

## MY WEIRD PROMPTS

Podcast Transcript

### EPISODE #254

# The Digital Shadow: Uncovering the Power of Metadata

Published January 20, 2026 • Runtime: 18:08

<https://myweirdprompts.com/episode/digital-shadow-metadata-privacy/>

## EPISODE SYNOPSIS

Think your "plain text" files are private? In this episode, Herman and Corn dive deep into the world of metadata—the invisible "digital shadow" that follows every photo, document, and interaction you create. From the ancient Library of Alexandria to the modern surveillance capitalism of 2026, they explore why metadata is essential for technology, how it's used to train AI, and why your "anonymized" data might not be as secret as you think. Join the conversation as they peel back the layers of the digital world to reveal the infrastructure that maps our lives.

## DANIEL'S PROMPT

### Daniel

How did metadata come into existence? If I create a simple digital file, like a plain text document, how much metadata is generated? Is the creation of metadata inevitable in modern computing? Are technology vendors becoming more aggressive about collecting metadata, or is there a shift toward greater privacy awareness? Additionally, why is metadata often left unencrypted even when the file content itself is protected? Ultimately, how much metadata are we generating on a daily basis?

# TRANSCRIPT

## Corn

You know, Herman, I was looking at a photo I took the other day of that old stone archway near the Jaffa Gate. It is just a simple image file on my phone, but when I swiped up, it told me exactly where I was standing, the exact second I pressed the shutter, the focal length of the lens, and even the altitude. It made me realize that for every bit of content we create, there is this invisible layer of context trailing behind it like a digital shadow.

## Herman

That is the perfect way to describe it, Corn. A digital shadow. Herman Poppleberry here, and I have to say, our housemate Daniel really hit on a fundamental nerve with this prompt. He was asking about where metadata comes from and just how much of it we are generating when we do something as simple as typing a note. It is one of those topics that feels technical on the surface, but once you peel it back, it touches on everything from the history of libraries to the modern surveillance state.

## Corn

It is fascinating because most people think of their files as just the stuff they put in them. The words in the document, the pixels in the photo. But Daniel's question about whether metadata is inevitable is a great starting point. Is it just a byproduct of modern computing, or was it a conscious design choice?

## Herman

It is actually both, but the concept is much older than computers. We often think of metadata as a high-tech invention, but the term itself was coined back in nineteen sixty-eight by a computer scientist named Philip Bagley. Long before him, though, humans were already doing this. If you think back to the Great Library of Alexandria around two hundred eighty B-C, librarians were attaching small tags to the end of scrolls with the title and author. That is metadata. The first library card catalog in seventeen ninety-one used playing cards to index books. The book itself is the data, but the card that tells you the author, the year, and the shelf location? That is metadata. It is data about data. We have always needed it to organize information. Without it, a library is just a giant pile of paper. In the digital world, that necessity just got automated and expanded by several orders of magnitude.

### Corn

So, when we transitioned from physical cards to digital bits, we just brought that organizational logic with us. But Daniel asked a very specific question that I want to dig into. If I open up a basic text editor, something like Notepad on Windows or Kate on Linux, and I just type "Hello world" and hit save, how much metadata are we talking about for that tiny file?

### Herman

That is where it gets really interesting, Corn. Most people think a plain text file is the cleanest form of data, and in a way, it is. But even a zero-byte file still has metadata. The moment you save that file to a disk, the operating system has to create an entry for it. On a Linux system, we talk about something called an inode. This is a data structure that stores everything about the file except its name and the actual content. On Windows, using the N-T-F-S file system, it is stored in the Master File Table.

### Corn

Okay, so what is actually inside that record for my "Hello world" file?

### Herman

You are looking at a lot of specifics. You have the file size, the owner's user identification, and the group identification. You have the permissions—who can read, write, or execute it. Then you have the timestamps. You have the creation time, the last access time, and the last modification time. In some file systems, there is even a changed time which tracks when the metadata itself was last altered. So even before we get into the application level, the operating system is already recording who you are and exactly when you were working on that file.

### Corn

And that is just for a local file on my hard drive. What happens if I create that same text file in a cloud environment, like Google Docs or Microsoft three sixty-five?

### Herman

Oh, then the metadata explodes into what we call telemetry. In a cloud environment, you are not just tracking a file; you are tracking a session. They are recording your internet protocol address, your browser version, your geographic location, how long you had the document open, how many times you paused while typing, and every single revision character by character. By the time you finish a one-page document, the metadata probably weighs ten times more than the actual text you wrote. In fact, by twenty twenty-five, the average person was making over four thousand nine hundred digital interactions every single day. Each one of those is a metadata event.

### Corn

That brings up Daniel's point about whether this is inevitable. It sounds like if we want features like undo history, or the ability to search for files by date, or even just basic security permissions, we cannot escape metadata. It is the price of functionality.

### Herman

Precisely. You cannot have a searchable, multi-user, secure operating system without metadata. It is the glue. But there is a second-order effect here. Because metadata is so useful for the computer, it becomes incredibly useful for anyone who wants to track what you are doing. Remember back in episode one hundred eighty-four when we talked about the Open Systems Interconnection model? Metadata is what allows those layers to talk to each other.

### Corn

Right, and that leads to another part of Daniel's prompt. He asked why metadata is often left unencrypted even when the content is protected. This seems like a massive security hole. If I send an encrypted email, the contents are safe, but the To and From fields and the timestamp are often visible. Why is that?

## Herman

Think of it like a physical letter, Corn. You can write your letter in a secret code and put it inside a titanium box, but you still have to write the destination address on the outside, or the mailman won't know where to take it. The routers and servers that make up the internet are like that mailman. They need the headers to know where to route the packets. If you encrypt the routing information, the network literally stops working. However, we are seeing a shift. There is a new standard called Encrypted Client Hello, or E-C-H, which is finally starting to close that gap by encrypting the server name you are connecting to. It is like putting that titanium box inside a second, generic envelope so the mailman only knows which building it is going to, not which specific person.

## Corn

It feels like a massive trade-off. We get this incredible global connectivity, but the cost is that every envelope we send is being logged. And this brings us to the question of whether technology vendors are becoming more aggressive about collecting this stuff. What is the trend you are seeing in the research, Herman?

## Herman

It is a tale of two cities. On one hand, you have the surveillance capitalism model. Companies like Google and Meta have built empires by mining metadata. They don't necessarily need to read your private messages if they know who you talk to, how often, and from where. That metadata is often more predictive of your behavior than the actual content. And in twenty twenty-six, metadata has become the ultimate training set for artificial intelligence. We call it context engineering. If you want to train an A-I to understand human social dynamics, you need the metadata that shows the hierarchy and the response times.

## Corn

But then on the other hand, we have the privacy-centric move, right?

## Herman

Exactly. We are seeing a massive regulatory push. The E-U A-I Act, which fully implemented in August of twenty twenty-six, and the E-U Data Act from twenty twenty-five are forcing companies to be much more transparent about what they collect. We have apps like Signal that specifically engineer their systems to avoid keeping metadata. They famously could only provide a creation date and a last connection date when subpoenaed. So, we are at a fork in the road. Most mainstream tech is getting hungrier for metadata to fuel A-I, while a vocal niche is trying to starve the beast.

### **Corn**

It is staggering how much we generate. Give us the breakdown, Herman. I am ready to be slightly terrified.

### **Herman**

Well, let's look at a typical smartphone user. Every time your phone checks for a signal, it is a metadata event. Some estimates suggest that a single smartphone user generates over four gigabytes of network-related data every single day. Now, that is not all metadata, but a huge portion of it is the background chatter of your digital life. Researchers have shown that you can uniquely identify a person out of a dataset of millions using just four spatio-temporal points. That is just four instances of where were you at what time.

### **Corn**

That really puts the anonymity of metadata into perspective. People often say, "Oh, don't worry, the data is anonymized," but if the metadata is rich enough, anonymity is an illusion. You can't really hide in a crowd if your shadow is unique to you.

### **Herman**

That is the big misconception. In many legal jurisdictions, the police need a higher level of authorization to intercept the content of a call than they do to get the call detail records. But for an investigator, the metadata is often more useful. It shows the network. It tells the story of your life without ever needing to hear a single word you said. And metadata has a much longer shelf life. It is small and structured, so it is very cheap to store forever. A company might delete your old video uploads to save space, but they will keep the metadata about those uploads until the end of time.

### **Corn**

So, Daniel asked if this is a shift toward greater privacy awareness. Do you think we are actually making progress?

### Herman

We are definitely more aware. Ten years ago, metadata was a word only nerds used. Now, it is part of the public discourse. But the sheer volume of devices is growing faster than our ability to regulate them. Think about the Internet of Things. Your smart fridge, your lightbulbs, your thermostat. They are all metadata factories. It is a race between engineers developing zero-knowledge proofs and the drive for seamless technology that requires more background data to function. If I want my house to know I am home, I have to give up the metadata of my location.

### Corn

Convenience versus privacy. And for most people, convenience wins every time. But the real takeaway is that metadata is not extra information. It is the primary information of the digital age. It is the map of our lives.

### Herman

It really is. And if someone wanted to actually see this metadata for themselves, I recommend a tool called ExifTool for photos. It is a command-line application that can read meta information in a huge variety of files. If you run it on a photo you took with your smartphone, you will see everything from the software version to the direction the camera was pointing.

### Corn

And for documents?

### Herman

For documents, you can often just change the file extension to dot zip and open it up. Modern Word or Google Doc files are actually just zipped folders full of E-M-L files. If you dig through, you will find files dedicated entirely to app metadata. You can see the names of every person who ever edited the document, the total editing time in minutes, and even the names of the printers the document was sent to.

### Corn

That makes it tangible. It is not an abstract concept; it is literally written into the file structure. I think we have covered a lot of ground here, from library cards to encrypted envelopes. It is clear that metadata is the infrastructure of our digital world.

### Herman

It really is. And this whole discussion is a form of metadata for our own lives, right? This recording, the length of it, the date we recorded it, the fact that we are two brothers talking in Jerusalem. It all gets logged.

### Corn

Speaking of which, if you are listening to this on Spotify or your favorite podcast app, you are generating some metadata right now. You are telling the platform what you like and how long you listened. If you made it to the end, we would really appreciate a quick review. It helps the algorithms understand that this is the kind of content people want to hear. It is the good kind of metadata, at least for us.

### Herman

Definitely. A quick rating or a comment really helps the show reach new people. And if you want to get in touch, you can always find us at our website, [myweirdprompts dot com](http://myweirdprompts.com). We have the full archive there, including that episode one hundred eighty-four we mentioned earlier.

### Corn

Thanks to Daniel for the prompt. I think I am going to go check the metadata on that Jaffa Gate photo again and see if I can find anything else hidden in the margins.

### Herman

Just don't get too lost in the weeds, Corn. Sometimes the photo is just a photo, even if the metadata says it is a three point five megabyte record of a Tuesday afternoon.

**Corn**

Fair enough. Well, this has been My Weird Prompts. I am Corn.

**Herman**

And I am Herman Poppleberry.

**Corn**

Thanks for listening, everyone. We will talk to you next week.

**Herman**

See ya!