Layers of CAN Data

SYSTEMS ENGINEERING
COLORADO STATE UNIVERSITY

Jeremy Daily
Colorado State University

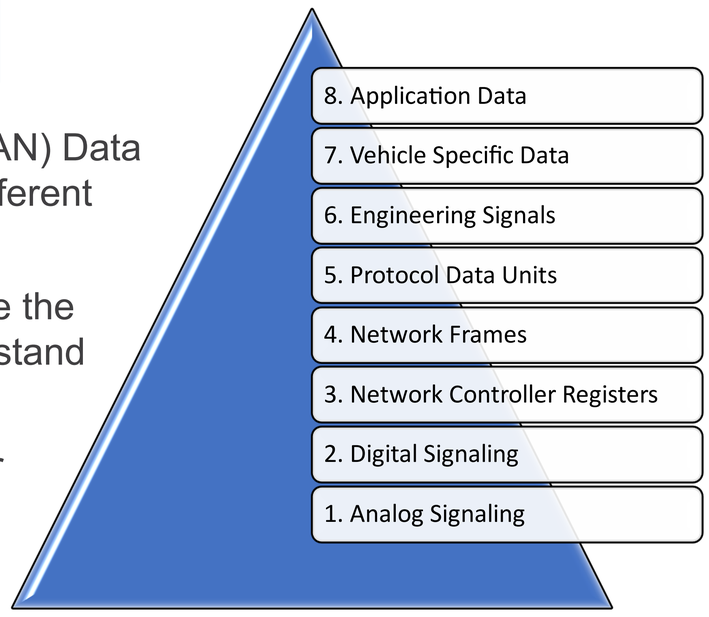
INCOSE | WSRC Western States Regional Conference

ABSTRACT

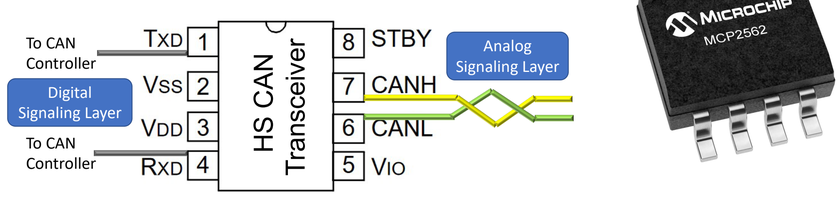
Advances in algorithms and utilization of vehicles is driven by the availability of training and test data. These data sets have traditionally been collected by the creators of vehicle technologies, but the pace of advancement has created a need for more diverse vehicular data sets. However, classification labeling of so-called vehicle data lacks industry cohesion. Furthermore, individuals and groups requesting data may have different notions for the data they need compared to what the producer of the data can generate. This poster addresses these discrepancies by introducing a vehicle data model and description to aid both producers and consumers of the data by facilitating communication through a common vocabulary. A layered approach to the data taxonomy is described in parallel with examples that accompany each layer. Furthermore, the context of the produced data is paramount to its utility, so additional recommendations for capturing and describing situational context are presented. An example of contextual data is presented based on an experiment with a Kenworth T270 research truck. Quality vehicle network data should be findable, accessible, interoperable, and reusable (FAIR). By following the recommendations and examples, vehicular datasets can be of higher quality, which translates to greater utility for projects and for promoting better archives for future research.

PYRAMID OF DATA

- Controller Area Network (CAN) Data means different things in different contexts
- A pyramid model to describe the data can help people understand what comprises data
- Layers build on one another
- The storage requirements are largest at the bottom



ANALOG AND DIGITAL DATA



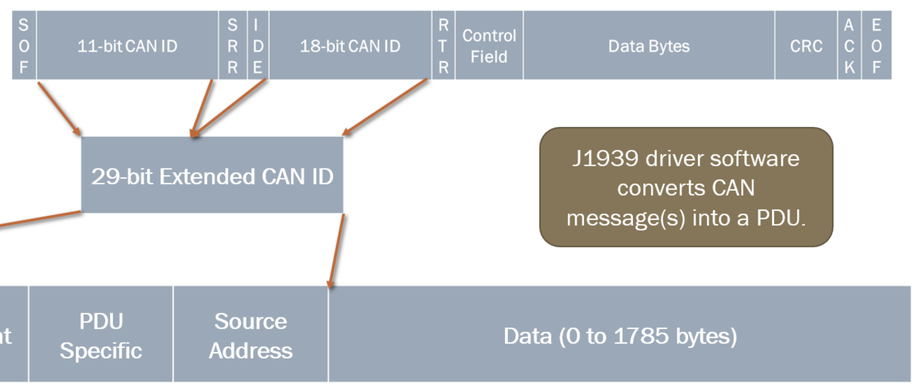
- Analog signaling is measured with an oscilloscope
- Digital wave forms show logical transitions
- Noise, timing, framing, and signaling artifacts can be analyzed

PROTOCOL DATA UNITS

- Encrypted streams may be decrypted
- Encapsulation and decapsulation

A J1939 message has all the elements in the protocol data unit (PDU)

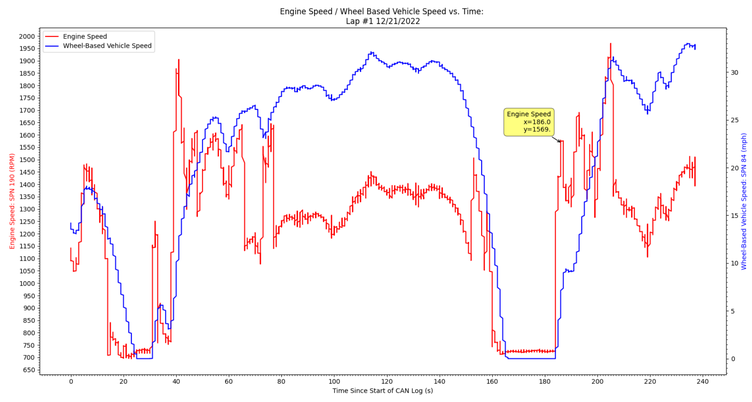
- 3-bit Priority
- 1-bit Extended Data Page (EDP)
- 1-bit Data Page (DP)
- 8-bit PDU Format (PF)
- 8-bit PDU Specific (PS)
- 8-bit Source Address (SA)
- Data Field up to 1785 bytes



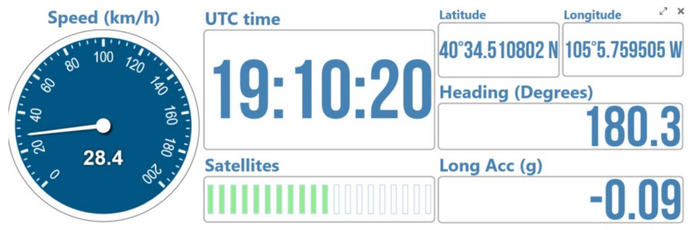
This work was funded by the National Science Foundation under award number 2213735. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

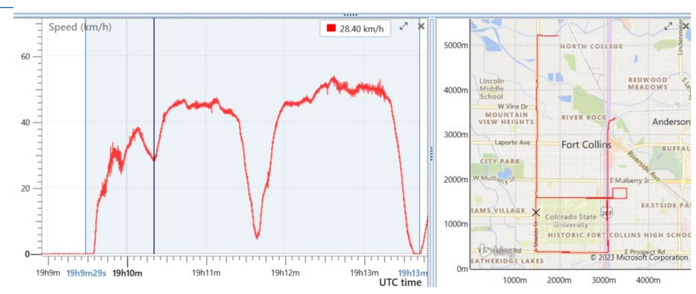
ENGINEERING SIGNALS

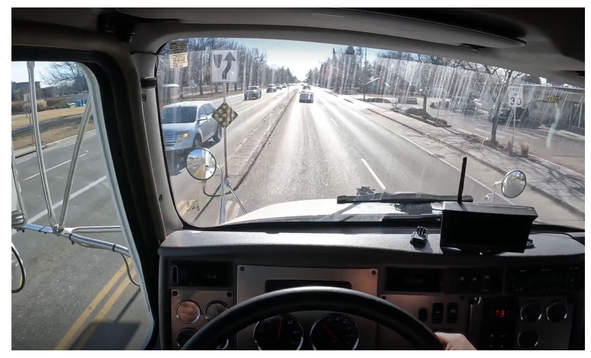
- Decode bits and bytes to bring meaning to the data
- Time history signals show operational characteristics
- Utilizes engineering units
- Meaning may be proprietary



CONTEXT DATA

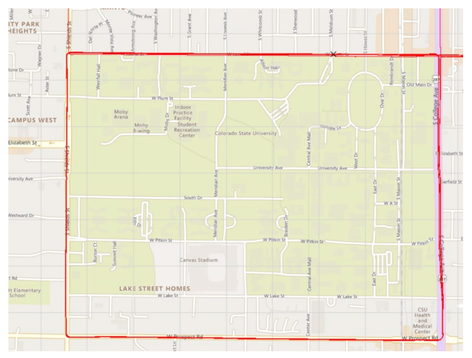






GoPro video (Top)

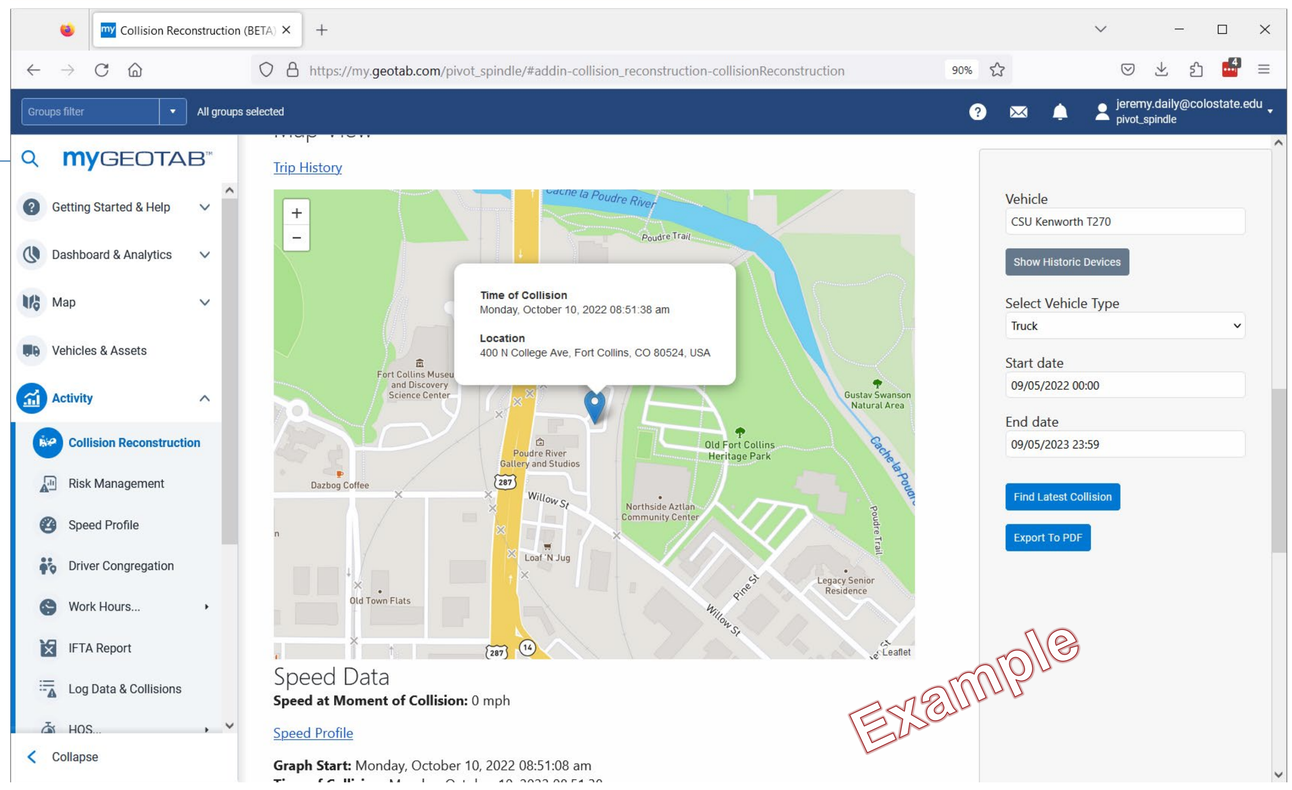
VBox3i GPS (Left)



Correlated time, geolocation, photos, and video provide additional value and meaning for the data

VEHICLE AND APPLICATION DATA

Telematics Service Provider (TSP) aggregates and presents data



CONCLUSIONS

Regardless of the type of data collected, to make it more useful, consider making FAIR:

- Findable – add metadata and keywords to enable better searches
- Accessible – have the ability to interpret and use the data
- Interoperable – the data is not tied to a specific platform or tool
- Resuable – free from royalties and encumbered from licenses

CONTACTS / REFERENCES

Jeremy.Daily@colostate.edu

D. Balenson, C. Papadopoulos, G. Atkinson, T. Guild, S. Prowell, and S. C. Hollifield, “Paving the Road to Future Automotive Research Datasets: Challenges and Opportunities,” Virtually on November 18-19, 2021, Workshop, Aug. 2022. [Online]. Available: https://drive.google.com/file/u/1/d/191ckNAtCriC6q9V4kzC7m53xl5wKJ9AX/view

D. Balenson, T. Benzel, E. Eide, D. Emmerich, D. Johnson, J. Mirkovic, and L. Tinnel, “Toward Findable, Accessible, Interoperable, and Reusable Cybersecurity Artifacts,” in Cyber Security Experimentation and Test Workshop. Virtual CA USA: ACM, Aug. 2022, pp. 65–70. [Online]. Available: https://dl.acm.org/doi/10.1145/3546096.35461

M. T. Campo, S. Mukherjee, and J. Daily, “Real-Time Network Defense of SAE J1939 Address Claim Attacks,” SAE International Journal of Commercial Vehicles, vol. 14, no. 3, pp. 02–14–03–0026, Aug. 2021. [Online]. Available: https://www.sae.org/content/02-14-03-0026/

P.-S. Murvay and B. Groza, “Source Identification Using Signal Characteristics in Controller Area Networks,” IEEE Signal Processing Letters, vol. 21, no. 4, pp. 395–399, Apr. 2014, conference Name: IEEE Signal Processing Letters