Hi, my name is Wendy Peterson. I'm a Vermont resident and a principal software engineer at Red Hat, which is part of IBM. I work on OpenShift.

I've been working in the software industry for about 25 years, and I'm one of the maintainers of Kubernetes, which is widely used by large companies and public institutions to run cloud and data systems.

I also organize locally in Vermont, so I hear regularly from people who are concerned about how their data is being used and who don't feel like they have much control over it.

My work focuses on systems that store, process, and move data at scale. So when I talk about this, I'm describing how these systems actually behave in practice.

I'm here to support keeping this bill as strong as it currently is.

I also want to share some personal context. I'm a trans woman. And right now, that's something that is being actively targeted at the federal level. So for me, data privacy is about risk. It's about whether someone can build a profile of my life - where I go, what care I seek - without my knowledge or consent.

This bill defines "sensitive data" to include things like gender identity, health data, and precise geolocation. To make that concrete - imagine a phone app collecting location data. If that app sees repeated visits to a specific clinic, and combines that with other signals, it can infer things about a person's health or identity, even if that person never directly shared that information. That's how modern systems work. They don't rely on a single piece of data. They combine many small signals and draw conclusions.

The bill also includes the concept of derived data - where systems take those signals and turn them into predictions about a person. That's especially relevant with AI. A lot of modern systems are designed to do exactly that - take large amounts of data and learn patterns from it. You can't build responsible AI on top of irresponsible data practices. If systems are trained on large amounts of sensitive or loosely controlled data, the risks increase - both for individuals and for the organizations using those systems. That's why data minimization matters. In practice that means collecting the data you actually need to provide a service - and not collecting everything else "just in case."

From a technical perspective, a "necessity analysis." is not a new or unusual requirement - it's just asking a simple question: do we actually need this data to provide the service? That's already how responsible systems are designed. For example: If you're processing a purchase, you need enough information to complete that transaction. You don't need precise location data, or unrelated behavioral data, to do that. In engineering, that question - "do we actually need this?" - is part of how systems are designed and approved. So this isn't introducing a new kind of burden. It's formalizing a practice that already exists.

I know there are concerns about how broadly this applies. This bill isn't triggered just by being a business. It's triggered by specific data practices. In particular: handling data at scale, processing sensitive data, or selling personal data. Those are not things that most small businesses or typical retail operations are doing.

To make that concrete: A local store processing purchases and maintaining customer relationships is using first-party data for a specific purpose. That's normal, and this bill allows that. What's different is when systems: track behavior across multiple websites, combine that with location or other signals, and build profiles that are shared or sold. That's a completely different category of system, both technically and in how data is used. And that's what this bill is addressing.

Another example of the kind of system this bill is addressing is when third-party services are used to profile customers using AI. For example, a retailer might use a service that analyzes customer behavior and assigns a risk score - for fraud, returns, or other purposes. On its own, using data to complete a transaction or prevent fraud can be reasonable. Where it becomes higher risk is when those profiles are reused across contexts, combined with other data sources, or shared beyond that original transaction. At that point, you're no longer just processing a purchase - you're building and potentially distributing a profile about a person. And that's the kind of practice this bill is designed to put boundaries around, especially when sensitive data or broader behavioral patterns are involved.

There are several versions of this bill. From a technical perspective, I think that some concerns comes from uncertainty about how these rules apply. But S.71 already reflects a balance. It preserves normal business activity, while still addressing higher-risk data practices - especially when sensitive data is involved.

I also want to highlight that this approach is actually good for businesses. Every piece of sensitive data you collect becomes something you have to store, secure, and manage. So one of the most effective ways to reduce risk is simply to not collect or retain data that you don't need. From a business perspective, that also reduces cost. Less data means: less infrastructure, less system complexity, and less exposure if something goes wrong Because when a breach happens - and they do - the cost is directly tied to how much sensitive data you're holding. So data minimization isn't just a privacy principle. It's a practical way for businesses to reduce both risk and cost.

You're also seeing this direction across the industry. There's a growing focus on data sovereignty - the idea that people should have meaningful control over how their data is collected and used. Companies like IBM are investing in this because it's becoming an expected baseline. This bill is aligned with that direction.

On advertising - there's a perception that companies need detailed personal data to advertise effectively. In practice, that's often overstated. For example, placing an ad next to relevant content - like advertising hiking gear on an outdoor website - works without needing a detailed personal profile. That's called contextual advertising, and this bill still

allows it. What the bill limits is the use of sensitive personal data in those systems. I think that's a reasonable boundary.

I also understand that the bill has already been negotiated and that some provisions have been weakened - including removing the private right of action and relying on the Attorney General for enforcement. Even with that, the remaining protections - especially around sensitive data - are still meaningful.

But only if they're not weakened further. Because for a lot of people, this isn't abstract. It's about whether their health decisions can be tracked. Whether their identity can be profiled. Whether their movements can be monitored and sold. This bill takes concrete steps to limit that, including banning the sale of sensitive data. From both a technical perspective and a personal one, I think those protections are important. And I'd ask you to keep the bill as strong as it currently is.

Thank you.