Detailed Description of Use Case — Synthetic Transcript Generation

1. USE CASE

Axon is developing an internal capability to generate **synthetic body-worn camera (BWC) transcripts** that can be used to evaluate and improve Axon products.

To build reliable and accurate products, Axon teams often need access to large collections of transcript data that reflect the kinds of conversations that occur during real-world interactions. However, Axon is committed to protecting the privacy of officers, agencies, and community members. To support this goal, Axon uses **differential privacy techniques** to train a model that can generate realistic but entirely synthetic transcripts.

During training, the model learns general language patterns from automatically generated BWC transcripts. Differential privacy techniques ensure that the model cannot memorize or reproduce individual transcripts. Once trained, the model can generate **new transcripts that resemble real conversations but do not correspond to real incidents, individuals, or locations**.

These synthetic transcripts are used internally by Axon teams to:

- Evaluate new and existing products
- Test and improve product performance
- Generate large datasets for product development and quality assurance

Customers do not interact with this model or its outputs directly. The system operates entirely as an **internal tool used to improve Axon products**.

---

2. TYPE OF AGENCY CONTENT (E.G. DATA ACCESSED)

Axon accesses **automatically generated transcripts derived from body-worn camera audio recordings**.
These transcripts may represent full interactions captured during body-worn camera recordings.

This use case involves access only to:

- Automatically generated **BWC transcript text**

This use case **does not involve access to**:

- Raw audio recordings
- Video footage
- Images

---

3. WHAT IS THE CUSTOMER BENEFIT?

Synthetic transcript generation allows Axon to build and evaluate products using **large, diverse datasets without exposing real customer interactions**.

By generating synthetic transcripts that reflect the patterns of real-world conversations, Axon teams can:

- Evaluate the performance of new and existing products
- Identify product gaps or edge cases
- Improve the accuracy and reliability of AI-powered features
- Develop new capabilities more quickly

Because these datasets are synthetic, Axon can scale development and testing while **protecting the privacy of officers, agencies, and community members**.

---

4. WHAT DATA WILL AXON ACCESS?

Axon will access **automatically generated transcripts from body-worn camera recordings**.

These transcripts are used **only during the training process** to teach the model general language patterns that occur in BWC interactions.

During training:

1. Transcripts are accessed within the **customer boundary environment**.
2. The transcripts are used to train a differentially private model.
3. Once training is complete, only the **trained model weights** are exported.

No transcript data is exported outside the customer boundary.

---

5. HOW WILL AXON USE YOUR DATA?

Axon uses transcripts only to train a **differentially private transcript generation model**.

The training process works as follows:

1. Automatically generated BWC transcripts are used during model training to learn general language patterns that occur during real-world interactions.
2. Differential privacy techniques are applied during training to prevent the model from memorizing individual transcripts.
3. After training, the resulting model is used to generate **entirely synthetic transcripts** that resemble real conversations but do not reproduce real events.
4. These synthetic transcripts are used internally to test, evaluate, and improve Axon products.

The original transcripts are not exported or reused outside the training environment.

---

6. WHAT PRIVACY PRESERVING TECHNIQUE WILL BE USED?

Axon uses **differential privacy** to protect customer data during model training.

Differential privacy is a technique that ensures models learn general patterns from data without retaining or revealing information about specific records.

During training:

1. **Noise is intentionally introduced during the training process** so the model learns overall language patterns rather

than memorizing individual transcripts.

2. A formal **privacy budget** is used to measure and limit the amount of information the model can learn from any single transcript.
3. Axon uses **canary testing methods** to verify that the model does not memorize or reproduce training data. Current testing shows less than **1% potential memorization** of training data.
4. These safeguards ensure the model generates new transcripts that are statistically derived from training data but **are designed not to reproduce real transcripts**.

---

7. HOW MUCH DATA AND FOR HOW LONG?

Training datasets may include **tens of thousands of transcripts**.

Transcripts are accessed only during the training process, which may take up to **one month** to complete.

Once training is finished:

- The transcripts are no longer used.
- Only the trained model weights are retained.

---

8. PRESERVATION OF ORIGINAL CONTENT & TEMPORARY COPIES

Original evidence and transcripts remain **unchanged in Axon Evidence**.

The system reads transcripts from controlled environments during training but does not modify or replace the original records. Any temporary processing copies created during training remain within the secure training environment.

---

9. CAN I GET MORE INFORMATION ABOUT WHAT AXON IS DOING AND WHY?

Yes. If you have questions about this use case or how your data is handled, please contact **aceip@axon.com**.

---

10. AM I ABLE TO WITHDRAW MY AGENCY FROM THIS USE CASE AND FROM ACEIP ALTOGETHER? WHAT WILL YOU DO WITH MY DATA IF I WITHDRAW AFTER THE FACT?

Absolutely. Participation in this use case is part of **ACEIP participation**.

If your agency chooses to withdraw from ACEIP:

- Axon will **no longer use your transcripts for future training**.
- Previously trained models may retain statistical influence from earlier training, but no new data from your agency will be used going forward.
- Your original evidence and transcripts will remain unchanged in Axon Evidence.

---

11. DO YOU HAVE EXAMPLES OF WHAT DATA AXON MAY OR MAY NOT EXTRACT?

Axon uses transcripts only during the training process to build a model that generates synthetic transcripts.

The example below illustrates how the process works. All examples are fictional and provided only for illustration.

Step 1: Original transcript (used during training only)
“Hey, can you step over here for a moment? We got a call about a disturbance near the parking lot.”

Step 2: Differentially private training
The training system learns general language patterns from many transcripts while applying differential privacy techniques to prevent memorization of individual conversations.

Step 3: Synthetic transcript generated by the model
“Could you come over here for a second? We received a report about a situation near the building.”

The generated transcript reflects common language patterns found in real interactions but **does not represent a real person, event, or location**.