurate? These are some of the "big concerns" about big data.

HEAR the language page 118

1 Well, banks collect information about their customers' credit card purchases. Insurance companies collect information about their patients' visits to the doctor. Internet providers collect information about what websites you visit.

2 And the list goes on and on.

3 What you do, what you buy, what you like, even who you like—it's all collected by someone.

4 Information becomes data when we save it—when we store it for future use.

5 Once the data is stored in byte streams, we can then use special software programs to access and analyze it—to find patterns and create models.

6 This creates *new* data. More and more data every moment.

7 Forecasts come from weather scientists who collect large amounts of weather data—like temperature changes and cloud movement.

8 Every time we use our phones, make a credit card purchase, or surf the web, we leave a trail of evidence.

9 Same thing when you apply for a credit card, or insurance, or a loan.

10 But is this fair? Who can collect this data? Who can use this data? And is the information about you accurate?

TALK about the topic page 119

Shelley: So, big data. It's a big topic. There's a lot to think about. But one thing really stood out to me. And that was the amount of big data we use every day. We use big data all the time! Were you guys surprised by anything?

Kenzie: Well, I was.

Shelley: Yeah? What surprised you?

Kenzie: The variety—all those examples. The weather, the news stories, airplane flights, maps. I never knew it was all big data.

Ben: Yeah, and what surprised me was how the lecturer really focused on all the positive, helpful ways that big data helps us day to day.

Hugh: Yeah. Me, too. That was interesting.

Shelley: What do you mean?

Ben, Hugh: Who? Me?

Shelley: Ben, what do you mean "helpful ways"?

Ben: So, like, I knew all the dangers of big data, like, privacy issues, but I never really thought about the positives.

Shelley: That's a good point. It can be positive or negative. Hugh, what about you? What impressed you?

Hugh: The thing that impressed me was, like you said, the *amount* of data. And how fast it grows. It increases more and more each year. And so I wonder, when will we have enough data? Or *will* we ever have enough?

Ben: Right, I mean the amount of data, just growing all the time—that just blows my mind.

Kenzie: And it's kind of creepy. Like probably right now this online study session, someone's collecting data on it. So we're all becoming a part of big data.