Data Collection and Management: How We Collect and Manage Data in the Field, Transmit Data, and Create Information

Ryan McWilliams, International Engineering
Andie Berry-Morris, Holland LP

You likely <u>often</u> wonder…

- **Why do we generate data?**
- **What types of systems collect data?**
- **What kinds of data are collected?**
- **How is data managed at the point of collection?**
- **Where does data go and how is it transmitted?**
- **How is data used to CREATE INFORMATION?**
- **How are data and information shared?**

So, Why DO We Data?

- **Generating data is driven by our need to monitor and learn more about the conditions of our assets and infrastructure**

- **We generate data so that we can CREATE INFORMATION**

- **INFORMATION drives our decision making process– e.g., maintaining versus replacing assets, changing our processes, implementing improvements**

- **INFORMATION gets us closer and closer to the truth– the HOW and the WHY**

What Do Mobile Systems Do?

- Mobile systems are traditionally used to gather data about structural assets and components:
  - Track
  - Rail
  - Wayside
  - Structures/In-Track Assets
- They are typically mounted to dedicated rail bound platforms, hi-rail vehicles, revenue cars, or even aerial vessels (drones, etc).

Mobile Systems

CSX's TGC3– Plasser USA and Gateway Rail Services

BNSF Car 90 UGMS– Holland

Ultrasonic Platforms

Photos courtesy of Union Pacific and Sperry Rail Service

Vehicle Based Systems

Video Imaging Drone

Holland Lightweight Geo

Holland TrackSTAR®

ENSCO V/TI

ENSCO Joint Bar Inspection

Commonly Used Vehicle-Based Systems

- Track Geometry Systems
- Joint Bar Integrity Vision System
- Ballast Integrity Vision System
- Gage Restraint Assessment
- Subgrade Monitoring (GPR)
- Rail Profile Measurement
- Rail Flaw Detection
- Vehicle/Track Interaction Monitoring
- Rail Surface Condition Assessment
- Coefficient of Friction Monitoring

- Track Component Visual Assessment (fasteners, ties, tie plates, etc)
- Clearance Measurement
- Aerial Survey and Imaging
- Right-of-Way Visual Inspection
- Lubrication System Monitoring
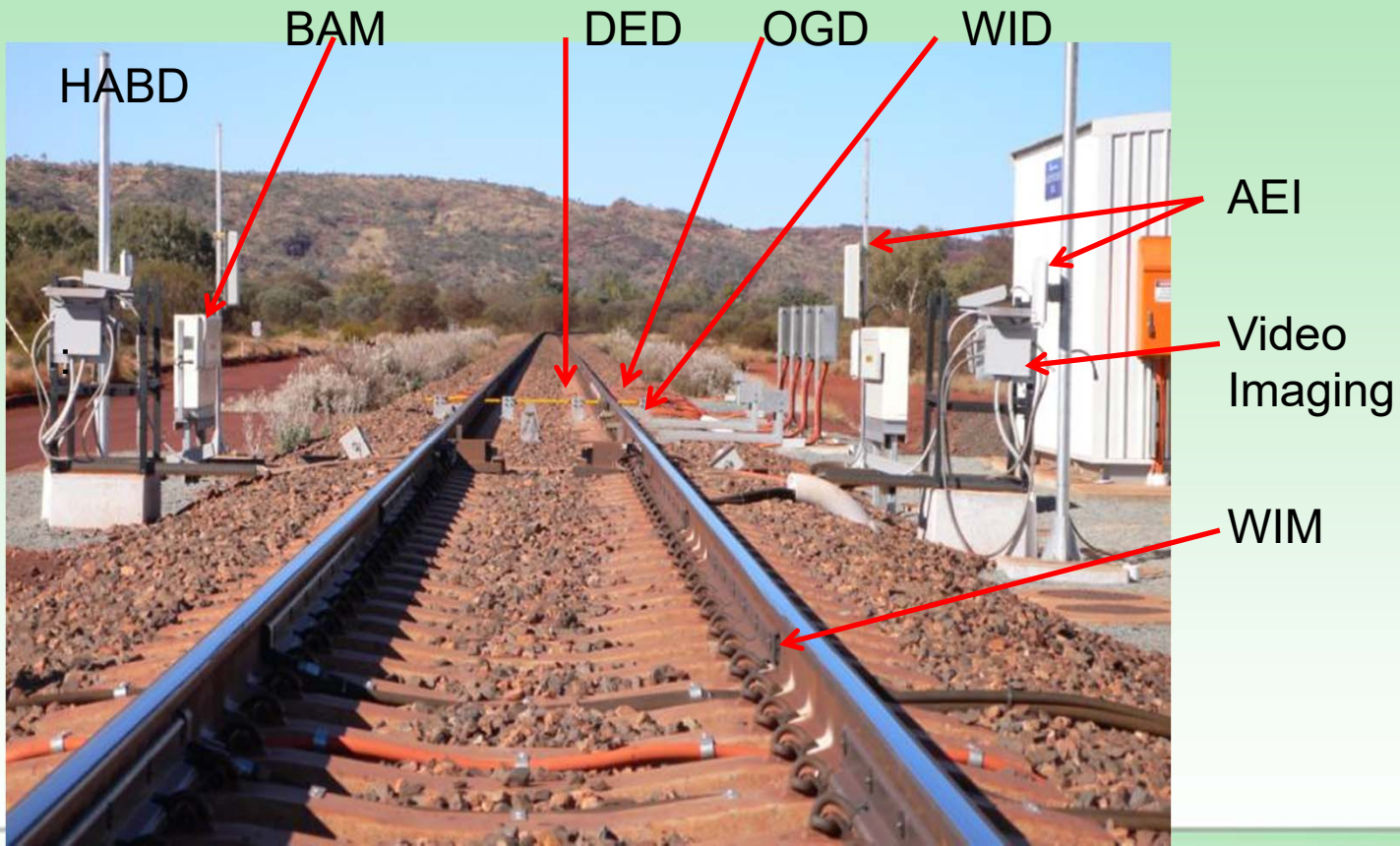- Ballast/Subgrade Movement Detector

What Do Wayside Systems Do?

- Wayside systems are traditionally used to gather data about vehicles and vehicle components:
  - Trucks
  - Axles
  - Wheelsets/Wheels
  - Undercarriage components
  - Brakes
- They are typically installed inside the track bed, along the wayside or under/alongside the rails

Integrated Sites

What Types of Data Are Produced?

- Raw signal/sensor data: accelerations, vibrations, images, applied signal responsivity (acoustic/ultrasonic, laser/light, LIDAR, x-ray, etc.)...

- Filtered/processed signal data

- Combined/integrated data from multiple sensors and modalities

- Exception/anomaly data

How Big is Big

- **Class 1 Railway: "Generating 100 Gb of Data Per day just from one on-board measurement program (9-systems)"**

- **100Gb per day ≈ 25 Tb per year per vehicle**

- **Facebook photos uploaded per day ≈ 350 million**

- **Railroad with 10 vehicles ≈ 350 million pictures**
  (4 cameras @ 30 frames per second)

- **Now add other technologies, measurements, & information.**

Deriving Informative Data from Raw Data

- Sensor data is typically filtered or processed to tease out the relevant signal

- Once the noise is filtered, that signal may be:

  – Analyzed to identify signatures representative of an issue, exception, or anomaly

  – Fed into additional processing layers and/or combined with filtered/processed data from other sensors
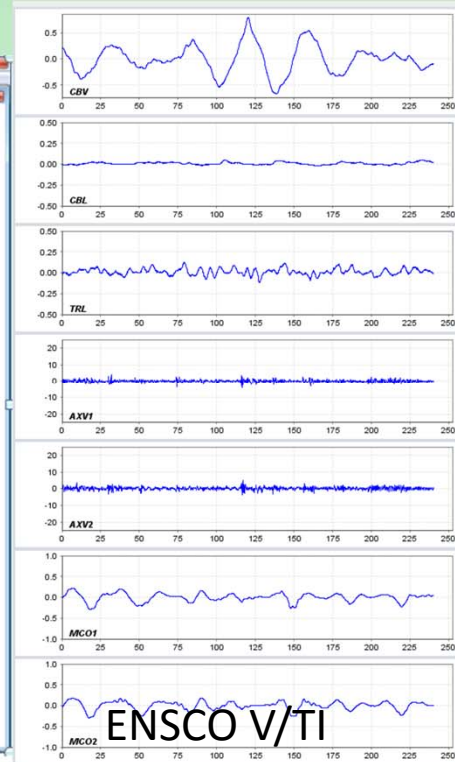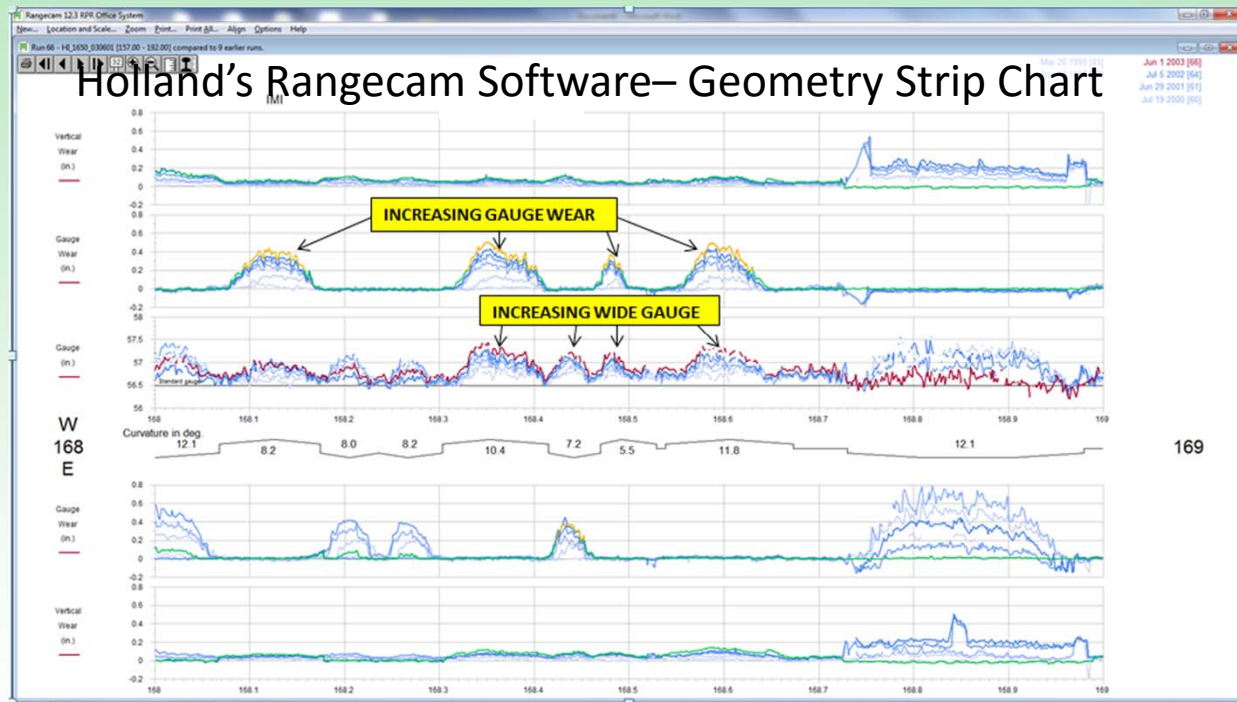
Deriving Informative Data from Data

- This process continues until an informative level of data is achieved (i.e., geometry foot-by-foot data processed for EXCEPTIONS)

- We can also "chain" or combine data streams and first-order exceptions to create new types of exceptions
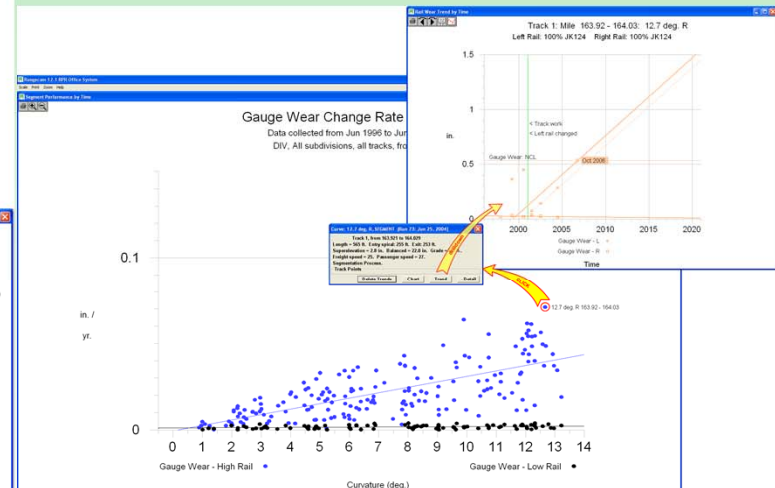
What Does Data Look Like?

Holland's Rangecam Software– Geometry Strip Chart

ENSCO V/TI

What Does Data Look Like?

Sperry Ultrasonic Inspection

ENSCO Joint Bar and Tie Inspection

PRINCIPLES COURSE • JUNE 6, 2017 Holland INTERNATIONAL ENGINEERING TECHNOLOGIES & GLOBAL CONNECTIONS WRI 2017
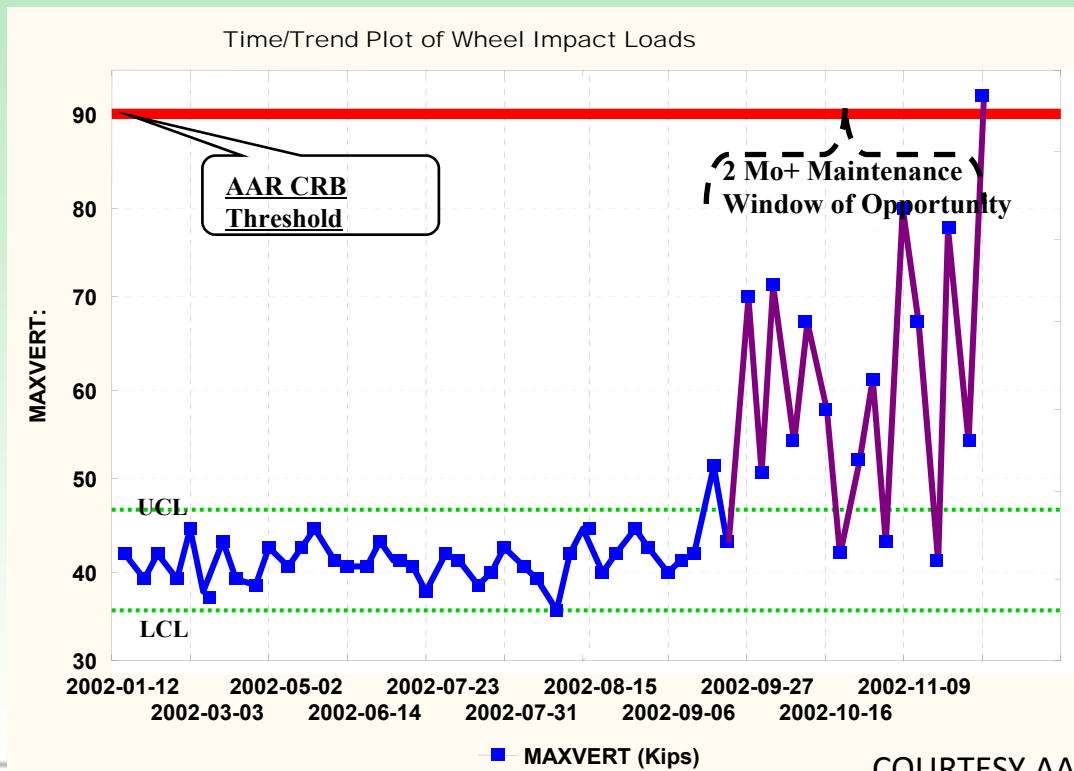
What Does Data Look Like?
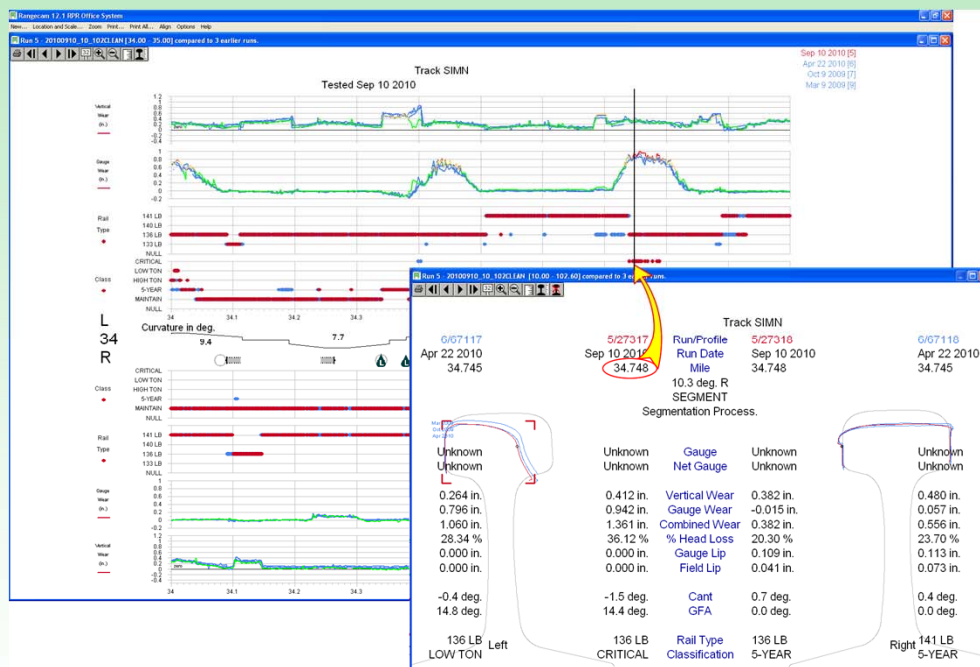
Holland's Rangecam Software with Rail Profile and Wear Analysis

WILD Trend Analysis for Predictive Maintenance

Time/Trend Plot of Wheel Impact Loads

AAR CRB Threshold

2 Mo+ Maintenance Window of Opportunity

UCL

LCL

MAXVERT:

MAXVERT (Kips)

COURTESY AAR

On-Site Data Handling and Storage

- Data is typically stored on-site (mobile platforms or wayside locations) on high-capacity servers

- Different systems store varying degrees of "raw" data: sensor, filtered, processed, exceptions, etc.

- Data can be stored in a variety of formats: flat files, rudimentary to sophisticated databases, video/images and associated metadata or file reference to database, etc.

On-Site Data Handling and Storage

- Many testing platforms and locations host systems from multiple vendors.

- Often, vendors will collaborate to synchronize their data streams and populate a common database.

On-Site Data Handling and Storage

- Some systems do not store data locally at all
- Unattended, Unmanned, Autonomous systems transmit data in real time directly from the field to a central location:
  - Some store data locally until such data is purged
  - Some store data locally temporarily, until wireless transmission is confirmed

Unattended/Unmanned/Autonomous

- Once data is transmitted, it is often further processed and/or validated before being acted upon

- Remote transmission of data leads to big questions about timing and data validity

Liability?

What happens when we experience data loss/overload/black hole?

    Railroad Shutdown?

What if we have the information but we haven't reviewed and/or acted on it?

Are we liable?

    Unwritten/ Unspoken understandings?

    Will those agreements change?

How Do We Reference Data from any Given Platform?

- Systems gather information about vehicle and wayside assets

- We need a way of knowing exactly WHAT/WHERE the data is describing

- We also need a way to reference multiple data streams to each other

How Do We Reference Data from a Mobile Platform?

- Systems residing on the same host vehicle can be time-referenced to each other, although processing delays can contribute to errors

- On-board systems typically rely on GPS and/or track-based referencing techniques to synchronize and report data

- Location information can be embedded at one or multiple points along the processing chain

Decoding the Decentralized Nature of Data Collection

- With thousands of vehicles collecting data all over their networks, railroads are faced with the issue of how to integrate or associate all of these disparate data streams
- This typically happens at a centralized location– "The Office"
- Requires data streams to have:
  - A common method of referencing the data (GPS, track location or asset, etc)
  - A known or anticipated margin of error for the accuracy of the location reference

What Happens to Data (Part 1)?

- ## Some data is immediately actionable
  - Ex.: Rail flaw detectors routinely stop and hand verify detected exceptions

- ## Some data is near-immediately actionable
  - Ex.: Geo cars radio back exceptions, hand off printed exception reports/strip charts to field personnel– during or at the end of a test

What Happens to Data (Part 2)?

- Generally, most data is transmitted back to a central repository for further action and storage after collection

- Data transmission happens in a variety of ways:
  - Physical– disk transfers via snail mail from vehicle to office
  - Electronic– email, upload to server/FTP

What Happens to Data (Part 2)?

- The central repositories are architected to host all different types of data sources, and accessibility provisions must be made

- High level users access data for follow up, audits, more detailed evaluation

- Most data management systems include provisions for "close the loop"

What Happens to Data (Part 2)?

- Because many modalities of data are hosted in these repositories, we start talking BIG DATA

- When combined with data from and about other assets, we move well beyond "find and fix" to predictive modeling and planning

The First Degree of Data

- A single survey yields data that is **immediately actionable**:
  - Rail Profile → Grinding
  - Rail Flaw → flaw removal
  - Imaging → Replace broken joint bars, bad ties, missing fasteners, etc.

The Second Degree of Data

- A second survey yields **HISTORY**, upon which a **FORECAST** can be built:
  - Rail Profile → Rail replacement planning
  - Rail Flaw → Internal defect growth
  - Imaging → tie replacement planning

The Third Degree of Data

- Multiple modalities of data at the same location yield **PREDICTIVE ANALYTICS** and provide the **WHY**:
    - Rail Profile + Rail Flaw + Rail Stress → predicting rail break
    - Rail Profile + Geometry → wide gage due to rail wear
    - Impact Monitoring + Visual Inspection → predicting rapid deterioration of joints

The Fourth Degree of Data

Lots of data about a specific Location or Asset

+

Lots of data about Vehicle/Mobile asset data

=

WHAT IF

The Fourth Degree of Data

- Using sophisticated modeling techniques and softwares, we can SIMULATE the **behavior of our track** as it responds to **different loading conditions**

- And vice versa– we can SIMULATE the **response of a particular vehicle/load** to a **specific track condition**

The Fourth Degree of Data…

- **Predicting responses** of our track and vehicles to each other based on ACTUAL or anticipated conditions means:
  - More effective maintenance standards and practices
  - More informed decisions about condemning limits
  - More efficient operational practices

...And Beyond

- The first four degrees of data optimize our operations
- By integrating our track and vehicle data with ERM/ERP's, we can optimize our BUSINESSES

Thanks

- Tom Toth and Pat Howser, Union Pacific
- Leo Kreisel, CSX
- Rob DiMatteo, Sperry Rail Service
- Bob Tuzik, Holland LP
- Matt Dick, ENSCO