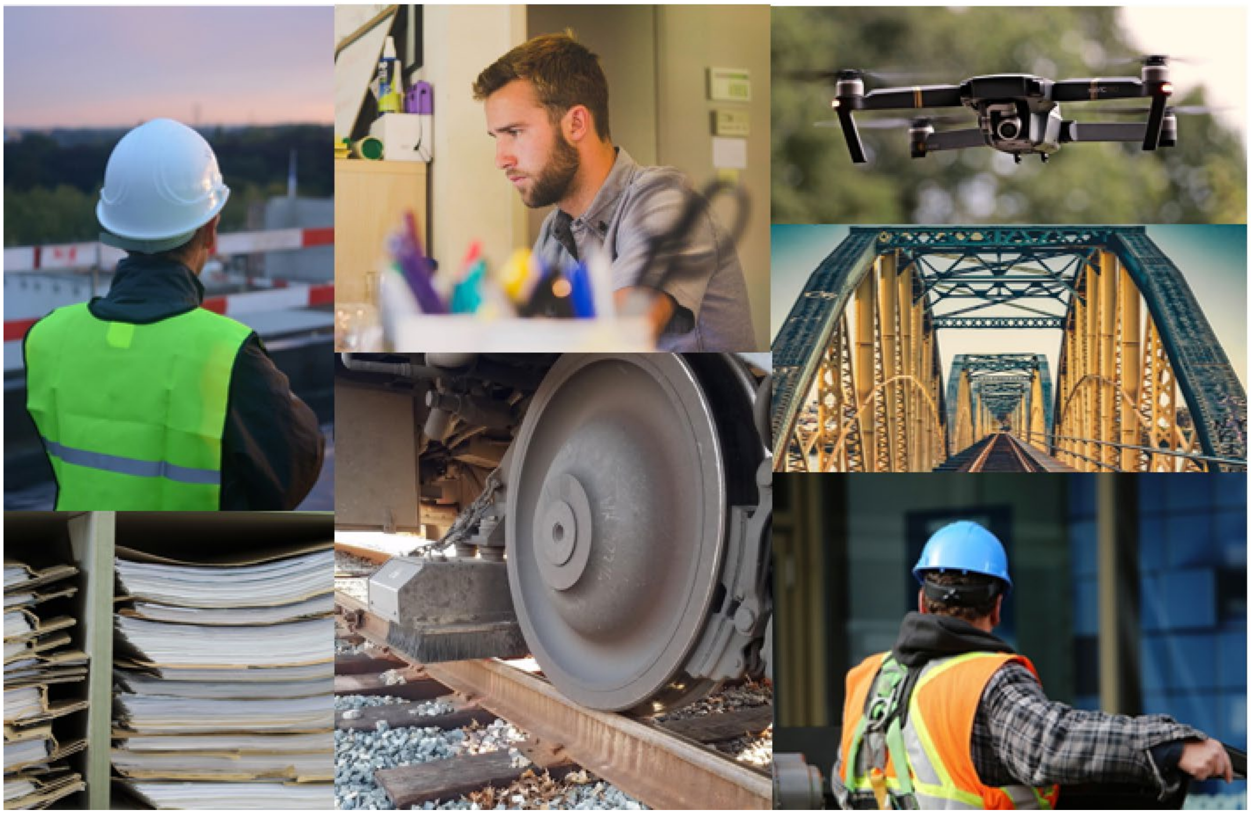
U.S. Department
of Transportation
Federal Railroad
Administration

Office of Research,
Development and Technology
Washington, DC 20590

Human-Automation Teaming in Track Inspection



DOT/FRA/ORD-23/07

Final Report | March 2023

NOTICE

NOTICE

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)* 08/06/2021	**2. REPORT TYPE** Technical Report	**3. DATES COVERED** *(From - To)* *04/01/2020 – 08/06/2021*

4. TITLE AND SUBTITLE Human-Automation Teaming in Track Inspection	**5a. CONTRACT NUMBER** **5b. GRANT NUMBER** **5c. PROGRAM ELEMENT NUMBER**
6. AUTHOR(S) Megan France-Peterson (0000-0001-7224-3982) Jordan Multer (0000-0002-7818-110X) Gina Melnik (0000-0001-6006-5189)	**5d. PROJECT NUMBER** **RR04AB** **5e. TASK NUMBER** **UK299** **5f. WORK UNIT NUMBER**
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Volpe National Transportation Systems Center 55 Broadway Cambridge, MA. 02142	**8. PERFORMING ORGANIZATION REPORT NUMBER** DOT-VNTSC-FRA-21-06
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Department of Transportation Federal Railroad Administration Office of Railroad Policy and Development Office of Research, Development, and Technology Washington, DC 20590	**10. SPONSOR/MONITOR'S ACRONYM(S)** **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** **DOT/FRA/ORD-23/07**

12. DISTRIBUTION/AVAILABILITY STATEMENT
This document is available to the public through the FRA website.

13. SUPPLEMENTARY NOTES
COR: Michael Jones

14. ABSTRACT
This report provides a foundational overview of how the rail industry can approach the design of human-automation teams (HATs), with the goal of improving safety and efficiency in the track inspection process. The authors present a general process for designing HATs, then explore how this design process can be applied to track inspection. The report addresses four track inspection tasks (data collection, data analysis, decision making, and action) and presents broadly applicable considerations for HATs for each task. Railroads and manufacturers may use the design process and general considerations in this report to develop requirements specific to a given inspection technology.

15. SUBJECT TERMS
Instrumentation, human-instrumentation teaming, HAT, human-centered design, human factors, track inspection, track instrumentation

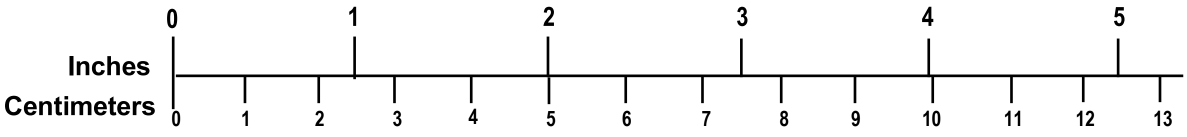
16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**		97	Megan France-Peterson
					19b. TELEPHONE NUMBER *(Include area code)* (617) 494-3674

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

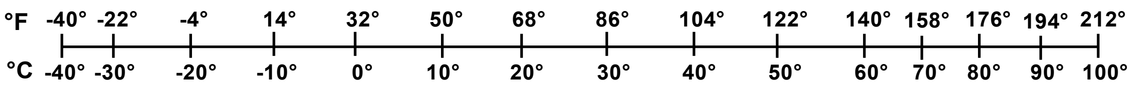
METRIC/ENGLISH CONVERSION FACTORS

ENGLISH TO METRIC			METRIC TO ENGLISH		
LENGTH (APPROXIMATE)			**LENGTH** (APPROXIMATE)		
1 inch (in)	=	2.5 centimeters (cm)	1 millimeter (mm)	=	0.04 inch (in)
1 foot (ft)	=	30 centimeters (cm)	1 centimeter (cm)	=	0.4 inch (in)
1 yard (yd)	=	0.9 meter (m)	1 meter (m)	=	3.3 feet (ft)
1 mile (mi)	=	1.6 kilometers (km)	1 meter (m)	=	1.1 yards (yd)
			1 kilometer (km)	=	0.6 mile (mi)
AREA (APPROXIMATE)			**AREA** (APPROXIMATE)		
1 square inch (sq in, in^2)	=	6.5 square centimeters (cm^2)	1 square centimeter (cm^2)	=	0.16 square inch (sq in, in^2)
1 square foot (sq ft, ft^2)	=	0.09 square meter (m^2)	1 square meter (m^2)	=	1.2 square yards (sq yd, yd^2)
1 square yard (sq yd, yd^2)	=	0.8 square meter (m^2)	1 square kilometer (km^2)	=	0.4 square mile (sq mi, mi^2)
1 square mile (sq mi, mi^2)	=	2.6 square kilometers (km^2)	10,000 square meters (m^2)	=	1 hectare (ha) = 2.5 acres
1 acre = 0.4 hectare (he)	=	4,000 square meters (m^2)			
MASS - WEIGHT (APPROXIMATE)			**MASS - WEIGHT** (APPROXIMATE)		
1 ounce (oz)	=	28 grams (gm)	1 gram (gm)	=	0.036 ounce (oz)
1 pound (lb)	=	0.45 kilogram (kg)	1 kilogram (kg)	=	2.2 pounds (lb)
1 short ton = 2,000 pounds (lb)	=	0.9 tonne (t)	1 tonne (t)	=	1,000 kilograms (kg)
				=	1.1 short tons
VOLUME (APPROXIMATE)			**VOLUME** (APPROXIMATE)		
1 teaspoon (tsp)	=	5 milliliters (ml)	1 milliliter (ml)	=	0.03 fluid ounce (fl oz)
1 tablespoon (tbsp)	=	15 milliliters (ml)	1 liter (l)	=	2.1 pints (pt)
1 fluid ounce (fl oz)	=	30 milliliters (ml)	1 liter (l)	=	1.06 quarts (qt)
1 cup (c)	=	0.24 liter (l)	1 liter (l)	=	0.26 gallon (gal)
1 pint (pt)	=	0.47 liter (l)			
1 quart (qt)	=	0.96 liter (l)			
1 gallon (gal)	=	3.8 liters (l)			
1 cubic foot (cu ft, ft^3)	=	0.03 cubic meter (m^3)	1 cubic meter (m^3)	=	36 cubic feet (cu ft, ft^3)
1 cubic yard (cu yd, yd^3)	=	0.76 cubic meter (m^3)	1 cubic meter (m^3)	=	1.3 cubic yards (cu yd, yd^3)
TEMPERATURE (EXACT)			**TEMPERATURE** (EXACT)		
[(x-32)(5/9)] °F	=	y °C	[(9/5) y + 32] °C	=	x °F

QUICK INCH - CENTIMETER LENGTH CONVERSION



QUICK FAHRENHEIT - CELSIUS TEMPERATURE CONVERSIO



For more exact and or other conversion factors, see NIST Miscellaneous Publication 286, Units of Weights and Measures. Price $2.50 SD Catalog No. C13 10286

Updated 6/17/98

Acknowledgements

The Volpe National Transportation Systems Center's research team would like to thank the subject matter experts who contributed to this effort by sharing their knowledge, reviewing sections of this report, or suggesting relevant literature and new ideas to the research team.

Contents

Executive Summary ... 1

1. Introduction ... 3
  1.1 Background ... 3
  1.2 Objectives ... 7
  1.3 Overall Approach ... 7
  1.4 Scope ... 7
  1.5 Organization of the Report ... 8

2. Background on Automation and Track Inspection ... 10
  2.1 Desired Impacts of Automation on Track Inspection ... 10
  2.2 Supporting Strategic Maintenance Decisions ... 11
  2.3 Augmenting Human Capabilities ... 14

3. Evolving Approaches to Function Allocation and Designing HATs ... 15
  3.1 Capability-based Approaches to Function Allocation ... 15
  3.2 Levels of Automation (LOA) Models ... 15
  3.3 HAT Approaches to Function Allocation ... 16
  3.4 Summary of Function Allocation Approaches ... 16

4. General Process for Designing Human-Automation Teams ... 18
  4.1 Step 1: Examining Process Requirements and System-Level Goals ... 19
  4.2 Step 2: Exploring Possible Roles for Humans and Automation ... 20
  4.3 Step 3: Considering Tradeoffs and Challenges ... 22
  4.4 Step 4: Developing Detailed Design Requirements ... 22
  4.5 Step 5: Implementing and Assessing HATs ... 23

5. HAT in the Track Inspection Process ... 25
  5.1 Step 1: Examining Track Inspection Process Requirements and System Goals ... 25
  5.2 Step 2: Exploring Roles for Humans and Automation in Track Inspection ... 27
  5.3 Step 3: Considering Tradeoffs and Challenges in Track Inspection HATs ... 30

6. Data Collection: Observe or Measure Track Condition ... 31
  6.1 Understanding the Data Collection Task ... 31
  6.2 Roles of Humans and Automation in Data Collection ... 32
  6.3 Considerations for Data Collection ... 37

7. Data Analysis: Analyze Track Condition and Identify Problems ... 41
  7.1 Understanding the Data Analysis Task ... 41
  7.2 Roles of Humans and Automation in Data Analysis ... 42
  7.3 Considerations for Data Analysis and Verification ... 47

8. Decision Making and Action ... 53
  8.1 Understanding Decision Making and Action Tasks ... 53
  8.2 Roles of Humans and Automation in Decision Making and Action ... 54
  8.3 Considerations for Decision Making and Action ... 59

9. Overall Considerations for HAT ... 61

9.1 Importance of Interdependence in Track Inspection Tasks 61
9.2 Design Considerations Across Track Inspection Tasks 62
9.3 Overall Considerations Associated with Increased Use of Automation 63

10. Conclusion 67

11. References 69

Appendix A. Glossary of Terms 72

Appendix B. Expanded Discussion of the General Process for Designing HATs 75

Appendix C. Questions to Guide Design 83

Abbreviations and Acronyms 86

Illustrations

Figure 1. Importance of identifying human-automation interaction issues early, adapted from Fleming (Fleming, 2015) 6

Figure 2. Potential impacts of automation on safety, maintenance requirements, and revenue service capacity 10

Figure 3. Relationship between track allocation for inspection and revenue service 11

Figure 4. Dual purpose of track inspection 12

Figure 5. “Find and fix” railroad maintenance cycle 13

Figure 6. Evolution of function allocation approaches 17

Figure 7. General process for designing HATs 19

Figure 8. Humans and technologies involved in track inspection 27

Figure 9. LOA can vary by inspection task 28

Figure 10. Iterative nature of the track inspection process, including repeated data collection activities for verification 30

Figure 11. Examples of teaming structures for data collection 35

Figure 12. Examples of teaming structures for data analysis 45

Figure 13. Approaches for integrating data and delivering it to field staff 49

Figure 14. Examples of teaming structures for decision making and action 57

Figure 15. Interdependence in track inspection 61

Tables

Table 1. Comparison of Two Approaches to System Design..6
Table 2. Summary of Track Inspection Process Tasks and Activities...26
Table 3. Example LOAs for Data Collection 33
Table 4. Example LOAs for Data Analysis..44
Table 5. Example LOAs for Decision Making...55
Table 6. Example LOAs for Action..56
Table 7. Relevant Design Considerations for Teaming Structures by Inspection Task 62
Table 8. Capability Analysis Categories, adapted from Johnson et al. (2020)............................78
Table 9. Characteristics Associated with Evolution of Automation, adapted from Billings (1997) 79
Table 10. Criteria for Evaluating HATs, adapted from Pritchett et al. (2014).............................81

Executive Summary

This report provides a foundational overview of how the rail industry can approach the design of human-automation teams (HATs) with the goal of improving safety and efficiency in the track inspection process. Through funding from the Federal Railroad Administration (FRA), a team from the Volpe National Transportation Systems Center conducted research for this study from April 2020 to August 2021.

There is no single correct way to structure HATs. Rather, it is important to understand the benefits and challenges associated with each possible design choice and take a deliberate approach to designing the right HAT structures for the desired context. Railroads, technology manufacturers, and track inspection researchers may use the information contained in this report to further their understanding of human-automation interactions, and to support work related to the design and use of specific track inspection technologies.

This report begins by providing background on automation in track inspection. The railroad industry seeks to use existing and emerging automated technologies to improve the efficiency and effectiveness of the track inspection process. The use of technology in track inspection can provide numerous benefits, including augmenting human detection of degraded track conditions and defects, speeding up track inspection, and, in some cases, inspecting track more frequently than human inspectors. In addition to improving railroads' ability to detect and address defects in the short term, automation may also facilitate more effective long-term planning for proactive maintenance activities. However, the use of automation introduces new challenges alongside these benefits.

Many technology developers use a technology-centered approach (i.e., automating as much as possible based on the capabilities of the technology and relying on humans to manage any limitations of the technology). This can introduce numerous problems, such as communication and coordination challenges, humans being required to intervene unexpectedly without adequate situation awareness, errors due to poor interface design, and deskilling of human operators.

In contrast, a human-centered approach to HAT design can mitigate many of the challenges associated with the use of automation. A human-centered approach means carefully considering how tasks and responsibilities should be allocated between humans and automation, taking into consideration the roles, abilities, and limitations of each, as well as the possible interactions between them.

Over the years, human-centered design approaches have evolved to better evaluate the range of possible human-automation interactions and implications of different design choices. Early approaches focused on assigning tasks to humans or machines in a binary fashion based on relative capabilities, while later, more sophisticated approaches recognize that humans and automation may share or trade off task performance and designers may then model those interactions. The most recent approaches to designing HATs involve understanding how humans and automation operate as a joint system or team and seeking to optimize the performance of that team.

This report presents a general process for designing HATs, which includes five steps:

1. Examining process requirements and system-level goals

2. Exploring possible roles for humans and automation
3. Considering tradeoffs and challenges
4. Developing detailed design requirements
5. Implementing and assessing human-automation teams

The report then discusses how this process can be applied to the design and use of track inspection automation.

This report uses a four-stage information processing model to describe the tasks involved in the track inspection process:

1. Data collection
2. Data analysis
3. Decision making
4. Action

For each of these four tasks in the inspection process, this report explores how designers can apply the general process for designing HATs. For each task in the process, this report examines the complexities of the task, identifies possible roles for humans and automation, discusses levels of automation (LOA) and possible HAT structures, and identifies important design considerations and tradeoffs (i.e., steps 1 through 3 of the general process for designing HATs).

Across track inspection tasks, some important areas to address in HAT design include technology and interface design, training and experience, communication and coordination between team members, and supporting both short-term and long-term goals.

This information may serve as a foundation for audiences looking to explore a HAT approach to the design and use of specific track inspection technologies. Railroads, technology manufacturers, and track inspection researchers may use the information contained in this report to adopt a more human-centered approach to the use of automation in track inspection.

1. Introduction

The way that railroads and manufacturers approach the design and integration of new track inspection technologies informs how those technologies will function as part of the track inspection process. Design choices that consider the humans and technologies operating *together* can mitigate negative impacts and support safety by improving the performance of the joint human-machine system. This report discusses human-automation teaming (HAT), which the authors use to describe humans and technology working together toward the same goals.[1]

This research examines the ways HATs can work together in the track inspection process, including focus on both current and future uses of automation. The authors address information gaps and examine challenges associated with various types of human-automation teams (HATs).

This report may be useful for a variety of stakeholders: railroads, technology manufacturers, academics, and the Federal Railroad Administration (FRA). Many of these stakeholder groups collaborate with one another to develop and test emerging technologies. The information contained in this report can help stakeholders adopt a more human-centered approach to the design and use of automation in track inspection.

A research team from the Volpe National Transportation Systems Center conducted this FRA-funded research between April 2020 and August 2021 and consulted with relevant subject matter experts to explore how a human-centered approach to design can mitigate many of the challenges associated with the use of automation and increase system safety and efficiency.

1.1 Background

A wide range of technologies currently exists to support the track inspection process. These include hi-rail vehicles commonly used to traverse tracks and perform visual inspections more quickly than on foot as well as more advanced technologies like track geometry measurement systems (TGMS) that use sensors and computers to take measurements and identify conditions such as wide gage. These more advanced technologies can be referred to as “automated,” because they automate the tasks of identifying track anomalies and evaluating whether they exceed thresholds indicating the need for repair. However, there is a wide range of capabilities among automated inspection technologies related to which types of track conditions they are designed to detect and the method used to detect those conditions (i.e., simple measurement or pattern recognition). As a result of these differing capabilities, there is variability in how much human involvement is required in the operation of automated inspection technologies.

Including such technologies in the track inspection process is generally aimed at supporting safety and efficiency. Automated inspection technology can support track safety by capturing conditions that cannot be easily detected by humans (e.g., ultrasonic detection of rail flaws). Automated inspection technologies also provide faster and more consistent measurements of entire sections of track (i.e., more data than can be collected manually) which allows railroads to

---

[1] Some researchers use the term “Human-Automation Teaming” or “HAT” to refer to interactions between humans and automation with specific characteristics, such as transparency (i.e., the ability for the human to understand what the automation is doing or suggesting), bi-directional communications, and human-directed actions (Battiste et al., 2018). This report uses the term HAT more broadly and does not require these specific teaming characteristics, though they may be desirable for some HATs.

identify emerging safety issues by monitoring changes over time. Augmenting human inspection with automated inspection technologies can also improve the efficiency of track inspection and maintenance. The use of automated inspection technology may allow railroads to make more cost-effective maintenance decisions because they have more data about track degradation over time. This data may also enable inspectors and maintenance crews to spend their time more efficiently by reducing the track time required for inspection and maintenance activities. In some cases, if permitted by FRA waiver, the use of such technologies can increase efficiency by allowing railroads to perform fewer visual inspections (i.e., inspections by a human alone) on the territory where they operate. These efficiency benefits may be especially notable if the technology does not require dedicated track time to operate.

However, there are also potential challenges to using automation in track inspection, particularly in relation to the interaction of humans and automation. One significant challenge is developing technologies which do not overlook degraded track conditions, but also do not have an excessive rate of false alarms. Currently, railroads send inspectors to verify conditions identified by automated systems, which could result in additional workload and inefficiencies if the false alarm rate is high. And because most automated technologies only detect a limited set of conditions, railroads must rely on information from multiple technologies and visual inspections and then synthesize that information to understand the state of their infrastructure.

Automation is not a replacement for human work. Humans and automation can share or trade off in the performance of tasks, allowing the automation to augment the capabilities of humans, relieve them of excessive workload, act as a backup to human actions, or replace some subset of the tasks a human normally performs (Sheridan & Verplanck, 1978). New automated technologies may change the way that track inspections take place, including changing the roles of the humans currently involved in track inspection (e.g., using data from automated inspections to target “problem areas” during visual inspection) or creating additional roles (e.g., programming and maintaining automated technology). In the human factors’ literature, “levels of automation” (LOA) models are used to describe the range of uses for automation and the roles that humans and computers can perform. Different LOAs have associated strengths and weaknesses related to the relative capabilities of humans and automation and the types of interactions between them.

The term “autonomous” is sometimes used when discussing automated track inspection technologies, which a layperson may interpret as the automation operating entirely on its own. However, most current track inspection technologies are partially automated, *not* fully autonomous, and rely on humans to perform tasks like setting their course, monitoring their location, or reviewing data outputs for accuracy. Even a technology that could operate autonomously during inspections would be programmed by humans, have output that is used by humans, and need to be maintained by humans. Therefore, human-automation interactions exist even for systems that railroads may view as autonomous.

Understanding how differing LOAs affect the track inspection process will help railroads and manufacturers make design decisions that reflect the role humans and technology will play as railroad systems evolve.

The current research takes the perspective that a human-centered approach, which carefully considers the design of the new track inspection automation as part of a team, will contribute to better safety and efficiency outcomes. The goal was to explore and provide guidance on how different design choices around automating track inspection activities may affect system safety and efficiency.

1.1.1 Moving Away from Technology-Centered Design Towards Human-Centered Design

This section describes the difference between current, technology-centered approaches to system design and human-centered or systems perspectives. Human-centered design can help railroads maximize the safety and efficiency of the track inspection process by using HATs for track inspection.

In the United States, the rail industry and technology developers have focused on how track inspection technologies can improve track safety and increase track availability for revenue service (i.e., by reducing track time required for inspection and maintenance). They have taken a technology-centered approach to achieve these goals. A technology-centered approach focuses on meeting technical hardware and software requirements needed for the system to complete tasks and does not address human system requirements. In a technology-centered approach, designers may not fully consider how the technology will fit into existing inspection processes and the overall context or environment where it will be used. For example, technology may be designed without consideration for the capabilities and limitations of the employees that will interact with the system, such as technology operators or inspectors that may use the data the technology produces; these employees must adapt to the the technology. This technology-centered approach is common in many industries and its limitations are well-documented (Hollnagel, 2001). Considering the role of the human *after* the automation has been incorporated may lead to failure modes that designers might have otherwise eliminated or mitigated and can cause suboptimal system performance (Ehrlich & Rohn, 1994; Grimes, Wright, & Hillier, 2009; Melnik, Roth, Multer, Safar, & Isaacs, 2018).

The technology-centered approach to designing track inspection technologies can be contrasted with a human-centered approach. In a human-centered approach, the human's role, abilities, and limitations are considered during the system design. The technology is designed to meet human requirements to complete tasks and match human mental and physical capabilities to operate the system.

Considering the role of the human and the goals of track inspection and maintenance may lead to different ways of incorporating automation, and support both more effective use of automation and safer systems (Bainbridge, 1983; Billings, 1997). Taking an approach in which the human is considered an integral part of the system can mitigate potential failure modes and increase safety (Dekker, 2011; Leveson, 2004).

Table 1 highlights differences between these two approaches.

Table 1. Comparison of Two Approaches to System Design

Technology-Centered	Human-Centered
• Technology comes first; the goals is to automate as much as possible • Tasks that automation can't do are allocated to humans • Human may be left with an illogical set of tasks • User interaction issues are often caught too late	• Human needs and capabilities in interacting with the technology are considered first • Technology is designed with optimal performance of the joint human-machine system in mind • User interaction issues are identified early & addressed iteratively

Addressing problems caused by the failure to consider the interactions between humans and technology tends to be more costly and problematic than considering the human's role during the initial design. The following section describes how to integrate human-centered perspectives early in the design process.

1.1.2 Integrate Human-Centered Design Early for Best HAT Outcomes

The best time to consider HAT is early in the design process (Fleming, 2015). Figure 1 illustrates how the cost and difficulty of addressing human-automation interaction issues increases at later stages of technology development.

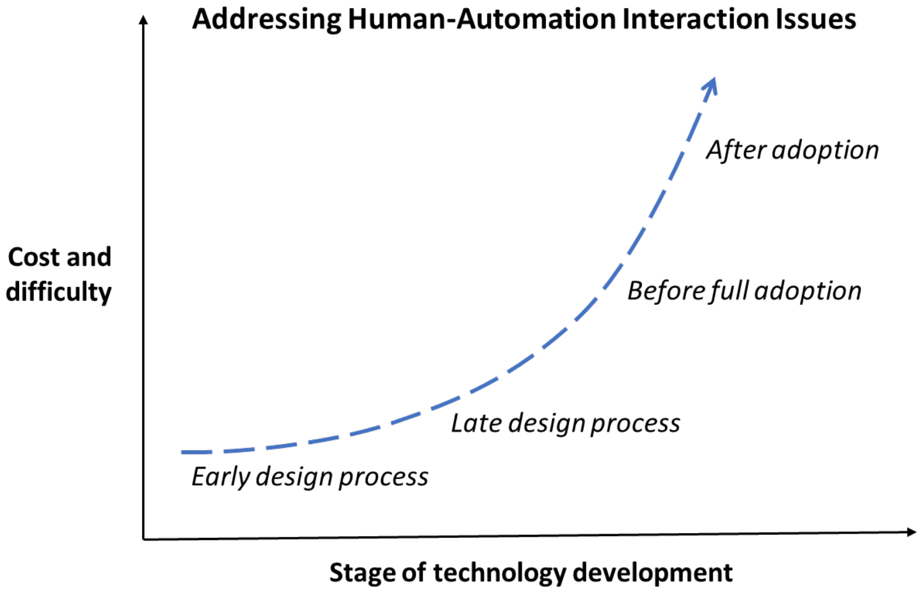


Figure 1. Importance of identifying human-automation interaction issues early, adapted from Fleming (Fleming, 2015)

Late in development, redesigns can be costly and inefficient, and once a technology has been adopted, options are much more limited. For example, if a poorly designed interface makes users prone to data entry errors, training may partly address but not eliminate the problem. To avoid this outcome, human-automation interactions should be examined as early as possible as part of overall systems engineering processes. However, if they have not been included in design, it is not too late to assess existing processes and implement improvements to make HATs more effective.

1.2 Objectives

This research aims to enable the railroad industry to approach the design of HATs from a human-centered perspective to improve safety and efficiency in the track inspection process.

There are multiple ways to approach assigning tasks to humans and automation. Traditional approaches do not always lead to the most effective use of human and machine capabilities. A human-centered HAT approach yields more effective design decisions.

This work aims to provide value for multiple audiences, including railroads, technology manufacturers, and FRA, by outlining a general process for incorporating human-centered perspectives in the design of HATs and discussing how this process can be applied to the design and use of track inspection automation. The report presents general considerations that apply to a wide range of inspection technology applications, including both current and future uses of automation.

Technology manufacturers and railroads may incorporate this information when designing or procuring new technologies to ensure that the technologies are well-suited to effective HAT. Railroads may use this information to inform how they integrate new and existing technologies into their track inspection processes. FRA may use this information to inform future research.

1.3 Overall Approach

First, the Volpe team sought to identify best practices for HAT, as well as how HAT is currently practiced in track inspection. The team explored how humans (e.g., track inspectors, supervisors, and technology operators) and automation currently interact in the track inspection process, and how they may interact as automation evolves.

The team reviewed existing research on HAT and human-centered design. The team also reviewed research literature on track inspection and automation and drew on the team's prior work on hazards associated with track inspection (France, Melnik, Safar, & Multer, 2020), which incorporated perspectives from three passenger railroads. The team conducted interviews with subject matter experts about how railroads use automated track inspection technologies. Interview discussions also addressed current HAT structures used in the track inspection process and challenges associated that may be addressed through effective design of HATs.

The team adapted existing research on HAT and human-centered design into a 5-step process for designing HATs. Then the team applied the first three steps to a track inspection context.

The team also identified HAT considerations applicable to track inspection tasks. The analysis considered the range of possible human-automation teams that could exist in each inspection process task, and human factors challenges that could apply to those teaming structures.

1.4 Scope

This work focused on human factors issues rather than technical specifications for automation. It discusses the design of automation when it is relevant to the operators of that automation (i.e., design of control and display interfaces) and discusses decisions related to functions the automation should perform. Ensuring that the capabilities of the automation are sufficient to perform those functions is left to the automation designers and manufacturers.

This work focused on ways that human factors considerations related to HATs may generalize across different technologies or within categories of technologies, rather than examining any specific technology in depth.

This report discusses automation uses that are not permitted under current regulation (e.g., those that do not meet current requirements for inspectors traversing the track, such as human-automation teams with a remote inspector). The report relates the breadth of possibilities for HAT (and associated design considerations) with the understanding that railroads will need to comply with FRA regulations for any new technologies they wish to implement in inspection practices. The purpose of discussing hypothetical automation use in this report is to identify potential impacts on HAT and overall system performance. FRA, railroads, and technology manufacturers may address these considerations if waivers or future regulations allow such uses.

1.5 Organization of the Report

This report is organized into the following sections:

- Section 2 provides background on automation and track inspection.
- Section 3 describes the evolution of function allocation approaches, including discussing three ways to allocate tasks and functions to humans and automation when designing HATs.
- Section 4 describes a 5-step process for designing HATs:
    - Step 1: Examining process requirements and system-level goals
    - Step 2: Exploring possible roles for humans and automation
    - Step 3: Considering tradeoffs and challenges
    - Step 4: Developing detailed design requirements
    - Step 5: Implementing and assessing individual teaming structures
- Section 5 discusses the first 3 steps in designing HATs as they relate to track inspection technologies. (The last two steps are beyond the scope of this report.) This discussion includes understanding how the track inspection process can be broken down into four tasks:
    - Data collection
    - Data analysis
    - Decision making
    - Action

  It then elaborates on how researchers can model relationships between humans and automation using LOA and HAT frameworks.
- Section 6 addresses the data collection task, including possible roles for humans and automation and associated considerations and challenges.
- Section 7 addresses the data analysis task, including possible roles for humans and automation and associated considerations and challenges.

- Section 8 addresses the decision making and tasks, including possible roles for humans and automation and associated considerations and challenges.
- Section 9 discusses relationships between the tasks and summarizes major themes identified in this research.
- Section 10 summarizes the report’s conclusions.
- Appendices:
    - Appendix A. Glossary of Terms
    - Appendix B. Expanded Discussion of the General Process for Designing HATs
    - Appendix C. Questions to Guide Design

2. Background on Automation and Track Inspection

This section provides context for understanding how the design of human-automation interactions affects track inspection safety and efficiency. Foundational background on the track inspection process and the goals of track inspection automation is also included.

2.1 Desired Impacts of Automation on Track Inspection

Railroads, vendors, FRA, and others have supported research to take advantage of evolving technology to improve the track inspection process and track and rolling stock maintenance (Baillargeon, 2020; Saadat & Sherrock, 2018). Railroads have adopted technologies including computer and communications technology, artificial intelligence (AI), modeling, sensors, and other options to improve productivity and safety (Roberts, 2003). This enables railroads to automate track inspection tasks using technologies that augment human detection of degraded track conditions, speed up track inspection, and, in some cases, inspect track more frequently than human inspectors while minimizing the impact on revenue service.

Increasing the use of automation in track inspection could provide many benefits to the railroad industry including:

- Allocating less time and resources to track maintenance
- Increasing track capacity for revenue service
- Maintaining safety at the current level or better

Figure 2 illustrates the desired relationship between these factors and increased automation. However, these represent *potential* benefits of automation; successfully realizing these goals depends on how the railroads implement that automation in practice. Realizing these benefits requires designing effective interactions between humans and automation in the track inspection process.

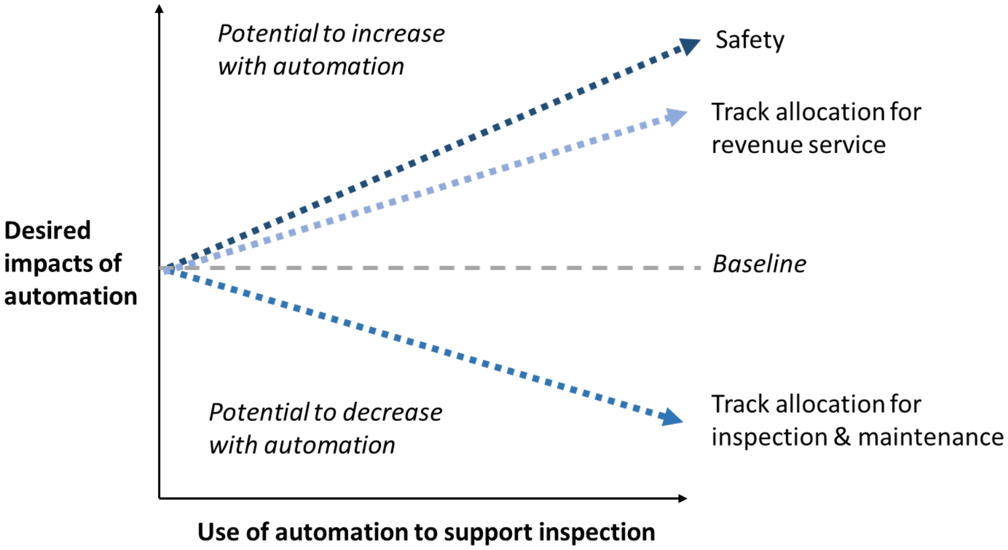


Figure 2. Potential impacts of automation on safety, maintenance requirements, and revenue service capacity

There is an inverse relationship between the track capacity needed for inspection and maintenance and the track capacity available for revenue service. Figure 3 depicts this relationship.

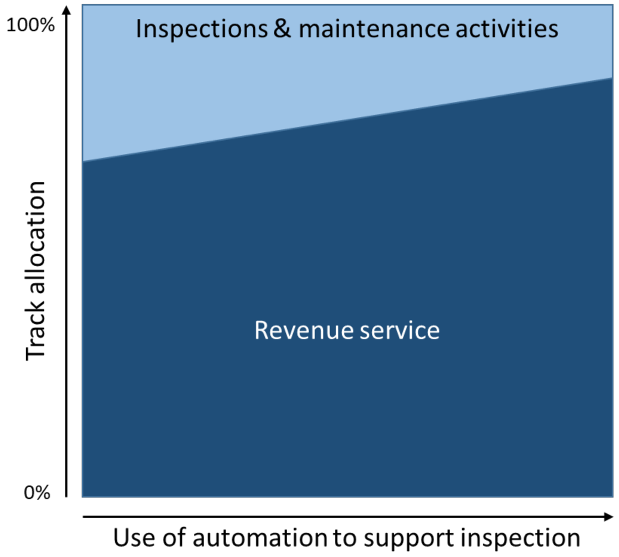


Figure 3. Relationship between track allocation for inspection and revenue service

Automation may support strategic maintenance decisions and augment the capabilities of human track inspectors.

2.2 Supporting Strategic Maintenance Decisions

Increased use of automation could help realize track inspection benefits through supporting strategic maintenance decisions. Track inspections serve two purposes:

1. They allow railroads to identify and address defects and/or track conditions which do not meet FRA’s track safety standards.
2. They allow railroads to gather data to inform their maintenance activities and maintenance planning. Maintenance planning involves determining what maintenance work should be done over time and how it should be carried out.

Track safety standards are a baseline for inspection frequency and set minimum requirements for track conditions. Railroads can choose to inspect track more often or hold their track to stricter standards than those set forth in the regulations. More frequent inspection or standards stricter than those required can result in greater maintenance flexibility. If a railroad identifies track degradation *before* it qualifies as a defect, they do not need to address it immediately and can be more strategic by also considering other maintenance needs at the same time.

Figure 4 provides a summary of the two track inspection purposes described above and their time scales. Addressing safety defects occurs on a shorter time scale, while maintenance planning occurs over longer time scales.

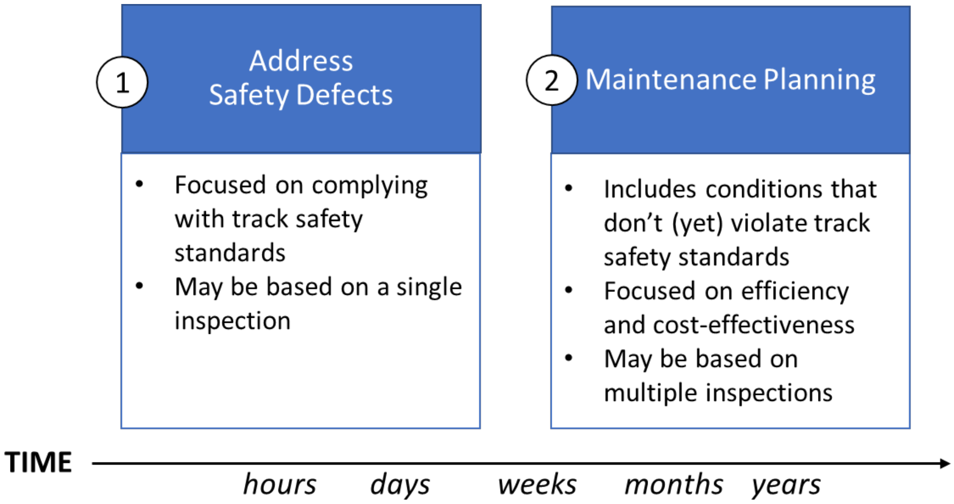


Figure 4. Dual purpose of track inspection

Track inspectors and technology operators use a variety of technologies and automated systems to identify degraded track conditions and defects during inspections. Track inspectors or maintenance crews repair these defects and degraded conditions during either planned or unplanned maintenance. The first inspection purpose shown in Figure 4 can be referred to as a "find and fix" approach. The focus is on fixing the symptoms (e.g., broken rail, misaligned switches, track geometry defects, etc.). The underlying causes that contribute to these symptoms may be unknown and/or unaddressed. This approach is common in the industry because it can be challenging to identify and address the causes of problems early.

Figure 5 shows the "find and fix" maintenance cycle. Track inspections can detect degraded conditions (e.g., track defects and maintenance conditions), which leads to either unplanned or planned maintenance. Following unplanned maintenance, track inspections continue to monitor track conditions and verify that the problem was solved.

Degraded track conditions also include maintenance conditions which do not require immediate action, but which railroads want to address before they become official defects and in a way that minimizes impact on revenue service. These documented conditions contribute to planned maintenance. During maintenance planning, railroads schedule maintenance activities, making more efficient use of limited track availability. Railroads consider not only present conditions but also anticipated future needs. They may perform larger-scale maintenance programs to replace track assets (e.g., tie replacement over a large area of track) or apply proactive treatments like rail grinding.

Railroads would like to shift the emphasis from a "find and fix" approach to a "predict and prevent" approach. In a "predict and prevent" approach, railroads identify underlying causes and prevent or minimize potential causes of track degradation. By predicting when track degradation will occur, railroads can better manage allocation of track between revenue service and maintenance and identify emerging issues that could affect safety (Kefalidou, Golightly, & Sharples, 2018; Golightly, Kefalidou, & Sharples, 2017).

As automation increases over time, the impacts of technology on track inspection frequency, revenue service capacity, and safety may change. For example, as track geometry inspection technology has become more automated, railroads shifted the responsibility for monitoring track geometry from exclusively relying on track inspectors walking the track or operating in hi-rail vehicles to operating conventional track geometry cars with onboard staff, and later to collecting geometry measurements using sensor systems attached to trains in revenue service ("Autonomous" Track Geometry Measurement Systems, or aTGMS). Each increase in automation has increased the amount of track that railroads can inspect for geometry defects within a fixed period. Collecting track geometry data during revenue service increases the amount of track inspected without adversely affecting revenue service. In fact, revenue service capacity may increase as this technology reduces the track time requirements for inspection. Lastly, there is the potential for this technology to increase safety. By inspecting more miles of track on a more regular basis, railroads can identify degraded track conditions sooner. For track geometry inspection, it is easy to see how the potential benefits of automation in Figure 2 could be realized.

Track geometry measurement is merely one component of track inspection. The safety considerations of automation and tradeoffs in track allocation between track inspection activities and revenue service must also be measured, particularly for track inspection activities beyond track geometry measurement. Nevertheless, this example illustrates the industry's vision for how automation can facilitate greater productivity and safety in railroad operations.

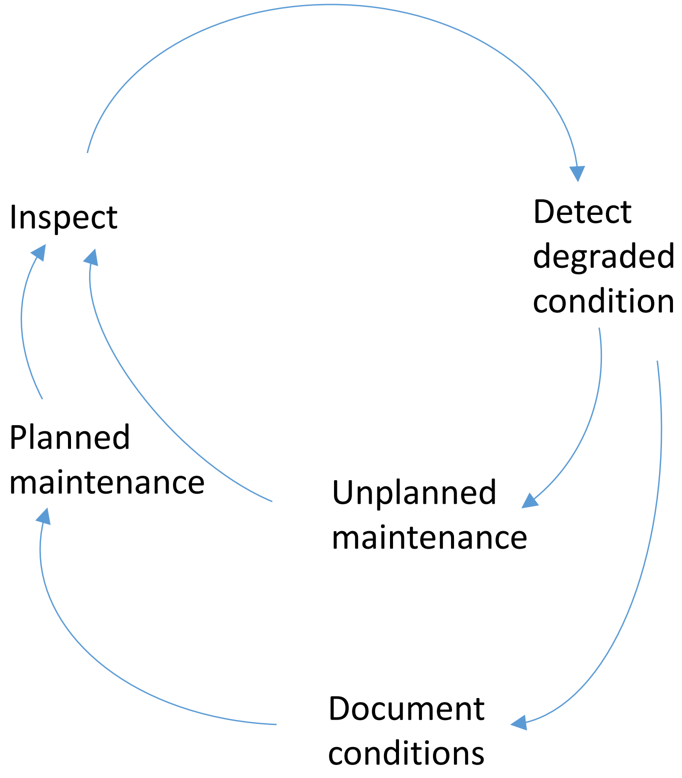


Figure 5. "Find and fix" railroad maintenance cycle

The railroad industry is striving to move from "find and fix" to "predict and prevent" through increased adoption of technology in the track inspection process. The use of automation may allow railroads to detect conditions earlier, granting them more flexibility in when and how to address issues. It also increases the amount of data available for analysis when making planning decisions and can provide other tools like predictive modeling to support decision making.

2.3 Augmenting Human Capabilities

Automation may realize track inspection benefits through augmenting the capabilities of human track inspectors.

When developing new technology to support safer and more efficient railroad operations, railroads and technology designers must consider how the joint track inspection system (i.e., humans and technology) functions. Should the human's role in track inspection change? If yes, how should the human's role change? Explicitly considering these questions is important for optimizing system performance.

The track inspection technologies that vendors are developing range from augmenting human track inspection tasks to conducting fully autonomous track inspection tasks.[2] Current technology has focused primarily on the detection and assessment of specific types of degraded track conditions. These technologies each focus on a particular type of track condition such as track geometry, rail flaw, tie, or ballast condition. Human track inspectors provide a more holistic perspective to track inspection. Designing for the performance of the joint system may result in technologies that make better use of the unique perspectives of human inspectors while augmenting them with the capabilities of automation.

---

[2] Although technology may autonomously perform certain tasks, such as data collection, humans may still play an important role in other tasks such as decision making, as well as in developing, programming, and maintaining the autonomous system. Therefore, human-centered design remains important for such systems.

3. Evolving Approaches to Function Allocation and Designing HATs

Designing HATs involves determining how to allocate tasks and functions to humans and automation, referred to as "function allocation."[3] Between non automated systems and autonomous systems, there is a complex range of ways that humans can interact with technology and automation.

Current approaches to designing HATs require considering the goals of the joint human-machine system, the tasks that humans and/or technology need to perform, and the strengths and limitations of the humans and technologies involved in that system. Understanding these factors and design tradeoffs can inform system requirements and contribute to better decisions regarding how to incorporate automation in track inspection. This is referred to as taking a "systems-based" approach, where the human is an integral part of the system.

This section describes three progressively more sophisticated approaches to function allocation to support the design of effective HATs:

- Capability-based approaches
- LOA models
- HAT approaches

3.1 Capability-based Approaches to Function Allocation

Early research focused on identifying the strengths and limitations of humans and automation and assigning task elements to whichever was better suited for each task. This idea is captured in the acronym MABA-MABA, which stands for "Men Are Better At – Machines Are Better At" (Fitts, et al., 1951). This approach for distributing tasks between humans and machines provides relatively limited support to designers because it addresses function allocation as an either/or approach (Roth, Shushereba, Militello, Diiulio, & Ernst, 2019). The common approach of designing the technology first and leaving the remaining work elements to the human may result in sub-optimal system performance where the humans are left with cognitively demanding tasks or tasks for which the humans are ill-suited (Dadashi, Wilson, Golightly, & Sharples, 2014). Although these early design approaches had limitations, they established that understanding the capabilities of both humans and automation is important when designing HATs.

3.2 Levels of Automation (LOA) Models

LOA approaches are more flexible in describing the range of ways that humans and automation can interact (Parasuraman, Sheridan, & Wickens, 2000; Save, Feuerberg, & Avia, 2012). This approach understands that humans and automation may both be involved in task performance, rather than allocating work solely to one or the other.

LOAs range from fully manual to fully autonomous task performance. However, most current track inspection technologies are *not* fully autonomous, and rely on humans to perform tasks like

---

[3] The authors of this report did not perform a function allocation, as the appropriate allocation of tasks depends on the goals and capabilities of the specific human-automation team. The process and considerations in this report are intended to be broadly applicable, and therefore do not address function allocation for any specific technology.

setting their course, monitoring their location, or reviewing data outputs for accuracy. Between the two extremes of manual and autonomous there are a range of possibilities for shared task performance and system designers must determine the appropriate interactions between humans and automation.

Additionally, even a technology that could operate autonomously during inspections would be programmed by humans, have output that is used by humans, and need to be maintained by humans. Therefore, it is important to understand how human-automation interactions exist even at the highest LOAs.

Section 5.2 explores how a LOA framework can be used to explain a range of human-automation interactions in track inspection context.

3.3 HAT Approaches to Function Allocation

More recent approaches to function allocation rely on a HAT perspective that involves understanding how humans and automation operate as a joint system, or team, and seeking to optimize the joint performance of that team.

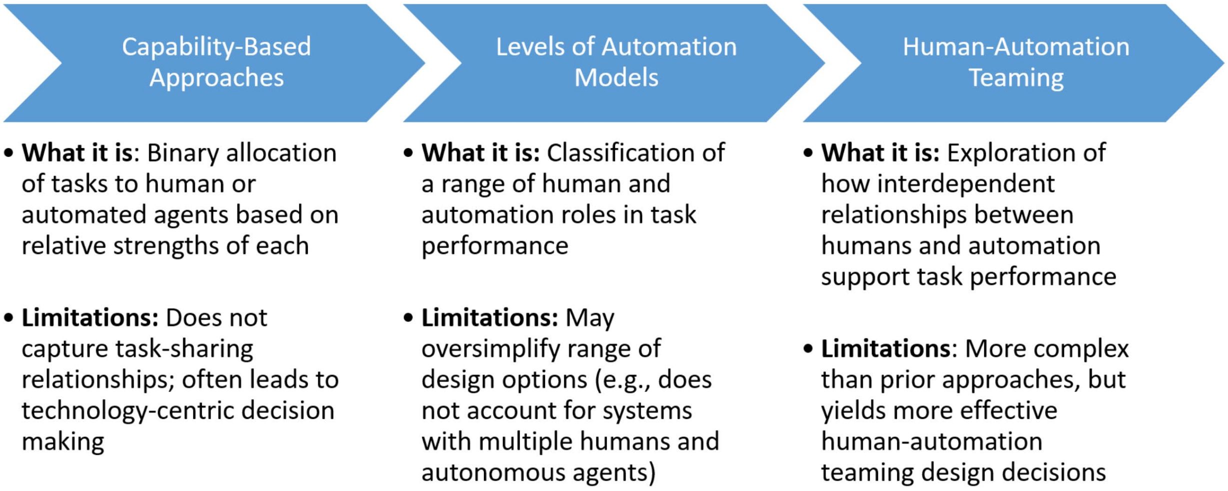
HAT approaches go beyond a LOA perspective in that a HAT approach recognizes that work may involve:

- Multiple human and automated agents across organizations (Singh, 2018)
- More dynamic distributions of workload between humans and automation (Gutzwiller, Caitlin, & Lange, 2018; Inagaki, 2001)
- More sophisticated communication requirements, similar to those of effective human-human teams (Joe, O'Hara, Hugo, & Oxtrand, 2015; Johnson, Vignatti, & Duran, 2020)

In other words, a HAT approach addresses real-world complexities more effectively than earlier approaches. There is a significant body of research on HAT and function allocation that seeks to understand these complexities and develop methods to manage them (Roth et al., 2019; Johnson et al., 2020; Cooke, Demir, & Huang, 2020). This report draws heavily on that literature in Section 4 to provide a general process for designing HATs and exploring function allocations for track inspection.

3.4 Summary of Function Allocation Approaches

Figure 6 summarizes the approaches to function allocation discussed in this section. The more sophisticated frameworks and tools may be better suited to the complexity of today’s technologies.



Capability-Based Approaches

- **What it is**: Binary allocation of tasks to human or automated agents based on relative strengths of each
- **Limitations:** Does not capture task-sharing relationships; often leads to technology-centric decision making

Levels of Automation Models

- **What it is:** Classification of a range of human and automation roles in task performance
- **Limitations:** May oversimplify range of design options (e.g., does not account for systems with multiple humans and autonomous agents)

Human-Automation Teaming

- **What it is:** Exploration of how interdependent relationships between humans and automation support task performance
- **Limitations**: More complex than prior approaches, but yields more effective human-automation teaming design decisions

Figure 6. Evolution of function allocation approaches

These approaches are not mutually exclusive. They build on one another to make up for the shortcomings of prior approaches. It is often useful to consider the relative capabilities of humans and automation or to refer to a LOA framework while taking a HAT approach to incorporate more nuanced design considerations.

The following questions can help system designers explore function allocation and HAT in track inspection:

- What are the strengths and limitations of both the humans and automation with respect to track inspection?
- To what extent should humans and/or automation be involved in various tasks and functions?
  - Are there tasks or functions for which it makes sense for the human to operate with little or no automation?
  - Are there tasks or functions for which it makes sense to augment or support human decision making and action with automation?
  - Are there tasks or functions for which it makes sense for automation to be the primary task performer (e.g., where the task is dangerous for humans)?
- How might humans and technology work together as a team to perform track inspection and address organizational goals around track inspection and maintenance?
  - How does automation change the nature of the work for humans, or the way the work is distributed? Does the automation enable the human to spend more time and attention on new tasks?
  - How might this human-technology team communicate and collaborate effectively? What is required to make this team successful?

To help support railroads and technology manufacturers in exploring these questions, this report uses several frameworks to explore the design of HATs for track inspection.

4. General Process for Designing Human-Automation Teams

As discussed in Section 3, there are a range of function allocation approaches that can be used to explore the relationships between humans and automation to design HATs. Regardless of the method chosen, it is important to understand how design decisions may affect system performance.

This section describes a general process for making HAT decisions and concepts from the research literature that can be used to support these decisions. This discussion is relevant to several audiences:

- Technology designers may follow this type of process to make sure these HAT considerations are included in the overall systems engineering process so that they can be addressed early.
- Railroads may follow this type of process to ensure that new technologies they purchase or procure will support effective HAT, or to assess their current inspection technologies and processes to determine if there are areas to improve.[4]
- FRA may use this type of process to pursue systematic evaluations when making decisions regarding the safety of novel automated technologies, or when funding research related to the development of new technologies.

The process described in this section provides a general overview for audiences that may be unfamiliar with these human factors, rather than a specific explanation for how to perform each of these activities. In other words, this is not a "how-to" manual. Readers may refer to existing design standards and guidelines for human-systems integration for more specific guidance related to these topics (Department of Defense, 2020; Federal Aviation Administration, 2016; National Aeronautics and Space Administration, 2014; National Aeronautics and Space Administration, 2014).

The Volpe team recommends seeking out a human factors professional to assist in performing the specific analyses mentioned in this section, or to assist in identifying the best analysis methods to use for a particular purpose. This section provides context to help readers understand more about the steps involved in designing HATs, why they are important, and how these activities might be integrated into existing systems engineering processes.

One way to frame the design of HATs is summarized in Figure 7. This process takes inspiration from the literature on HAT, function allocation, and interdependence analysis (Roth et al., 2019; Johnson et al., 2020; Cooke et al., 2020).

---

[4] For further guidance on including human-systems integration in the procurement process, see Melnik et. al (2018).

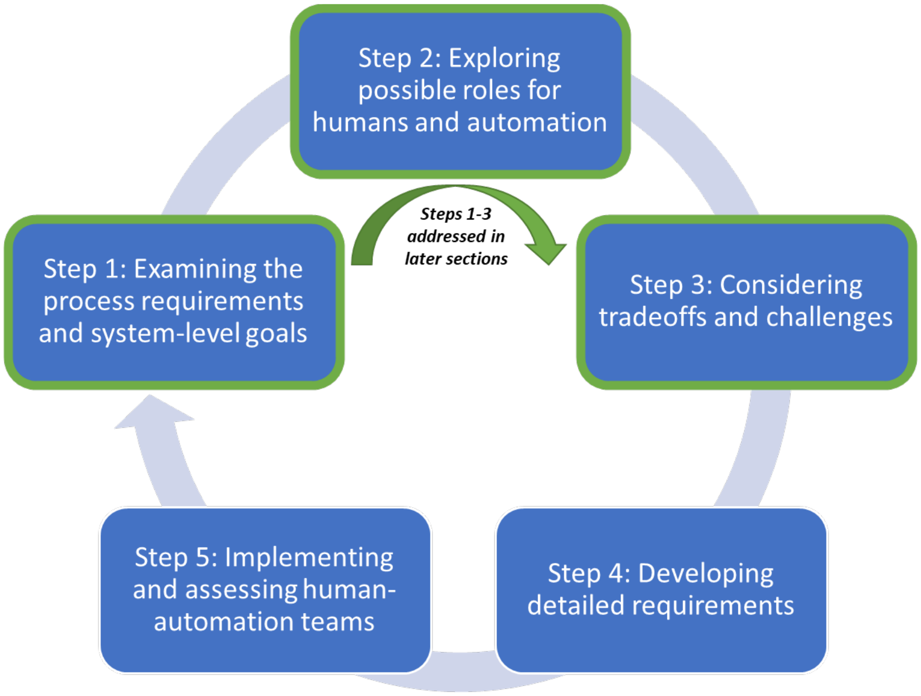


Figure 7. General process for designing HATs

The remainder of this section discusses each step of the process for designing HATs along with questions to consider at each step. A more detailed version of this discussion can be found in Appendix B. To see the full list of questions along with additional questions that can be used to guide design, see Appendix C.

4.1 Step 1: Examining Process Requirements and System-Level Goals

In order to allocate work effectively, the system designer must be very clear about all the work that must be done. Examining the work that the team will perform requires thinking beyond what the technology itself can do and thinking about all the work that must be done by the users of the technology and what must be accomplished by the HAT more broadly.

Roth et al. (2019) suggest two questions that system designers must ask at the beginning of the function allocation process: “what is the nature of the work to be done?” and “what makes it challenging?” Understanding the answers to these questions should precede asking “who does what?”

Roth et al. (2019) note that this examination of operational demands and challenges should include not only the typical functions to be performed, but also functions that may occur in non-routine situations. For example, in track inspection, consider not only routine inspections but also those that may be required after severe weather events or other circumstances that may result in damage to the track.

Two important considerations when examining process requirements and system-level goals are task decomposition and task interdependencies.

- **Task Decomposition.** Understand how work can be broken down into tasks and subtasks. This helps ensure each individual part of the process is considered and allows decision makers to consider how each task can be divided or shared among humans and automation.

  This report uses a simple task decomposition for track inspection based on a four-stage cognitive information processing model (Parasuraman et al., 2000). The model breaks the track inspection process into the tasks of data collection, data analysis, decision making, and action. Task decomposition is detailed in Section 5.1.1.

- **Task Interdependencies.** When tasks are decomposed into subtasks, there may still be critical interactions between them. The relationships between tasks have implications for how to structure the HAT and the design requirements. For example, if all activities must be completed in sequence, then humans and technologies will need to coordinate to hand off each task to the other.

Designers may consider the following questions when examining process requirements:

- What are the overall goals of the process?
- What makes that work process challenging?
- How can the process be broken into smaller tasks for analysis?
- In what ways are these subtasks interdependent?

4.2 Step 2: Exploring Possible Roles for Humans and Automation

Once system designers have examined the process requirements and system level goals, the next step is to explore possible roles for humans and automation for each task. This includes examining these possible roles, examining various ways humans and automation could work together to perform the task, and their capabilities to perform or support various activities, as well as noting interdependencies between team members.

- **Exploring roles using LOA.** LOA models provide a common way to describe and categorize human-automation interactions (e.g., sharing or trading off work, using automation as a backup to human actions, or using automation to replace some subset of the tasks a human normally performs).
    - This report uses a four-stage information processing model (Parasuraman et al., 2000) to examine the roles involved in the track inspection process for each of four inspection tasks. The four tasks are data collection, data analysis, decision making, and action. Section 5.1.1 describes these inspection tasks in more detail.
- **Exploring roles using teaming models.** In addition to using LOA models, it is often valuable to examine possible roles for humans and automation that do not fit neatly into such a framework. As noted in Section 3.3, recent developments in the HAT literature provide holistic perspectives on the ways that humans and automation can perform work.
    - In examining the roles for humans and automation within a system, consider systems where work may be distributed across a range of people and automated technologies (Singh, 2018). Where appropriate, designers may also examine ways that the workflow can be altered dynamically (e.g., adaptive automation). This

    ability to adjust who does the work based on the work context represents a more recent consideration in the design of HATs (Gutzwiller, Caitlin, & Lange, 2018; Inagaki, 2001). These approaches call for examining the system in a more holistic fashion to address the interdependent nature of the work.

- **Assessing capabilities of humans and automation for each role.** As noted in Section 3.1, the capabilities of humans and automation were some of the earliest criteria used in function allocations. Technology has evolved significantly since these early capability-based function allocation methods were developed, and automation can now perform many tasks that were previously only possible for humans. Examining the capabilities of humans and automation remains an important aspect of system design.
    - Modern techniques take a nuanced approach to integrating human and machine capabilities into function allocation to capitalize on the strengths of each, providing a wider range of possibilities for HATs. Examining these possibilities also allows system designers to consider whether they are creating supporting roles that would be very difficult for humans to perform, a common pitfall of technology-centered design approaches.
- **Teaming interdependencies.** Once designers identify possible roles for humans and automation and examine their ability serve in primary or supporting roles, it is important to determine where teaming interactions or coordination between humans and automation are necessary, as well as where coordination is optional but could improve efficiency or reliability. Johnson et al. (2020) discuss the importance of interdependence and present a method for analyzing interdependencies, as well as a set of principles that can be used in designing HATs.
    - If the human or automation performing the task has limited capability or reliability, a supporter may be important to make sure the task is successful. For example, a technology with low reliability in identifying defects may require a human to carefully review its outputs. Similarly, a human with low ability to detect internal rail flaws may require an automated technology to perform data collection to effectively detect such defects.
    - In cases where a supporting actor is not required, “opportunistic teaming” can contribute to system resilience. For example, following a visual inspection with an automated inspection technology may increase the likelihood of detecting defects. Simply allocating a task to the human *or* automation would limit the HAT’s success. Using a combination of human and technology to perform a task may be beneficial or even necessary for optimal task performance.

Designers may consider the following questions when exploring possible roles for humans and automation:

- What roles can humans and automation perform or support?
- What LOAs are possible or desirable?
- In what ways could humans and automation interact?
- In what ways does this task depend on or influence others?

4.3 Step 3: Considering Tradeoffs and Challenges

When considering tradeoffs and challenges associated with the desired teaming structure, each option has strengths and weaknesses. System designers must determine which option is best suited to meet the system goals and requirements (as discussed in Section 4.3). This requires assessing human factors impacts of possible teaming structures and selecting options that meet the system requirements.

- **Assessing human factors impacts of possible function allocations.** In assessing tradeoffs and challenges associated with possible roles for humans and automation, consider how these decisions affect the humans in the system.
    - For example, one possible impact of increasing the use of automation is loss of situation awareness (Endsley, 1995). As organizations introduce more automation, the potential exists for humans to become more removed from the details of their work. These changes contribute to loss of situation awareness and skill loss that make it challenging to take over when the automation fails. There are several considerations that should be examined when assessing the viability of design decisions related to HATs (e.g., the characteristics described in Billings (1997), such as employee selection and training).
- **Selecting options that meet system requirements.** Roth et al. (2019) suggest that tradeoffs may take into consideration several factors, such as maturity of the required technology, human-system integration considerations (e.g., personnel selection and training costs), and operational considerations (e.g., ability for humans and/or automation to perform under unusual circumstances).
    - For track inspection, system designers, researchers, and railroads should consider which combinations of automated technologies and human actors best meet their needs. This may be partly shaped by constraining factors such as costs of inspection technologies, compliance with regulations, requirements for track time, and staging and training requirements.
    - Requirements such as those described by Pritchett, Kim, & Feigh (2014) can be used to examine whether the function allocation or HAT structure is likely to result in effective work or encounter problems.

Designers may consider the following questions when considering tradeoffs and challenges:

- What are the benefits and challenges for each option?
- Do the roles for both humans and automation make sense?

4.4 Step 4: Developing Detailed Design Requirements

Once system designers determine which HAT structures are best suited to meet the system goals and requirements (based on the tradeoffs and considerations discussed in Section 4.3) they must determine the system's design requirements.

Designers may derive some design requirements from the interdependencies identified in prior steps, as well as the expectations for what tasks the HAT will perform. Some requirements may address activities that occur outside the inspection process itself, such as training. These design requirements can help address limitations associated with the selected HAT structure.

- **Develop requirements based on task interdependencies and expectations for the HAT's performance.** For tasks that are shared among humans and automation, interdependencies of the process may shape design requirements. For example, Johnson et al. (2020) cite requirements for observability, predictability, and directability as important parameters to consider. These factors describe "who needs to observe what from whom, who needs to be able to predict what, and how members need to be able to direct each other for a given aspect of the work." These requirements may dictate design decisions such as the locations of the humans and automated technologies, the timing of their activities, and the design of human-machine interfaces.
    - Examples in the research literature can be used to develop detailed design requirements. For example, researchers have developed criteria for HATs based on what is known about human-human teams (Joe et al., 2015; Johnson et al., 2020). Designers may establish work agreements to describe the roles of humans and automation and the tasks they will perform, and to establish shared expectations (Gutzwiller, Espinosa, Kenny, & Lange, 2018).
- **Consider requirements beyond the inspection process itself (e.g., training requirements).** Requirements may also extend beyond the four track inspection tasks discussed in this report. For example, researchers have examined how requirements for training may change as with increased use of autonomous systems. One study of unmanned aerial vehicle (UAV) operators found that humans may require less skill-based training to achieve the same performance when supported by automation. However, the study also found that novice operators may have a greater bias toward trusting the automation and overlooking when it makes errors, whereas experienced operators may be better able to catch these errors (Cummings, Huang, Zhu, Finkelstein, & Wei, 2019). New forms of training may be required to develop appropriate levels of trust in automation and maintain the expertise needed to identify automation errors.

Designers may consider the following questions when developing detailed design requirements:

- What is necessary to make the HAT successful, particularly given any interdependencies between tasks?
- What challenges can be addressed through system design?
- What additional requirements (e.g., training) should be considered?

4.5 Step 5: Implementing and Assessing HATs

After design decisions have been made, it is important to assess whether the HAT is functioning as intended. This involves selecting assessment criteria and determining whether to adjust the design or operation of the HAT.

- **Choosing assessment criteria.** Criteria for evaluating HATs, such as those developed by Pritchett et al. (2014), can be used to assess proposed HAT structures before they are implemented and to assess existing HAT structures. Designers should select assessment criteria that consider the impacts on human team members as well as the performance of the HAT.

- **Making changes to the design of new or existing HATs.** Although integrating HAT considerations into the systems engineering process early on is most effective, designers can still make improvements after a certain technology is implemented, particularly if unforeseen challenges related to human-automation interaction arise. However, when assessing existing systems, the ability to make improvements in system performance may be more limited and/or more costly than if designers had conducted such an assessment earlier in the design process.

Designers may consider the following questions when implementing and assessing HATs

- Once implemented, how will system performance be monitored?
- What unanticipated challenges have been encountered?
- Where is there room for improvement?

5. HAT in the Track Inspection Process

This section discusses the following aspects of designing for HATs as they can apply to track inspection:

- Step 1: Examining process requirements and system-level goals
- Step 2: Exploring possible roles for humans and automation
- Step 3: Considering tradeoffs and challenges

The last two steps are beyond the scope of this report, as they must be considered with respect to specific technologies and are difficult to discuss in a generalized way:

- Step 4: Developing detailed design requirements (as discussed in Section 4.4)
- Step 5: Implementing and assessing specific teaming structures (as discussed in Section 4.5)

5.1 Step 1: Examining Track Inspection Process Requirements and System Goals

The Volpe team previously documented some of the operational demands associated with the track inspection process for visual and TGMS inspections (France et al., 2020). Challenges associated with track inspection include:

- Maintaining knowledge of track defects (e.g., managing training and qualifications for inspectors, passing along hands-on knowledge, being able to verify outputs of automated technologies, etc.)
- Complying with track safety standards (e.g., meeting inspection frequency requirements and identifying and addressing defects within required timeframes)
- Communicating information effectively (e.g., coordinating track usage with dispatchers, relaying data from automated inspections, reporting to supervisors)
- Managing resource and production pressures (e.g., managing requirements for track time to perform inspections and maintenance, allocating optimal territory sizes for inspectors, minimizing track outages and speed restrictions)

This list illustrates some of the complexities of the process. Railroads or technology designers may wish to perform additional analyses to understand other specific aspects of the track inspection process and the challenges involved.

5.1.1 Decomposing the Track Inspection Process into Four Tasks

Decomposing a process into tasks is one way to better understand the requirements and complexities of that process.

The authors divided the track inspection process into four tasks based on the four-stage information-processing model proposed by Parasuraman et al. (2000). This model provides a simple way to understand how humans and automation can work together to process information. The four tasks include the following:

- **Data Collection** involves observing or measuring track condition using human senses, sensor technology, or some combination of the two.
- **Data Analysis** involves analyzing track conditions to determine whether a degraded track condition is present, and if so, the severity of the condition. This analysis can be done by humans, computers, or a combination of the two.
- **Decision Making** involves determining what actions to take to address the conditions identified in the data analysis. These decisions may include making repairs, putting speed restrictions in place, or taking track out of service. The details of these decisions (e.g., how much to restrict speed or what type of maintenance to perform) depend on the type and severity of the problem.
- **Action** involves humans and/or technologies taking the action(s) decided upon during decision making. This might include contacting the dispatcher to impose a speed restriction or remove track from service. It may also include performing a repair or contacting a maintenance-of-way crew to instruct them to do so. Any actions taken must be documented.

These four tasks are interdependent, and do not typically occur in a linear sequence; track inspection tasks form a continuous cycle. For example, data analysis may reveal additional information needed to reach a decision, leading to additional information acquisition.

5.1.2 Activities in the Inspection Process

As noted in Figure 4, track inspections serve a dual purpose: to identify safety defects that railroads must address in the near term and to inform maintenance planning that takes place over a longer time frame. Table 2 summarizes specific activities that may occur within the track inspection process tasks, and how they relate to long-term and short-term goals.

Table 2. Summary of Track Inspection Process Tasks and Activities

Cognitive Tasks in the Track Inspection Process	Short-Term Focused Activities	Long-Term Focused Activities
Data Collection	• Human observations (visual, auditory, kinesthetic) • Sensor measurements • Image or video capture • Simple highlighting or filtering of data	• Repetition of data collection • Continuous monitoring
Data Analysis and Verification	• Data manipulation through computation and comparison • Pattern recognition • Comparison to allowable thresholds	• Integration of data from multiple sources • Identification of trends • Predictive modeling
Decision Making	• Make severity judgments • Decide on actions (repair, remove, or restrict) if needed	• Maintenance planning
Action	• Remove track from service • Restrict track speed • Perform maintenance • Log inspection	• Proactive and preventative maintenance

Activities also occur outside this sequence. These activities do not *directly* contribute information to the process of inspecting track, but still shape the inspection process. Examples include:

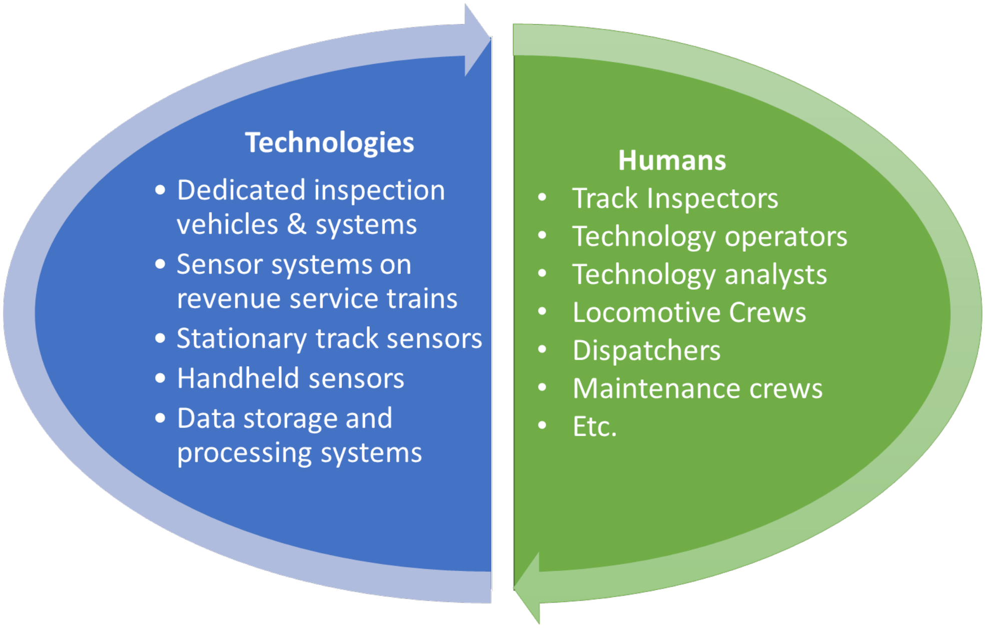
- Preparing employees to perform inspection duties (hiring, training, job assignment, etc.)
- Determining what tools or technology to use to perform the inspection
- Scheduling inspections to meet the time frames mandated by FRA regulation
- Obtaining track time for inspection, if needed
- Traveling to/from inspection location

Many of these activities occur prior to the inspection itself. Others, like traveling along the track, may occur simultaneously with the inspection but do not directly contribute to collecting and analyzing information to make decisions. Though this report does not dedicate specific sections to these activities, it does discuss them in the context of design considerations for HATs.

5.2 Step 2: Exploring Roles for Humans and Automation in Track Inspection

When exploring roles for humans and automation, consider all the humans and automation involved in inspection, not just a single technology and/or a single inspector or operator. Figure 8 shows that there are many people and technologies involved in track inspection to consider when making decisions regarding HAT and the design of effective track inspection processes.

The following sections of this report explore ways to examine these roles and determine how to share inspection tasks among humans and technology. This section first explores these roles using a LOA framework, followed by more holistic HAT framework.



Technologies

- Dedicated inspection vehicles & systems
- Sensor systems on revenue service trains
- Stationary track sensors
- Handheld sensors
- Data storage and processing systems

Humans

- Track Inspectors
- Technology operators
- Technology analysts
- Locomotive Crews
- Dispatchers
- Maintenance crews
- Etc.

Figure 8. Humans and technologies involved in track inspection

5.2.1 LOA Using the Four-Stage Model

LOA models are one of the most common ways of exploring roles for humans and automation. These models are often used to discuss automation of vehicle operation. The automotive industry uses five levels to describe the role that automated features play in the driving task (SAE 3016). Similarly, the International Association of Public Transport (UITP) defines five Grades of Automation (GoA) according to the level of responsibility assigned to the train control system. These range from manual operation with no automation (GoA0) to unattended train operation (GoA4). Each level is defined by the tasks allocated to the human and to the technology.

In both rail and automotive domains, these LOA models are simply ways of *classifying* automated features and allowing comparisons among features with different capabilities. There may be many different automated features that fit into the same level but operate differently. Knowing the LOA is not the same as having a complete description of the feature, how it works, or the human-machine interface.

Figure 9 depicts a track-focused LOA model using the four-stage information processing model described in Section 5.1.1. It illustrates that for each of the four track inspection tasks, the LOA may range from fully manual to fully autonomous.[5]

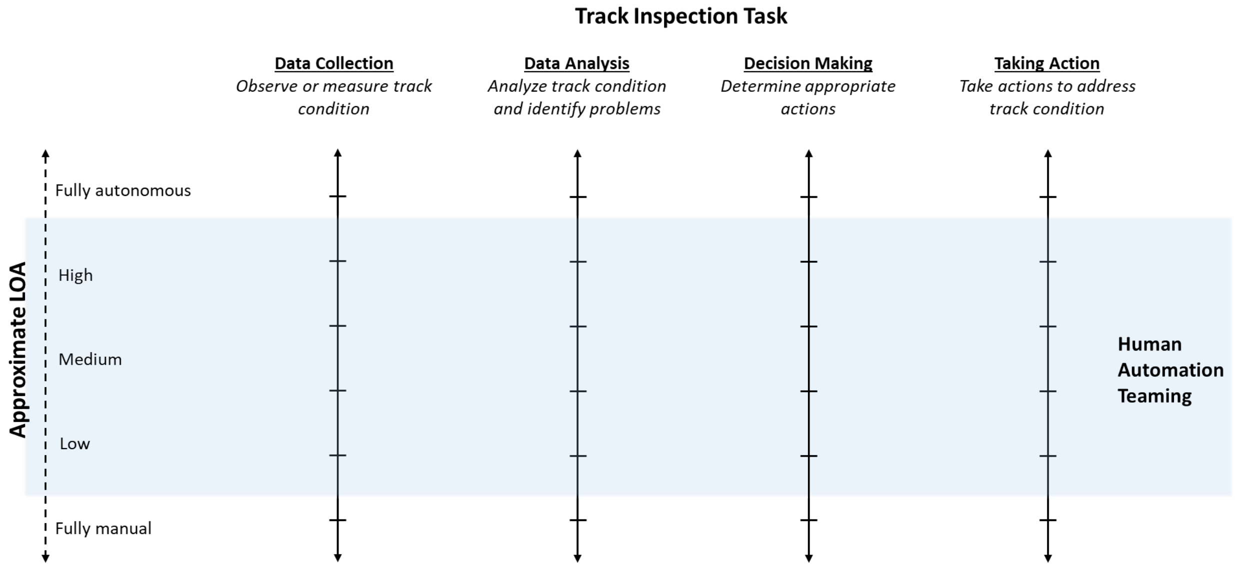


Figure 9. LOA can vary by inspection task

The region of greatest interest is highlighted in blue. This region represents the area where both humans and automation are involved in the performance of the task, and it encompasses many different possible combinations of humans and technology. At the lower end of the scale are tasks where the human is primarily responsible for performing the task with some simple tool or technology support, while at the higher end, automation may operate with a human monitoring the automation. This model shows that there are a wide range of ways that HAT can interact.

---

[5] This report does not examine or advocate for fully autonomous performance of track inspection tasks. This figure is merely included to demonstrate the range of automation levels.

The LOA may vary across these four tasks. Both high and low LOAs have advantages and disadvantages that must be considered and balanced when designing HATs.

Upcoming sections discuss how LOA models can be used to explore roles for humans and automation in data collection (Section 6.2.3), data analysis (Section 7.2.3), and decision making and action (Section 8.2.3).

5.2.2 Exploring Alternate Roles for Humans and Automation in Track Inspection

The choice of automation level is one of many design decisions that shapes the interactions between the human and automation. LOAs define what the human and automation do, but do not define how they interact with each other and the track environment. Some important aspects of human-automation interactions in track inspection not covered by LOA models include:

- **Location of the human and technology.** Is the human present when the data is collected? Or are they in a remote location?
- **Timing of human and automation receiving the data.** If data is collected by sensors and reviewed by humans, when does review happen? In real time or after the fact? Is this true for all data, or do different pieces of data arrive at different times?
- **Interface design considerations**. If it is collected by automation but some or all of the data is reviewed by a human, how does this occur? Does the interface introduce any challenges to detecting defects and maintenance conditions?
- **Collaboration among multiple automated and human entities.** LOA models typically focus on a single human and a single automated system. However, railroads may collect data from multiple technologies, and may have multiple operators and analysts supporting those technologies, in addition to the human inspectors conducting regular visual inspections.

The four-stage information processing model does not depict the ways that tasks are interdependent, or ways that they may occur outside of a linear sequence.

Railroads may require a human to verify that a defect is present after an automated technology identifies an exception. This could trigger additional (manual) data collection activities following data analysis to flag a potential defect. Figure 10 illustrates that this verification process, shown in green, can be thought of as a second data collection activity, separated in time from the initial automated data collection. Both initial data collection and verification are important parts of the inspection process, though they do not follow a purely linear path through the four inspection tasks. For the purposes of this report, this verification task is discussed under “data analysis” as it follows the initial analysis of automated data.

The tasks in the track inspection process also occur in a regular cycle. Data collection is regularly repeated after the action task to confirm that those actions were effective and to monitor for further track degradation over time. Therefore, it is inaccurate to consider the process as “complete” after the action task.

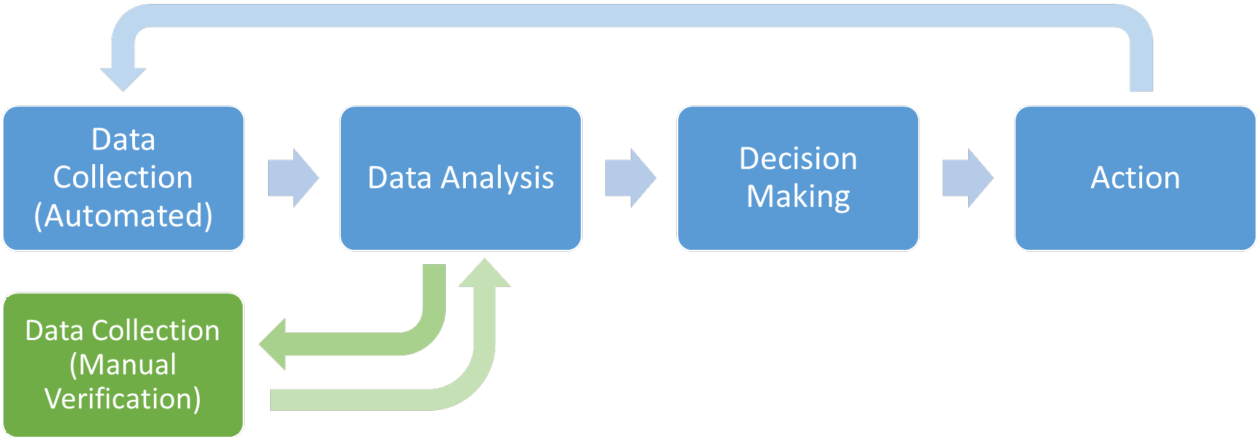


Figure 10. Iterative nature of the track inspection process, including repeated data collection activities for verification

Those who wish to understand human-automation interactions and optimize the performance of the track inspection process should consider additional characteristics of HATs that are not included in a LOA model. Upcoming sections discuss possible teaming structures for data collection (Section 6.2.4), data analysis (Section 7.2.4), and decision making and action (Section 8.2.4) that go beyond a LOA framework.

5.3 Step 3: Considering Tradeoffs and Challenges in Track Inspection HATs

As discussed in Section 4.3, it is essential to consider tradeoffs and challenges associated with any desired teaming structure.

These considerations relate to both the short-term goals of inspection (maximizing the effectiveness of defect detection) while others focus on long-term goals (seeking ways to use track inspection data to improve maintenance planning).

To support railroads and technology developers in understanding the types of human factors considerations associated with the track inspection process, upcoming sections of this report discuss considerations and tradeoffs for various methods of data collection (Section 6.3), data analysis (Section 7.3), and decision making and action (Section 8.3).

6. Data Collection: Observe or Measure Track Condition

Data collection is the first cognitive task in the track inspection process. In this task, humans and automation acquire information about the track condition using human senses, sensor technology, or some combination of the two.

This section addresses steps 1 through 3 of the general process for designing HATs as summarized below. Refer to Section 5 for details on these steps.

- **Step 1: Examining track inspection process requirements and system goals.** This section describes key aspects of the data collection task.
- **Step 2: Exploring roles for humans and automation in track inspection.** This section describes how human-automation interactions involving data collection can be understood using two frameworks: a LOA framework and a more nuanced HAT framework that considers the possible roles of multiple humans and automated systems.
- **Step 3: Considering tradeoffs and challenges in track inspection HATs.** This section describes considerations for railroads and manufacturers when making decisions about how to allocate data collection tasks and design effective HATs.

6.1 Understanding the Data Collection Task

Data collection is the foundation of track inspection. Railroads need track condition data to make informed maintenance decisions. This data is essential for both purposes of track inspection: finding and addressing any defects that need immediate attention and informing the railroad's future maintenance plans. The later track inspection tasks (i.e., data analysis, decision making, and action) depend on the quality and quantity of data obtained during data collection.

Aspects of data collection to consider include:

- **Location of data collection**: Data collection may be continuous or discrete. Continuous data collection activities are performed along the full length of the track (e.g., assessing track geometry and looking for defects such as wide gage), while discrete data collection activities are focused on examining specific track features (e.g., inspecting switches and highway-railroad grade crossings).
- **Type of data**: The federal Track and Safety Standards (Track Safety Standards 49 CFR Part 213, 2019) includes 21 categories of track defects in 4 areas (roadbed defects, track geometry defects, track structure defects, and track appliances and track related devices defects). Type of defect can also be considered with respect to whether the defect classification is objective (i.e., based on comparing measurements to a threshold) or subjective (i.e., based on pattern recognition and often relying on multiple parameters).
- **Method of travel:** Human inspections may be conducted on foot or in a hi-rail vehicle, and automated track inspections may happen using a system attached to a train in revenue service, a train not in revenue service, or using a self-propelling technology. Track inspections that occur on foot or in a hi-rail vehicle take longer than automated inspections that operate in revenue service.

6.2 Roles of Humans and Automation in Data Collection

During data collection, human inspectors observe the track condition, gathering information they can see, hear, or feel. When inspectors acquire information manually (i.e., no technology or a minimal level of technology), they observe the track using their senses and hand-tools, such as measuring tapes. They may walk or use hi-rail vehicles to traverse the track while gathering the data, but any technologies they use to support data collection are relatively simple and augment their ability to detect degraded track conditions.

Automated technologies support data collection using a range of sensors and data capture devices, such as cameras (capturing images or video), accelerometers, LiDAR sensors, radar, X-rays, etc. These devices automate the data collection task by aiding humans in gathering and perceiving information. In gathering information, automation can increase the speed at which data can be collected. In perceiving information, the sensors can detect degraded track conditions that humans would not otherwise detect.

Regardless of the combination of humans and automation used in track inspection, railroads must comply with FRA's track safety standards. Currently, track safety standards mandate visual inspections at a given frequency for each track class. However, these standards do not restrict railroads' ability to perform more frequent inspections or to supplement visual inspection with automated technologies.

The following section will discuss some of the roles and capabilities of humans and automation to perform or support data collection tasks. This is not an exhaustive list of roles. However, it illustrates the relative strengths that each brings to the process.

6.2.1 Roles of Humans in Data Collection

Collecting data about track defects is primarily a human role, in accordance with current track regulations requiring regular visual inspections. In cases where automation is involved, humans remain needed to support the data collection.

- **Performing data collection.** Human inspectors acquire a holistic understanding of the interrelationships between track conditions that current automated inspection technologies are currently not designed to address. Human track inspectors can integrate information across inspections including their knowledge of historical data along with operational considerations that may explain current track conditions. Automation does not understand or make sense of the data it collects. The task of comprehending and creating meaning from the collected data is one that only humans can do.
- **Supporting data collection.** Humans can also play several roles in supporting data collection. As noted above, humans can make sense of the data that automation collects. Often, automated systems will have onboard operators or data analysts to help review the gathered data for quality and accuracy. Humans may also review this data remotely.

  Other humans involved in supporting data collection include dispatchers and supervisors. Dispatchers facilitate data collection by granting access to the track, and supervisors oversee inspectors and provide instructions, such as when a special inspection is needed.

6.2.2 Roles of Automation in Data Collection

Each automated inspection technology relies on specific sensors to measure aspects of track condition and identify a subset of the defects specified in track safety standards.

Automated technologies may be used in track inspection for the following purposes.

- **Dedicated automated systems for different track conditions.** It may be necessary to integrate different types of defect data so that humans sent to verify the presence of a defect or perform repairs do not need to travel to the same area multiple times.
- **Using automated technologies to collect multiple types of track condition data.** There may also be interactions between these defect types, where one type of degradation contributes to another (e.g., loose ballast leading to a geometry issue), or where multiple types of track degradation occur in the same location to create a more significant hazard.

  Some rail industry stakeholders have expressed interest in developing inspection vehicles that integrate sensors on a single vehicle. This could improve efficiency of data collection and may allow for other benefits, such as data outputs that help integrate information from these multiple sources. (Section 6.3 discusses how integrating data from multiple sources poses challenges for data collection).
- **Role of multiple automated technologies in verifying track defects.** Technologies that address the same type of defect may be used together. For example, ultrasonic inspection systems can take several forms, from vehicle-based units to systems that can be pushed along the track while walking, to small handheld units for very localized detection. Each of these can play a distinct role, as vehicle-based units offer greater speed and handheld units may offer greater accuracy. A vehicle-based unit can be used to hone in on a location, whereas a human may use a handheld unit to verify the presence of a defect. These technologies, together with the humans in the system, support one another to collect data more effectively.
- **Increasing data availability about a particular condition.** Another example of using multiple technologies to supplement one another is the use of UAS to inspect railroad crossings, which are normally inspected using LiDAR. Adding UAS allows these inspections to occur more frequently.

6.2.3 LOA for Data Collection

This section uses LOA to explore ways that data collection functions may be allocated to humans and automation. Table 3 provides examples of LOA for data collection ranging from low to high.

Table 3. Example LOAs for Data Collection

Approximate LOA	Description	Example
Low	Automation supports the human by using sensors to scan and observe the environment and presents the complete set of raw data to a human.	*A system that takes images or videos of the track and transmits these directly to a human without highlighting or filtering any information.*

Approximate LOA	**Description**	**Example**
Medium	The human does not need to interact with the full set of raw data. The technology may assist human perception by sorting or highlighting information, while the human retains access to the full dataset and can choose to look at things the technology has not highlighted.	*A conventional TGMS geometry car, where the technology highlights measurements that exceed a programmed threshold, but the onboard operator can view the full dataset and strip charts of raw measurements.*
High	The technology may also *filter* the data. The full dataset is not visible to the human, only that which the technology has selected. In some cases, at the highest LOAs, the technology may not present any information to a human at all until performing further analysis.	*An aTGMS system that runs continuously and notifies employees only when it identifies an exception or potential defect.*

Technology that operates without human intervention is referred to as *autonomous*. However, technology that collects data in this way still involves humans in other track inspection tasks, including data analysis, decision making, and action. The degree of automation of the data collection task affects human performance in these related inspection tasks. Format of data presentation affects how the track inspector makes sense of the data.

Sometimes data collection occurs in ways that do not fit neatly into this framework.[6] This is why it is also useful to consider additional ways of thinking about automation, such as HAT.

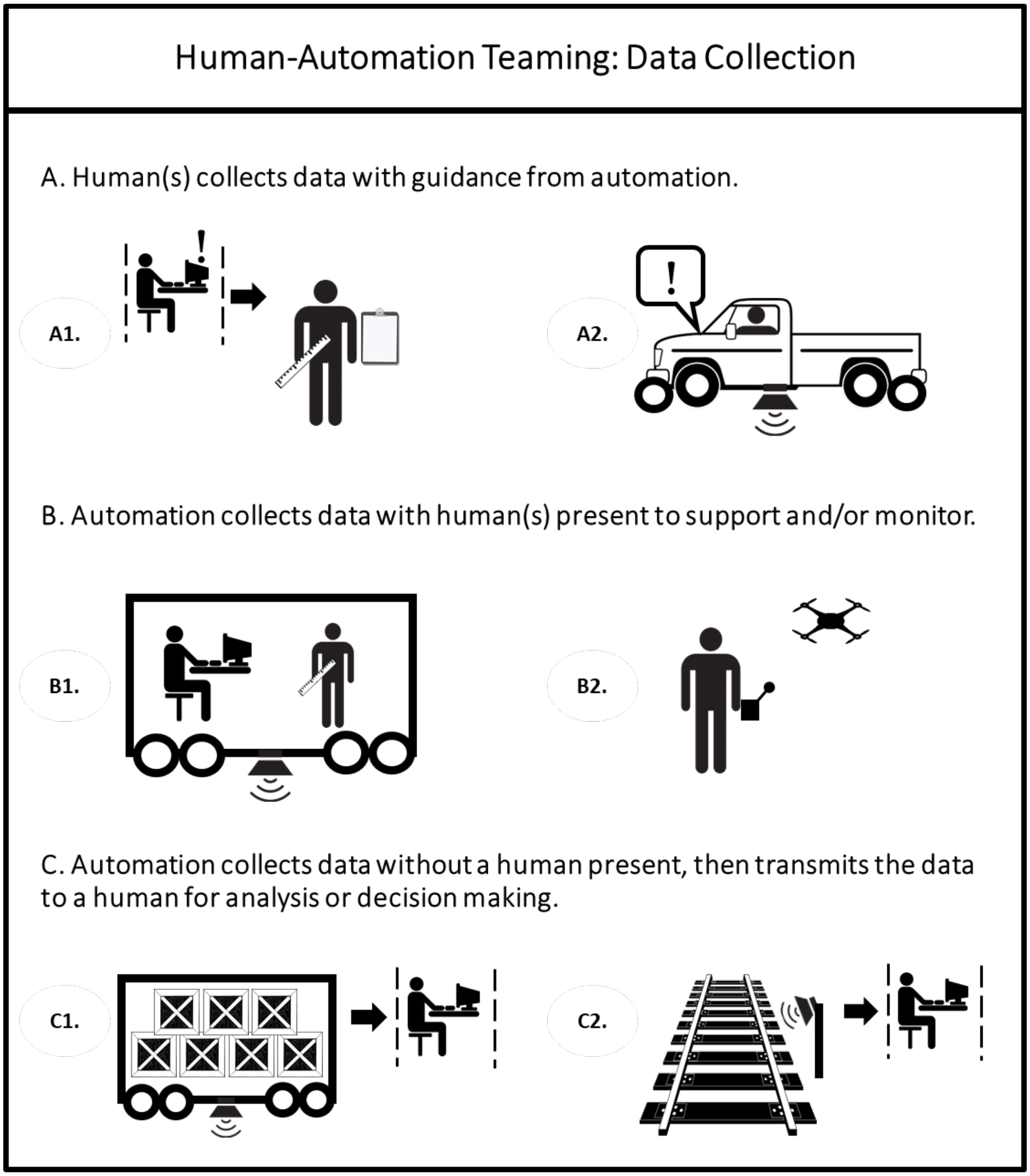
The remainder of Section 6 explores the roles that humans and automation can play in data collection and interactions between them.

6.2.4 Possible Teaming Structures for Data Collection

Figure 11 depicts possible teaming structures, or interactions between humans and automation, that can be used in data collection. These teaming structures vary both in their LOA and in other characteristics, such as the location of humans and technology and the timing of their activities.

---

[6] For example, LOA models do not: (a) capture the extent to which multiple data collection activities focused on different defect types may influence one another, (b) address interactions between systems involving multiple humans and automated systems, nor (c) describe the location of the human and automation and the relative timing of their involvement in data collection.



Human-Automation Teaming: Data Collection

A. Human(s) collects data with guidance from automation.

B. Automation collects data with human(s) present to support and/or monitor.

C. Automation collects data without a human present, then transmits the data to a human for analysis or decision making.

Figure 11. Examples of teaming structures for data collection

- **Human(s) collect data with guidance from automation.** The first set of teaming structures, labeled "A," involve human(s) that collect data with guidance from automation. In example A1, an inspector may use data from past automated inspections to identify suspected problem areas that should receive extra attention during their inspection. In example A2, the inspector traverses the track in a hi-rail vehicle equipped with a sensor system that provides alerts under certain conditions, and the inspector may get out of the vehicle to look more closely at the track. In both examples, the inspector is primarily responsible for the inspection while the automation provides guidance.

- **Automation collects data with humans present to support or monitor**. The second set of teaming structures, labeled "B," involve automation collecting data with human(s) present to support or monitor data collection.

  In example B1, staff onboard an inspection vehicle monitor the data as it is collected and can exit the vehicle to verify the data's accuracy. This is how data is collected in a conventional TGMS geometry car.

  In example B2, a human operator uses an automated system to collect data, such as by overseeing the flight of a UAS that has been preprogrammed to fly a particular route and photograph the track.[7] A similar example would be operating a handheld ultrasonic testing device, which requires a human operator and is not mounted to a vehicle.

  These are all examples of humans supporting automated data collection that are similar in their LOA yet differ in their method of operation.

- **Automation collects data without humans present and transmits the data.** The third set of teaming structures, labeled "C," involve automation collecting data without human(s) present and transmitting that data elsewhere for analysis and decision making. While humans may not be involved in the data collection, they may be involved in later inspection tasks.

  In example C1, a sensor system is mounted to a vehicle in revenue service that traverses the track and transmits data to a server located in an office. aTGMS collects and transmits data this way. Technologies such as aTGMS can operate without an onboard operator and dedicated track time; they collect data while the equipment is in revenue service. An operator reviews the data stored on the server. Review of this data does need to take place in real time.

  In example C2, a sensor system is installed on or near the track and transmits data about that location elsewhere for monitoring. Railroads use these types of sensors to monitor rail stress and rail temperature.

  Although these examples may use a similar LOA, they pose different operating requirements. In example C1, differences could occur in the amount of data transmitted (e.g., full sets of track measurements, or alerts only in the case of an exception), and the timing of the transmission (e.g., in real-time or later, in summary format). In C2, the operating requirements are minimal, but the sensors only collect data where they are installed.

Railroads may use several of these teaming structures in different contexts or to collect different kinds of data. Using multiple data collection methods, whether manual or automated, can strengthen railroads' track inspection process by adding redundancy, reducing the likelihood of missed defects, and ignoring false alarms.

---

[7] This is not the only way that UAS can be used in track inspection. Though Federal Aviation Administration regulation stipulates that UAS cannot be operated beyond visual line of sight (Small Unmanned Aircraft Systems 14 CFR Part 107, 2016), some railroads have obtained waivers to operate UAS along the length of their track without direct human supervision. This corresponds to the teaming structures labeled "C" in the figure.

Each of the teaming structures in this section requires different elements to function effectively (e.g., staffing levels, training, automated detection capabilities, communication capabilities, human-machine interface displays). Furthermore, each teaming structure poses implications for the safety and efficiency of inspection (e.g., type and volume of data collected, accuracy of data collection, amount of track time required).

6.2.5 Questions to Explore Teaming Structures for Data Collection

When designing new track inspection technologies or considering how to use them most effectively, consider the following questions related to teaming possibilities for data collection:[8]

- Who (or what) is primarily responsible for data collection?
- Who (or what) is involved in supporting data collection (e.g., monitoring the quality of data, etc.)?
- Where is the human(s) and/or automation involved in data collection located when inspection is taking place?
- When are humans involved in data collection (e.g., throughout the process, periodically when they receive alerts, or after it is completed)?
- How often is data collected (e.g., continuously during revenue service, routinely occurring on a fixed interval, periodically but irregularly, etc.)?
- What type of data is being collected?
- How does the method of data collection affect *analysis* of that data?

Each teaming structure has relative strengths and weaknesses. Decisions around teaming during data collection can influence track inspection system performance, including data collection efficiency, data accuracy and defect detection rates, and workload.

Section 6.3 presents considerations for data collection that railroads and manufacturers may use to decide between teaming structures and develop design requirements. These are not pros and cons, as each consideration can be managed through system design choices.

6.3 Considerations for Data Collection

This section explores interdependencies, challenges, and design requirements related to data collection.

6.3.1 Obtaining Track Time

Many forms of data collection require obtaining track time for the inspection. When making decisions regarding HAT in track inspection, consider whether the teaming structures require track time and how those inspection teams will coordinate with dispatchers.

- **Track time for inspections.** An inspection requiring dedicated track time (e.g., hi-rail or TGMS inspection) affects revenue service. An inspection may lead to delays for revenue

---

[8] For additional questions that can be used to guide design, see Appendix C.

service trains around the inspection because inspection vehicles may move more slowly than revenue service trains and may need to stop to examine possible exceptions.

Railroads may reduce the requirement for track time for inspections by using technologies that collect data in revenue service such as aTGMS. However, the differences in how these systems function may lead to tradeoffs in other areas. For example, there are no onboard staff to verify the data's accuracy during the automated inspection run.

- **Track time for verification and maintenance.** When using technologies that do not require dedicated track time during inspections (e.g., aTGMS or UAS-based inspections), field staff may still need track time to verify that the exceptions those technologies generate are indicative of a track defect, or to repair the track defect. Section 7 further explores the process of verifying exceptions.

6.3.2 Capturing Accurate Location Data and Inspection Parameters

It is important to make sure that data collected using automation is accurate and has sufficient information to trace back to the location where the data was collected. This entails gathering parameters like GPS location, which track(s) the inspection covers, and the track class at that location.

This can be accomplished in several ways:

- A human inspector or technology operator is responsible for entering these parameters.
- An automated system gathers them using GPS or other sensors.
- An automated system gathers these parameters, and a human then reviews their accuracy.

Each method has benefits and limitations and may provide criteria to inform system design.

- **Accuracy of manual data entry.** If an onboard operator is responsible for entering certain parameters, such as which track is currently being inspected, there is a risk that they could enter information incorrectly.

  The design of the user interface is particularly important. A poorly designed interface could increase the risk of errors. There are many user interface design guidelines available to assist in designing data entry systems.

- **Accuracy of automated parameters.** To avoid the risk of human error in collecting track parameters, or in cases where there is no onboard operator, inspection systems may collect these parameters automatically.

  However, this automated parameter collection may also have challenges related to accuracy. Poor GPS quality can make it difficult to find the exceptions that automated systems identify. To address this issue, railroads and manufacturers may make the accuracy of automated parameters a priority during technology development.

  Alternatively, in teaming structures with an onboard operator or analyst, that person may review the parameters for accuracy. This activity will be most effective if the operator is using a well-designed interface.

6.3.3 Sensitivity and Bias of Inspection Technologies

Both sensor sensitivity and bias affect the detection of degraded track conditions. Sensors, whether human or technology-based, have limits on their ability to detect energy in the form of light, sound, etc. Sensitivity can also change with conditions, such as when ambient light changes the ability of a light sensor to detect illumination levels in the dark compared to daylight. The decision for the sensor to indicate an exception reflects a judgment on what is and is not a degraded track condition. This judgement represents a bias by the observer and may change depending upon the conditions in which the judgment is made.

The track detection task can result in two kinds of errors: false alarms and misses. In a false alarm, a judgment is made that a degraded track condition exists when it does not. In a miss, a judgment is made that no degraded track condition is present when one is present.

There is a tradeoff to make in deciding the rate of misses versus the detection of false alarms. Does the technology provide a way for the railroad to manage the rate of missed exceptions compared to the number of false alarms? Missing an exception affects safety while false alarms affect resource allocation.

Sensitivity challenges can be approached in several ways with respect to teaming:

- If the teaming structure relies on redundancy (e.g., using both visual and automated inspection routinely or repeated measurements close in time), it may be more acceptable to increase the miss rate to increase efficiency by maintaining a low false alarm rate, as a later pass of the technology may detect the defect. Teaming may help automation and human inspectors compensate for one another in situations that render human or automated inspections less capable, such as in poor weather.
- If a teaming structure operates with less redundancy (e.g., only using human inspectors for verification), prioritize detection and allow for higher false alarm rates to avoid misses. Be aware of false alarms have on the system, and that reduced redundancy may lead to system brittleness. If automation encounters situations that render it less effective at detection, there may be less opportunity for a human to compensate.

Railroads must determine what level of sensitivity and how much redundancy is appropriate to meet their safety and efficiency goals.

6.3.4 Workload for Track Inspectors and Other Railroad Employees

Automation can affect workload for track inspectors or other railroad employees.

If inspectors are given outputs from automated inspections that help direct them to problem areas, this could allow them to complete inspections more quickly and reduce their workload. However, if inspectors are required to verify exceptions found by automated inspection technologies, and the volume of exceptions increases (as may occur with continuously operating inspection systems like aTGMS), the workload for track inspectors could increase as they are required to verify many potential defects. In this situation, the job of the track inspector changes from detecting exceptions to verifying exceptions.

Increased data volume affects the workloads of track supervisors and data analysts who process and react to the track condition data they receive.

6.3.5 Consistency of Data from Different Technologies

Some characteristics of inspection technology may lead to differences in measurement results. Variations in sensor calibration, vehicle weight, and power type could lead to slight differences in measurements even between technologies that fulfill the same function (e.g., between two different aTGMS vehicles). Just as humans vary in their reliability to detect track exceptions, so too can technology vary in the reliability of the measurement process. Be aware of this measurement issue when comparing measurements over time, (e.g., by planning to use the same inspection vehicle in the same territory when possible or being aware of and accounting for differences in vehicle measurements).

7. Data Analysis: Analyze Track Condition and Identify Problems

After track condition data is collected, humans and/or automated technologies examine the data to identify problems related to track condition. Data analysis cannot occur without data collection. In many cases these tasks occur simultaneously. In other cases, data analysis reveals further questions that require additional data collection to answer, such as the verification of potential defects.

The HAT structures and considerations discussed as part of data collection (see Section 6) exert a significant influence on those discussed in this section on data analysis.

This section addresses steps 1 through 3 of the process for designing HATs as summarized below. Refer to Section 5 for details on these steps.

- **Step 1: Examining track inspection process requirements and system goals.** This section describes key aspects of the data analysis task.
- **Step 2: Exploring roles for humans and automation in track inspection.** This section describes how human-automation interactions involving data analysis can be understood using two frameworks: a LOA framework and a more nuanced HAT framework that considers the possible roles of multiple humans and automated systems.
- **Step 3: Considering tradeoffs and challenges in track inspection HATs.** This section describes considerations for railroads and manufacturers when making decisions about how to allocate data analysis tasks and design effective HATs.

7.1 Understanding the Data Analysis Task

Track inspection data analysis addresses both short term and long term maintenance.

7.1.1 Short-Term Analysis: Identifying and Classifying Conditions

In the short term, track inspection data analysis involves identifying the presence of degraded track conditions and classifying them according to their severity.

In the short term, railroads group degraded conditions into two categories based on severity:

1. **Track safety defects** are degraded conditions that exceed FRA's track safety standards and must be addressed immediately.
2. **Maintenance conditions** are degraded conditions that comply with FRA's track safety standards but exceed railroad thresholds or raise concerns for the inspector. Railroads may choose to address these conditions in the short-term or include them in longer-term planning.

For track safety defects, severity can be further classified for those defects that are "class limiting." If the measurement would *not* be considered a defect at a lower track class, the defect can be described by how much the track class would need to be reduced to eliminate the defect, for example, a "two-class drop."

7.1.2 Long-Term Analysis: Examining Trends and Making Predictions

In the long term, track inspection data analysis involves tasks such as identifying trends and patterns in track conditions, predicting future degradation, or attempting to understand the root causes of recurring track problems.

This type of analysis requires data from multiple inspections. It may draw on data from multiple inspections of the same track over time to monitor for changes. It may also integrate data from several different kinds of inspections to see where problems may be co-located. Identifying co-located conditions can help railroads understand the underlying causes of these conditions and inform maintenance planning.

7.2 Roles of Humans and Automation in Data Analysis

This section discusses some of the roles and capabilities of humans and automation in performing or supporting data analysis tasks. It illustrates the relative strengths that humans and automation bring to the process.

Human inspectors perform data analysis by comparing the track condition information they observed to information stored in memory, such as their experience with past inspections, their training, and the thresholds set by track safety standards. Human data analysts who work with the outputs of automated technologies (e.g., railroad employees or contractors serving the railroads) may use similar cognitive processes to review and manipulate the data outputs.

Technology may contribute to the analysis by using computational algorithms to process information. This may include performing simple computations or comparisons and flagging exceptions that exceed programmed thresholds or using more complex algorithms such as AI and machine learning to recognize patterns.

7.2.1 Roles of Humans in Data Analysis

This section discusses some of the roles and capabilities of humans related to data analysis.

- **Analyzing data to identify degraded track**. As inspectors observe track conditions, they also analyze that data, sometimes integrating many factors simultaneously and sometimes focusing on a single factor. They may note when concurrent factors that are all below threshold for an FRA defect combine to create a problematic condition.
- **Managing data quality.** The role that humans play in managing data quality depends in part on the type of automated system. Some technologies rely on automation to perform the initial data collection and analysis, flagging potential defects or "exceptions," and then use a human analyst to review all the data in real time to manage deviations from desired technology performance (e.g., false alarms, calibration errors, etc.). Often this happens onboard the inspection vehicle, as is the case for some TGMS geometry cars.
    - Alternatively, the role of managing data quality may happen later or in a different location. Some railroads use a "remote desk loop," where an analyst reviews data from automated technologies like aTGMS to make sure it looks reasonable before sending it to field staff for verification and maintenance. In other cases, railroads may not perform any specific data quality checks before sending data to the field.

- **Role of humans in verifying exceptions.** If there is no human involved in managing data quality prior to the data being sent to the field, the field staff, at a minimum, play a role by confirming that a degraded track condition or defect is present at the specified location prior to performing maintenance activities (e.g., repairing or replacing components). If no one else has reviewed the data, as may be the case for some railroads with aTGMS data, the field employees must identify any false alarms or other erroneous data.

  The logistics of these verification activities depend on how the inspection is conducted. For technologies with onboard staff, such as conventional TGMS geometry cars, the vehicle can be stopped, and a human can verify the exception on the spot. For technologies that lack onboard staff, such as aTGMS, verification must occur separately from data collection. In these cases, verification could be performed during separate visual inspections, or could be coupled with performing maintenance activities.

7.2.2 Roles of Technology in Data Analysis

- **Role of technology in analyzing data to identify degraded track conditions.** The type of data that a remote analyst or field staff receive can vary significantly depending on the role that technology plays in analysis. In some cases, analysts may receive raw or processed data that they are required to review or manipulate to check for accuracy and look for problem areas (e.g., strip charts, images of track, or lists of exceptions for review). In other cases, the technology may produce detailed dashboards that require little manipulation.

  These variations in degree of data analysis automation also affect HAT in decision making, as more sophisticated analysis outputs may support decision making by helping humans narrow down their options, as discussed in Section 8.

- **Role of technology in verifying exceptions.** Multiple technologies may also play a role in verifying exceptions in some cases. When a technology identifies flaws that are difficult for a human to detect, the exception may need to be verified using a more accurate version of the technology. For example, exceptions found using vehicle-based ultrasonic technology (e.g., Sperry car) may be verified by handheld versions of this technology. As with TGMS and aTGMS, the timing of this verification may depend on whether there are human staff onboard the vehicle. If not, verification must occur separately from data collection and analysis.

7.2.3 LOAs for Data Analysis

To explore some of the ways that data analysis tasks may be allocated to humans and automation, this section describes task allocation in terms of LOA. Table 4 provides examples of LOAs for data analysis ranging from low to high.

Table 4. Example LOAs for Data Analysis

Approximate LOA	Description	Example
Low	The technology can help the human in comparing and combining information to make it easier for the human to understand the status of track condition.	*A TGMS display that shows multiple measurements on the same strip chart for the onboard analyst to review.*
Medium	The technology can perform analyses as instructed by the human analyst, who sets the parameters for various calculations and comparisons. These technologies support the human by manipulating information in ways that are difficult for a human to do using only mental or paper computations.	*A program that allows an analyst to review track condition data over the railroad's territories and creates reports to show areas that exceed a particular threshold of interest to the analyst.*
High	The technology can perform analyses according to preprogrammed criteria and may trigger alerts if human attention is needed. The human does not necessarily take active part in these analysis activities. At the highest LOAs, a human could use this information to make decisions without taking any active role in the analysis.	*An aTGMS system that notifies the railroad when it identifies an exception or potential defect.*

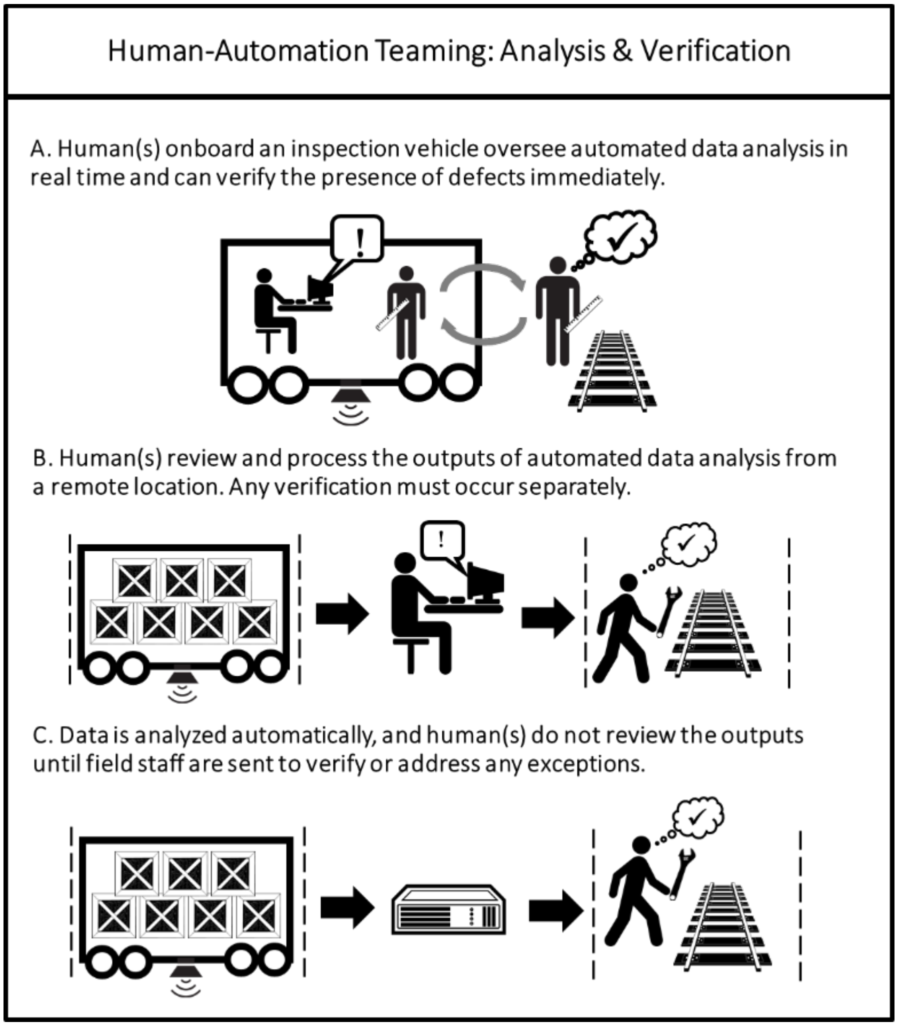
The degree of automation used in the data analysis task can affect human performance for other tasks in the inspection process, depending on the level of human involvement in these later tasks. For example, whether a human has access to a full raw dataset or a more limited list of exceptions may shape how they analyze that data and make decisions later.

As with data collection, while LOA models are useful for explaining the role of technology in data analysis, there are nuances not explored in these models. One example is the role that a human may play in quality control during the analysis task. Track inspection technologies vary in their likelihood of registering "false alarms," (i.e., exceptions that are not the result of an actual degraded track condition) (Al-Nazer, Raslear, Wilson, & Kidd, 2017). For example, a leaf blowing past a sensor could appear as an exception or passing over a frog could cause an erroneous wide gage reading. If the technology involved in analysis is not capable of recognizing such issues, railroads may require a human monitoring or reviewing outputs to find and correct them.

Other factors, such as when and where analysis and data cleaning activities occur, the format of the data output, and user interface design, also have implications for human and system performance.

7.2.4 Possible Teaming Structures for Data Analysis and Verification

Figure 12 depicts possible teaming structures that can be used in data analysis and verification. These examples vary in their LOA and in other characteristics, such as the location of humans and technology and the timing of their activities. This figure focuses on teaming, therefore it does not depict scenarios in which a human performs data analysis in a completely manual way (e.g., where an inspector notices a defect and assesses its severity to make decisions without any automated support).



Human-Automation Teaming: Analysis & Verification

A. Human(s) onboard an inspection vehicle oversee automated data analysis in real time and can verify the presence of defects immediately.

B. Human(s) review and process the outputs of automated data analysis from a remote location. Any verification must occur separately.

C. Data is analyzed automatically, and human(s) do not review the outputs until field staff are sent to verify or address any exceptions.

Figure 12. Examples of teaming structures for data analysis

- **Human(s) onboard an inspection vehicle oversee automated data analysis**. The first teaming structure, labeled “A,” involves humans overseeing and supporting automated data analysis onboard an inspection vehicle, such as a conventional TGMS geometry car. In this type of teaming structure, analysis and verification can be performed together because the analyst is present at the location where data is being collected and is reviewing the analyzed data in real time. The analyst or other onboard staff (e.g., a track supervisor) can exit the vehicle to verify the accuracy of the automated analysis, confirming whether a defect is present at the location. Ultrasonic testing vehicles operate similarly, using onboard staff to inspect certain areas more thoroughly once the automated analysis identifies a potential problem.
- **Human(s) review and process data outputs from a remote location.** The second teaming structure, labeled “B,” involves human(s) reviewing and processing the outputs of automated data analyses from a remote location. For example, aTGMS uses a human analyst to remotely review (e.g., in an office environment) either the full dataset of track measurements or a subset flagged by the automated analysis. This review determines

whether the exceptions identified by aTGMS warrant sending field staff to verify or address the issue.

- **Automated systems analyze data without a dedicated human reviewer.** The third teaming structure, labeled “C,” involves data that is analyzed automatically with little human involvement. Humans do not review these outputs until field staff are sent to repair any defects. These field staff support a primarily automated analysis by verifying that a defect is truly present where the automated system has identified an exception.

These examples of teaming structures illustrate that there are a range of possible interactions between humans and automation in the analysis of track condition data. As with data collection teaming structures (as described in Section 6.2.4), railroads may use several of these teaming structures in different contexts. They may use different teaming structures to analyze different kinds of data, though it may also be important for these analyses to integrate data from multiple sources. Each teaming structure has relative strengths and challenges that should be assessed prior to adopting that teaming structure or automated data analysis tool.

7.2.5 Questions to Explore Teaming Structures for Data Analysis

When designing new track inspection technologies or considering how to use them most effectively, consider the following questions:[9]

- What is the purpose of the analysis (e.g., identifying safety concerns, identifying long-term maintenance needs, or both)?
- What types of track conditions or defects are being analyzed? Are they objective or subjective assessments (i.e., is the severity classification based on a single threshold, or more complex pattern recognition)?
- What kind of data is being analyzed (e.g., data from visual inspection, data from a single automated technology, data from multiple automated technologies, data from several inspections using the same technology, etc.)?
- Does the analysis require gathering any additional data for verification purposes? If so, when and how does this occur (e.g., at the time of inspection and analysis, or later)?
- Who (or what) is primarily responsible for analyzing the data (e.g., track inspectors evaluating defects during visual inspection, automated systems doing the bulk of data processing from an automated data collection run, etc.)?
- Who (or what) is involved in supporting data analysis (e.g., data analysts reviewing analysis outputs, field staff verifying presence of defects, etc.)?
- Where does data analysis take place (e.g., at the location of the inspection, or elsewhere)?
- When does data analysis take place (e.g., at the same time as inspection, or later; all at once, or spread out over time)?
- How many humans are involved in data analysis, and how do they coordinate with one another?

---

[9] For additional questions that can be used to guide design, see Appendix C.

These questions may help define the complexity of the analysis task so that railroads and technology designers may make appropriate decisions to manage that complexity.

7.3 Considerations for Data Analysis and Verification

The way that data is collected has implications for data analysis. Design decisions in one area can have implications for other areas or for the track inspection system. Some of the potential challenges associated with track inspection data analysis and verification are detailed in the remainder of Section 7.3.

7.3.1 Developing Appropriate Levels of Trust

The concept of trust in HATs is equally as important as it is in human teams. The level of trust that is appropriate depends on the capability of the automation, the nature of interactions between the human and the automation, and the allocation of responsibilities.

Both overtrust and lack of trust can contribute to problems. Overtrust in automation may lead to humans overlooking problems with the automated output, while lack of trust can reduce the acceptance of technologies.

- **Overtrust in automation**. As technologies mature and reliability increases, humans tend to develop higher levels of trust in these technologies. For example, railroads may be willing to trust that the output of TGMS and aTGMS are likely to be accurate and may not feel the need to verify the analyses these technologies perform. This is due in part to the maturity of the automation (e.g., geometry detection systems) and familiarity with the types of problems that these technologies detect and the output that they produce.

  There is a risk that if humans come to trust automated technology output *too much*, they will rely on the automated analysis even when it contradicts their normal judgment. Generally, high levels of trust develop when technology is highly accurate, which makes humans less likely to respond when infrequent mistakes occur.

  Additionally, humans who have taken on a monitoring role may not be accustomed to reacting to problems in the same way as those who have had to take a more active role in analysis. Over time, analysts' ability to recognize problems with the accuracy of data output may diminish as such errors become less frequent and trust in the technology grows.

- **Insufficient trust in automation.** With newer technologies, such as machine vision technologies that use algorithms to flag locations with potential defects, trust in the technology may be lower. The algorithms may not provide transparency in how judgments are being made. This lack of clarity may contribute to mistrust in the automation.

  Technologies that use neural networks and deep learning methods for pattern recognition may be more difficult for humans to trust, as the way these methods work are not always clear to humans. The lack of transparency in how these methods operate mean that the human does not have an accurate mental model of what the system is doing.

  While one implication of reduced trust may be the decision to have humans review data and manage false positives and missed exceptions, if trust is *too* low, humans will be inclined to disregard the output and the technology will no longer add value to the track inspection process.

Therefore, it is important for humans to understand both the strengths and limitations of any automated technology they are working with so they can calibrate appropriate levels of trust.

7.3.2 Developing Automated Capabilities for Data Analysis

Just as with human limitations, technology limitations can compromise the ability to find and address individual defects. One such issue is working with inaccurate geolocation data. Accurate GPS data is critical to enable field staff to locate exceptions from automated inspections. Some automated technologies have experienced challenges associated with collecting accurate GPS data. Without accurate location information it is difficult to find and verify exceptions, as well as to align data from multiple passes over the same location.

Another challenge occurs when track conditions and defects are difficult to classify. For example, inspectors must sometimes make qualitative assessments that require classifying track components (such as ties) as "good" or "bad," without a quantitative threshold for classification. While technologies that use machine vision and pattern recognition may be able to automate these assessments, detecting such conditions presents unique challenges because humans may disagree about whether the technology is accurate, unlike systems that rely on comparing measurements to quantitative thresholds. One subject matter expert the team interviewed described challenges in determining the accuracy of an automated system for assessing tie degradation, given that human inspectors may have varying opinions about how to classify certain ties.

Other challenges limit the ability to monitor the track condition and identify trends. For example, data collected across multiple inspections may not always be consistent, particularly if those inspections used different automated systems, occurred under different lighting conditions, or if the data collected is not sufficiently accurate (like the geolocation example described above).

Development of highly accurate automated analysis capabilities for examining trends and making predictions, rather than simply identifying individual defects, can also be challenging. Currently, railroads primarily use regression analyses to examine trends, which cannot always provide accurate assessments for all types of track conditions. Researchers are working on developing methods with greater accuracy for a wider range of defect types using AI and machine learning. However, there are some challenges in implementing these analysis techniques, including the need for large, high-quality datasets for training the AI.

7.3.3 Volume of Data Generated by Automated Technologies

In general, the more information available about track conditions, the better equipped a railroad is to manage safety and make effective maintenance decisions. However, as the number of automated data sources that railroads use increases, the amount of data collected and railroad employees must manage also increases.

In a typical visual inspection, the human primarily documents degraded track conditions. As the use of track inspection automation increases, railroads have access to data regarding track conditions for the entire length of track, not only areas with degraded track conditions or exceptions. Many automated inspection technologies can collect data continuously during revenue service, increasing the frequency of data collection. The result is a much larger data set that poses both opportunities and challenges to the track inspection process.

Whether in real-time or following the collection of track inspection data, the data is analyzed to detect exceptions and potential defects. Without effective interfaces and tools to manage this data, the volume of information may easily become overwhelming to the analysts or those responsible for prioritizing track maintenance activities. The opportunities and challenges depend upon whether this larger quantity of data is analyzed by humans, automation, or some combination of the two. Some of this data may be in a raw form, while other data may be processed in a form that is specifically designed for humans to use in decision making.

Several considerations related to managing the volume of track inspection data are:

- **The need to integrate data across multiple sources.** Whether data is collected by a single inspection vehicle or from multiple automated technologies, there may be a need to compare data across technologies to examine areas where different types of track conditions co-occur.

  This data integration could occur in several ways:

  - **Manual data integration.** A human analyst or supervisor could comb through multiple reports and look back and forth between them to make comparisons, perhaps manipulating the data manually to create an integrated report. This process would likely be time-consuming and prone to error.
  - **Automated data integration.** The data integration process could be automated, and information would be presented to the human analyst or supervisor in a user-friendly format. In this case, the design of the human-machine interface would be important, since it would shape how the human perceives and interprets the data. Such a system could be highly beneficial in supporting human decision making, but it could also have negative consequences if it is not well-designed.

  Figure 13 illustrates several ways that data integration could occur and then be disseminated to field staff for use. Note that without data integration, field staff will need to handle outputs of multiple technologies simultaneously.

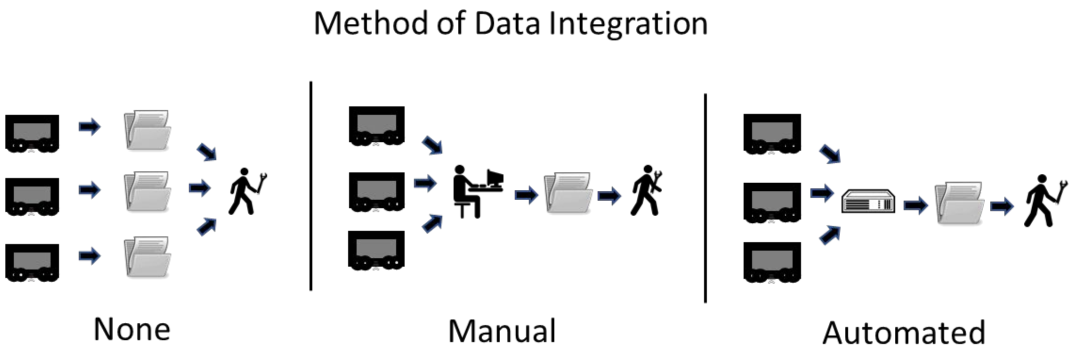


Figure 13. Approaches for integrating data and delivering it to field staff

- **The need to review data with reduced context.** Teaming structures that use a remotely located analyst or inspector can pose a challenge in that this employee would not have the same context that they would have if riding onboard an inspection vehicle or performing a standard visual inspection (via walking the tracks or hi-rail).

  For remotely located analysts, it may be harder to understand the "big picture" for that area of track and understand the conditions in context, especially if there is a large volume of data and the human only looks at the portion of the data flagged as exceptions. However, a greater level of detail could be overwhelming and may make it difficult to review all the necessary information in a reasonable time.

  The amount of data that automated collection can generate may make it impractical for humans to be fully responsible for data analysis. This volume of data creates a constraint that may shape design decisions about how to analyze the data, such as the number of analysts required and the level of detail that those analysts receive.

- **The need to determine the best way to distribute data analysis tasks.** Another consideration is how tasks are distributed among data analysts involved in managing the output of multiple technologies. Railroads will need to assess how best to staff and organize their data analysis process depending on their specific resources and requirements.

  One option would involve multiple data analysts responsible for analyzing data from distinct automated technologies. Railroads may establish this type of teaming structure if using technologies that already include an onboard analyst, such as a conventional TGMS geometry car, or for which the analysis requires specialized expertise. However, the downside of analyzing each dataset separately is that this teaming structure may require significant coordination between these employees to establish a global picture of the track condition.

  Another option involves one analyst, or team of analysts, examining data across all the technologies the railroad uses. If the way that the system presents data requires specialized knowledge and training to understand, this may limit the number of humans who are able to support data analysis to those specifically trained as data analysts. During interviews with railroads, the team noted that relying on a single person or limited number of analysts could create significant challenges if an analyst leaves the railroad, becomes ill, or simply goes on vacation. If only a few people can make use of the data, this can create a bottleneck or constraint.

7.3.4 Challenges Associated with Verification

The process of following up on track safety issues can have its own challenges, including: (1) preparing information for field staff and (2) ensuring that field staff maintain the knowledge and skills needed to perform this task.

- **Preparing information for field staff.** Regardless of the degree of data analysis automation, information from the data analysis process must eventually be passed on to field staff (inspectors and/or maintenance crews) for verification and/or maintenance.

Field staff may also deal with challenges related to the volume of data generated by multiple automated inspection technologies. If data from different systems are presented as separate reports or outputs, rather than a single consolidated track status report, these field staff may struggle to keep track of this information. They may need to review multiple reports to see where on the track to look for defects of different types, even if those defects are in the same location. This could make it more likely that they will overlook something or need to backtrack to revisit something they previously missed.

Depending on the type of data output that any automated analysis produces, field staff may experience challenges in interpreting this data. Field staff may not have training or experience in reading and interpreting strip charts and other automated data outputs. Currently, some railroads have a supervisor or data analyst prepare information for inspectors so that they know what to look for and where to find the potential defect. This requires a significant amount of work that may only increase as railroads increase their use of automated technologies.

An alternative would be if the automated analysis tools were able to integrate information into a user-friendly interface for field staff. However, giving this data to field staff directly will create new problems if the data output and interface are not well-designed.

- **Maintaining adequate knowledge and skills for analyzing severity and verifying defects.** Another challenge posed by increased automation is maintaining adequate knowledge and skills for those railroad employees responsible for analyzing track condition data and verifying defects (i.e., inspectors and maintenance crews).

  Many of the skills that track inspectors and foremen use to recognize degraded track conditions are obtained through hands-on field experience over many years. Though automation may in some cases make their jobs easier and reduce the workload required to identify defects, it may also lead to deskilling, or reduction of certain skills if those skills are no longer routinely practiced.

  Particularly if field staff conduct a reduced number of visual inspections, it may be more difficult to maintain a team of experienced staff. Practices such as on-the-job training and mentorship may also be affected. These important ways of passing on knowledge rely on more experienced staff available to pass on their expertise, which could be more challenging in an environment with fewer hands-on inspections**.**

  One possible impact of increasing the use of automation or reducing the number of visual inspections is that inspectors could be expected to cover a larger territory when looking for track degradation or verifying exceptions from automated systems. Currently, inspectors are required to maintain an intimate knowledge of the territory they inspect and its features so that they can recognize changes and degradation when they occur. If inspectors are required to cover a larger territory, or are assigned to inspect less frequently, it may be challenging for them to maintain the degree of familiarity that is currently required.

  Even if the number of inspectors and frequency of inspections remains the same, challenges in maintaining inspectors' skills could occur if inspections are more heavily guided by automation and inspectors no longer need to rely on their experience. (This is closely related to the issue of overtrust described earlier.)

Deskilling can be combatted through periodic retraining or refresher training. However, in an environment that relies more heavily on automation, these efforts may need to be deliberately undertaken to ensure that field staff possess the required knowledge and skills. Railroads will need to evaluate whether their training practices for inspectors are sufficient to address the job demands associated with a more heavily automated inspection process.

8. Decision Making and Action

Once humans and/or technology analyze track condition data to identify degraded conditions, they decide on and take any appropriate actions. The tasks of decision making and action are closely related, and the railroad industry currently uses little automation for these tasks, relative to data collection and analysis. This section presents a combined discussion of decision making and action.

This section addresses steps 1 through 3 of the general process for designing HATs as summarized below. Refer to Section 5 for details on these steps.

- **Step 1: Examining track inspection process requirements and system goals.** This section describes key aspects of the decision making and action tasks.
- **Step 2: Exploring roles for humans and automation in track inspection.** This section describes how human-automation interactions involving decision making and action can be understood using two frameworks: a LOA framework and a more nuanced HAT framework that considers the possible roles of multiple humans and automated systems.
- **Step 3: Considering tradeoffs and challenges in track inspection HATs.** This section describes considerations for railroads and manufacturers when making decisions about how to allocate decision making and action tasks and design effective HATs.

8.1 Understanding Decision Making and Action Tasks

Decisions and actions associated with the track inspection process may be short-term or long-term. Short-term decisions and actions depend on the severity of the track condition, while long-term decisions and actions involve maintenance planning.

8.1.1 Short-Term Decisions and Actions

If a degraded track condition exceeds an FRA safety standard, the severity of the defect informs decision making. The possible decision alternatives are determined by regulation, which specifies a timeframe within which the defect must be addressed. Appropriate actions to take after identifying a defect may include repairing the defect, restricting the track speed, or removing the track from service. These short-term decisions tend to involve unplanned maintenance, as discussed in Section 2.2.

If the track condition does not exceed FRA safety standards, the railroad has greater discretion in when and how to act and can decide to defer actions until later, during planned maintenance.

Railroad employees also keep records of the track conditions and any actions they perform.

8.1.2 Long-Term Decisions and Actions

In the longer term, railroads make decisions related to maintenance planning and large-scale maintenance projects, including proactive and preventative maintenance. This is an opportunity to address documented track conditions that do not exceed FRA safety standards in more efficient ways.

The railroad industry would like to increase its focus on this type of long-term decision making (Shukla, 2021). This is the shift from “find and fix” to “predict and prevent” approaches to track maintenance described in Section 2. Although addressing defects promptly remains important, the goal is to reduce the number of defects that require immediate action through preventative measures so that railroads can act proactively rather than reactively in their decision making.

8.2 Roles of Humans and Automation in Decision Making and Action

The following sections discuss the possible roles of both humans and automation in decision making and action tasks in the track inspection process.

8.2.1 Roles of Humans in Decision Making and Action

Typically, railroad employees make decisions about how to address track conditions. These decision makers include track inspectors, track supervisors, technology operators, analysts, and maintenance way crews, depending on how the inspection is conducted and how automation is used.

Humans are accountable to others for their decisions and their consequences (Billings, 1997). Subject matter experts indicated that railroads typically value having a human as the final decision maker, even when they rely heavily on automation for data collection and analysis. Humans add value by verifying that the analysis and suggestions from the automated system make sense and by making decisions about how to allocate funding and resources as part of maintenance planning.

8.2.2 Roles of Automation in Decision Making and Action

Automation can support decision making and action in a variety of ways, both in determining how to address a particular defect and in making larger-scale decisions related to maintenance planning.[10]

- **Supporting short-term decisions and actions.** Automation can assist with decisions and actions related to addressing defects (i.e., reactive or unplanned maintenance) by helping humans classify defect severity and select appropriate actions. For example, TGMS may indicate the severity of a defect (e.g., a "two-class drop," which suggests that reducing the track class for that section by two classes would address the defect) and aids the TGMS operator in making decisions.

  This decision making may also depend on the role of automation in data collection and analysis:

  - Automation of the data collection task may aid decision making by presenting the human with more information than they would have from visual inspection alone. For example, x-ray-based systems provide data about internal rail defects that humans cannot see, enabling them to decide when it is necessary to act.
  - Automation of the data analysis task may make information easier for humans to understand, particularly if the automation helps manage the level of detail so that the human is not overwhelmed without hiding or obscuring information. For

---

[10] Note that "decision making" for automation is distinct from human decision making. It does not imply a requirement for consciousness on the part of the automation. In this case it refers simply to selecting a viable course of action from a set of possible alternatives.

example, a system that allows an analyst to move easily between strip charts and images of the track at a given location may help them understand the information better than if they had to manually compare the two data sources.

Action may include contacting the dispatcher to impose speed restrictions or remove track from service, performing repairs or other maintenance activities, and logging or documenting the track condition and other actions taken. Most of these tasks are performed by the same humans responsible for decision making; it may also be possible for automation to support taking action, though the team is not aware of current systems that use a higher LOA for the action task (e.g., recommending a specific course of action or taking actions without human involvement.)

- **Supporting long-term decisions and actions.** Automation may be particularly valuable in supporting longer-term decisions regarding maintenance planning. One of the challenges associated with using automation in data collection is the increased volume of data that railroads must handle in data analysis and decision making. This problem is magnified when that data comes from multiple automated technologies. Automation can help integrate data from these multiple sources and can use that integrated data to help humans make better decisions.

  Automated data analysis and decision support can be used to look at trends and model when track parameters are likely to exceed some limit. A succinct report of these predicted track degradations help railroads make strategic maintenance decisions. Automation could also help railroads identify underlying causes of certain defects so they can find more effective long-term solutions, rather than continuing to make the same types of repairs as inspections identify them.

  Railroads may also use decision support automation in the form of maintenance planning tools that explore possible courses of action and visualize the impacts of those decisions.

8.2.3 LOA for Decision Making and Action

This section describes LOAs for deciding on and taking any actions needed to address degraded track conditions. Table 5 explores some of the ways that decision making may be allocated to humans and automation by describing several examples ranging from low to high LOAs.

Table 5. Example LOAs for Decision Making

Approximate LOA	Description	Example
Low	The technology may propose a set of decision alternatives for a human, and potentially narrow down possible options.	*A TGMS system that indicates whether a geometry exception requires a one-class or two-class drop, which narrows the options available to the human operator.*
Medium	The technology may propose a single option, and in some cases, may be able to execute that action with the human approves.	*A hypothetical technology could inform the track supervisor that a speed restriction is needed, and with the supervisors' approval, send a message to the dispatcher.*
High	The technology decides on an appropriate action without human input.	*A hypothetical technology could determine that a speed restriction is needed without any human involvement.*

Table 6 provides examples of LOAs for action ranging from low to high. Except the first, these examples are hypothetical, as there is not currently a significant focus on automating these types of actions in the track inspection process.

Table 6. Example LOAs for Action

Approximate LOA	Description	Example
Low	A human performs all actions mostly manually or with the use of simple tools.	*An inspector calls the dispatcher to impose a speed restriction or uses simple tools to repair a defect.*
Medium	The technology may perform actions automatically once directed or approved by a human.	*A hypothetical technology that, with an analyst or supervisor's approval, can send a message to the dispatcher to impose a speed restriction.*
High	The technology may perform actions while a human supervises and is able to intervene if needed. At the highest levels of action automation, the technology acts without any human involvement.	*A hypothetical technology that takes actions, such as contacting the dispatcher to impose a speed restriction, without any track inspector or technology operator instructing it to do so.*

8.2.4 Possible Teaming Structures for Decision Making and Action

Figure 14 and the text that follows describe a range of possible teaming structures that can be used in decision making and action. This description is neither a formal classification system nor a comprehensive list of possibilities. It provides a simple depiction of the range of possible interactions between humans and automation for these tasks.

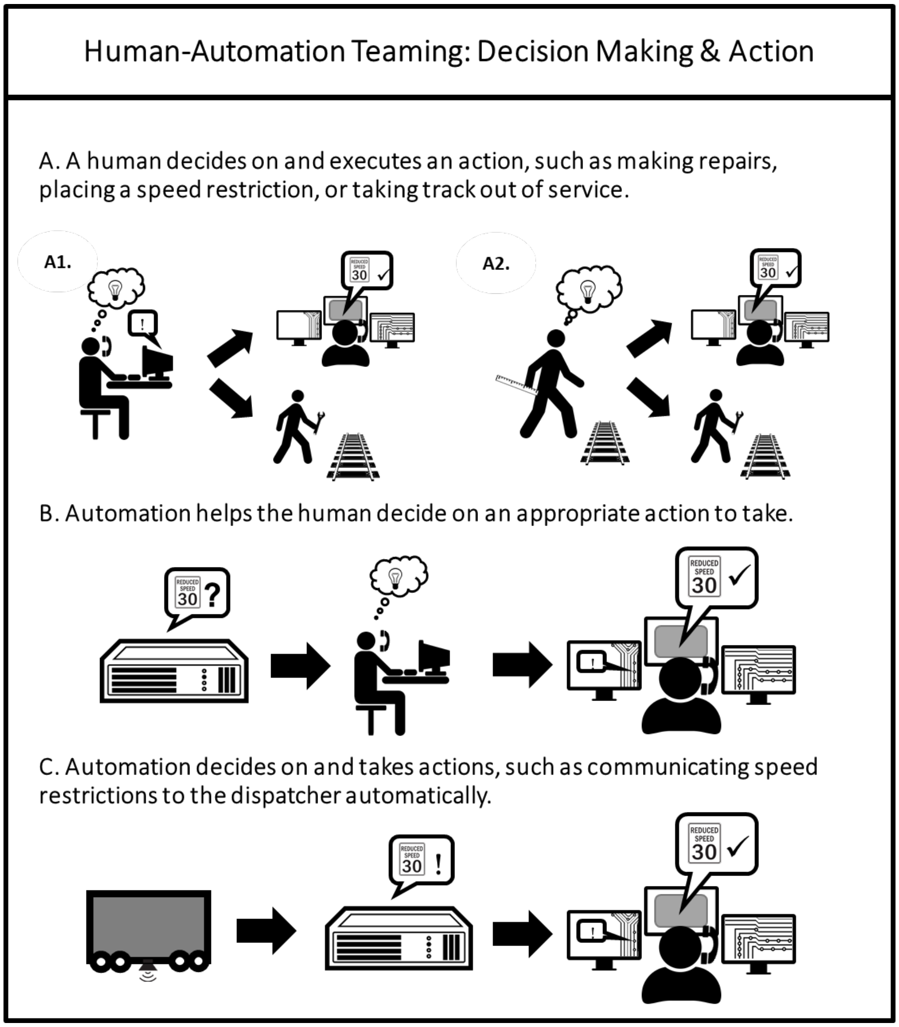


Figure 14. Examples of teaming structures for decision making and action

- **A human decides on and executes an action.** In the first set of teaming structures, labeled "A," a human decides on and executes an action.

  In example A1, a supervisor or data analyst makes a decision based on data from an automated system. They may decide to impose a speed restriction or take the track out of service by calling the dispatcher, or to order a maintenance crew to repair or replace the track component.

  In example A2, the person making the decision is the track inspector, who is either noticing a defect for the first time or verifying an exception from an automated system. They may decide to impose a speed restriction or take the track out of service by calling the dispatcher, or they may choose to repair or replace the track component themselves if they have the time and resources to do so.

  Both teaming structures use a similarly low LOA according to Table 5 and Table 6, but differ in who is making the decision and where they are located, which in turn affects the types of action available to them (e.g., the inspector can perform repairs on the spot, while the analyst cannot).

- **Automation helps the human decide on an appropriate action.** In the second teaming structure, labeled "B," automation helps the human decide on an appropriate action. In this case, a hypothetical technology suggests to the track supervisor or data analyst that a speed restriction would be an appropriate mitigation for a particular defect. The supervisor or analyst then contacts the dispatcher.

  At a *slightly* higher LOA, the technology could contact the dispatcher with approval from the supervisor. This is consistent with the example for medium automation in Table 5 and Table 6.

- **Automation decides on and takes the appropriate action.** In the third teaming structure, labeled "C," automation decides on and takes the appropriate action. In this case, a hypothetical technology takes actions, such as contacting the dispatcher to impose a speed restriction, without any track inspector or technology operator instructing it to do so. This is akin to the high LOA example in Table 5 and Table 6.

These examples of teaming structures illustrates that there are a range of possible interactions between human and automation in decision making and action. Each teaming structure has relative strengths and challenges that should be assessed prior to adopting it.

8.2.5 Questions to Explore Teaming Structures for Decision Making and Action

The following questions may be considered when exploring different teaming possibilities related to decision making and action:[11]

Decision Making Questions

- What type of decision is being made (e.g., determining appropriate short-term actions or planning for long term maintenance needs)?
- What types of track conditions or defects are included in this decision making? Are they objective or subjective assessments?
- What kind of data sources informed the analysis on which these decisions are based (e.g., data from visual inspection, data from a single automated technology, data from multiple automated technologies, data from several inspections using the same technology, etc.)?
- Who (or what) is primarily responsible for decision making (e.g., track inspectors, other field staff, supervisors, data analysts, or an automated system)?
- Who (or what) is involved in supporting data analysis (e.g., automated analysis tools, decision support tools such as automation that narrows or proposes options, or other humans)?
- Where does decision making take place (e.g., at the location of the inspection, or elsewhere)?
- When does decision making take place (e.g., at or near the same time as inspection, or later)?

---

[11] For additional questions that can be used to guide design, see Appendix C.

- How many humans are involved in decision making, and how do they coordinate with one another?

Action Questions

- What type of action is being taken (e.g., maintenance or repair of track infrastructure vs. communication with dispatcher to impose a speed restriction or remove track from service)?
- Who (or what) is primarily responsible for this action (e.g., track inspectors, other field staff, supervisors, data analysts, or an automated system)?
- Who (or what) is involved in supporting data analysis (e.g., automated technologies for performing maintenance, automated communications with dispatch, or additional humans to support action execution)?
- Where does the action take place (e.g., at the location of the inspection, or elsewhere)?
- When does action take place (e.g., at or near the same time as inspection, or later)?
- How many humans are involved in the action, and how do they coordinate with one another?

These questions reveal the complexities of the decision making and action tasks that railroads and technology designers may wish to consider when designing HATs.

8.3 Considerations for Decision Making and Action

Many of the considerations for decision making and action are closely related to those for data analysis. For example, trust in automation is an important factor, as human decision makers could overly rely on automated decision aids or lack trust in these technologies and ignore their recommendations. Similarly, maintaining the skill level necessary to make good decisions could be a challenge if humans have a reduced role in analysis and decision making in the future. Although the ultimate responsibility may remain with the human, if they are responsible for making decisions without being sufficiently "in the loop," they may struggle to make good decisions. Two considerations specific to decision making and action are responding to an increasing volume of data and addressing root causes rather than symptoms.

8.3.1 Responding to an Increasing Volume of Data

Railroads may find it challenging to keep up with the volume of data generated by automated technology. In some cases, this data may include a greater number of exceptions than railroads are prepared to verify and address, particularly if the railroad has set stricter maintenance standards than required by regulation. Over time, the rate of new exceptions would likely decrease as the railroad addresses them, but during initial implementation of new technologies it may be challenging to address them within the time frame required by regulations or railroad policy.

8.3.2 Addressing Root Causes, Rather Than Symptoms

A major challenge associated with decision making is determining how railroads know whether they are taking the appropriate actions to prevent a defect from recurring. In some cases, railroads may see the same issues appear multiple times and address them as they occur, without realizing that there is some underlying cause triggering the degradation. Until the underlying cause is identified and addressed, the issue may return.

A HAT approach to inspection may make it easier for railroads to find and address these underlying causes by optimizing the joint performance of humans and automation and making use of the strengths of each. Automation can support finding underlying causes of defects by increasing the volume of data available, including the potential for collecting and comparing data about how issues at the same location change over the course of several automated inspection runs. Automation can also make it easier to synthesize and analyze that data. The humans involved can do what no automation can, that is derive meaning from the data and interpret the underlying causes based on their experience and expertise.

9. Overall Considerations for HAT

This section discusses the interdependence between the tasks described in the previous sections, as well as overarching considerations for HAT in track inspection, particularly as the use of automation in the industry increases.

9.1 Importance of Interdependence in Track Inspection Tasks

The previous sections discussed how HATs can be composed for each track inspection task (i.e., data collection, data analysis, decision making, and action). However, considering these tasks in isolation is not sufficient to design an effective track inspection program that accounts for HAT considerations. As noted in Section 3, examining the interdependencies between related tasks and between human and machine team members is an important aspect of designing HATs. Figure 15 depicts some of the different humans and technologies that interact with one another throughout the tasks of data collection, data analysis, decision making, and action.

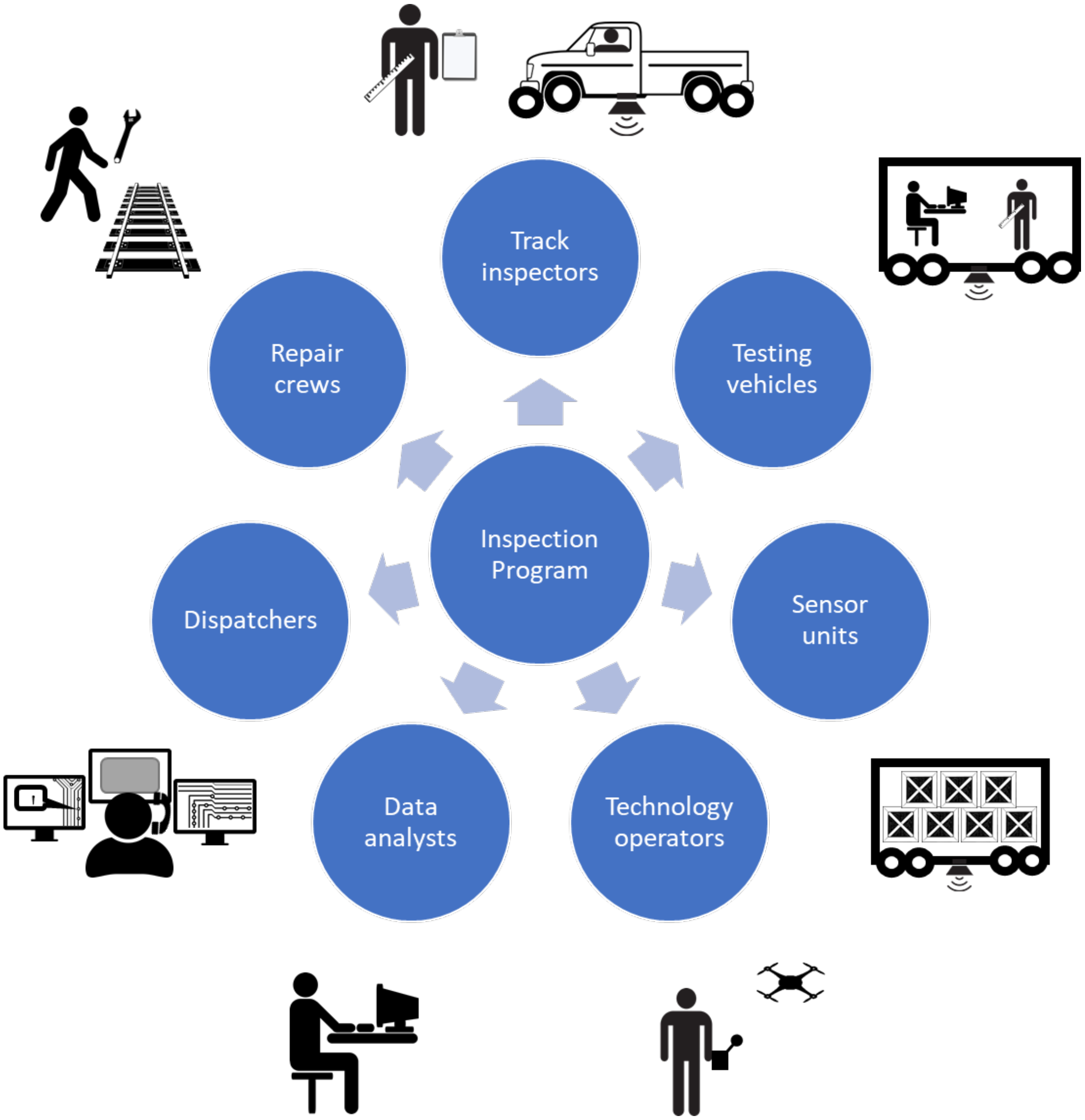


Figure 15. Interdependence in track inspection

When considering how tasks are interrelated, examine the overarching context of track inspection, in addition to the four inspection tasks that this report discusses in detail. Factors such as staffing and training, procurement of new technologies, allocation of track time, and other aspects of railroad operations affect HAT in the inspection process.

9.2 Design Considerations Across Track Inspection Tasks

Each teaming structure has relative strengths and weaknesses that should be assessed prior to adopting that teaming structure or any new automated data analysis method. Some considerations are summarized in Table 7.

Table 7. Relevant Design Considerations for Teaming Structures by Inspection Task

Track Inspection Task	Type of Teaming Structure	Examples of Relevant Design Considerations
Data Collection	Human primarily responsible (either walking or using hi-rail vehicle)	• Sensitivity of human inspectors to defects of different types • Obtaining track time / coordinating with dispatcher
	Automation with human(s) present to support and/or monitor data collection	• Sensitivity of technology to various defects • Data consistency / accuracy of location data • Design of user interface for managing data collection • Obtaining track time / coordinating with dispatcher
	Automation without a human present during data collection	• Sensitivity of technology to various defects • Data consistency / accuracy of location data • Design of data outputs • Workload of verifying defects at a later time
Data Analysis	Human performs data analysis in real time (e.g., TGMS)	• Level of trust • Interface design / design of data output • Maintaining skills for verification
	Humans review and process data remotely (e.g., aTGMS with remote desk loop)	• Integrating data from multiple sources • Level of trust • Interface design / design of data output • Preparing information for field staff • Workload of verifying defects at a later time • Maintaining skills for verification
	Humans sent to verify and address exceptions without prior review (e.g., aTGMS without remote desk loop)	• Integrating data from multiple sources • Level of trust • Design of data output for field staff • Workload of verifying defects at a later time • Maintaining skills for verification
Decision Making and Action	A human decides on and executes an action	• Workload • Communications
	Automation helps the human decide on an appropriate action	• Technical specifications / automation capability • Interface design for decision support • Trust / overreliance • Maintaining skills for decision making
	Automation decides on and takes the appropriate action	• Technical specifications / automation capability

9.3 Overall Considerations Associated with Increased Use of Automation

This report discussed different considerations related to HAT in track inspection for the four inspection tasks: data collection, data analysis, decision making, and action. Themes that are common include:

- Human-centered technology and interface design
- Maintaining skills and expertise of human team members
- Communication and coordination between team members (both human and automated)
- Balancing long-term and short-term priorities

These themes are particularly important to consider as railroads increase their use of automation in track inspection.

9.3.1 Human-Centered Technology and Interface Design

One of the most important influences on the success of HAT is the design of the technology and any human-machine interfaces. Shifting from a technology-centered design approach to a human-centered or systems perspective when developing these technologies and interfaces is a critical step toward future inspection processes that integrate the strengths of both humans and automation to maximize safety and efficiency.

Designing the technology without also considering the human role and interaction is likely to lead to a suboptimal system design. Considering the design of the parts in the absence of the interactions of those parts can contribute to a system that operates less efficiently and less safely than a system that was designed to consider these interactions. The emergence of new properties occurs at the level of the joint human-automation system or HAT. Taking these new properties into account can only be done by considering interactions between the human and the automation.

- **Technology Development.** Human-centered technology development means examining during the development process where technology can add the greatest value and how it will interact with humans during the inspection process (e.g., not just track inspectors, but also operators and data analysts, maintenance workers, managers, and dispatchers). This approach also includes when and how the technology will be used and how the data will be integrated with data from other inspections. When developing or procuring new technologies, railroads and other stakeholders can promote a HAT approach by including these considerations in their initial requirements.

  There are many technical factors that affect HAT but are beyond the scope of this report to discuss in detail. These include the sensitivity of the inspection technology to detect track defects, the ability to collect geolocation data and other parameters to compare data outputs across runs, and the process of calibrating the technology. Such factors should be considered when designing HATs. Humans may need to help compensate for any technology limitations.

- **Interface Design.** Interface design should be included in the technology development process. Controls and display interfaces mediate how humans interact with automated technologies and the data they generate during data collection, data analysis, decision making, and action. The design of these interfaces can help humans manage complexity by making it easy to gather and interpret track condition data and make decisions based on that data. A poorly designed interface can make these tasks more difficult or error prone.

The importance of good interface design becomes more important as the use of automation in track inspection increases, leading to an increased volume of data for humans to manage. Interface designers must determine the appropriate level of information to display. Too much can be overwhelming, while too little may make it difficult for operators or analysts to make informed decisions. For example, a dataset could be presented as a table of raw values, a table with highlighted rows, a plot or strip chart showing variation in measurements, or a set of notifications showing only exceptions. Each of these has different implications for the human user. It may be useful to use alerts or highlighting to draw the user's attention to particular areas, while making additional details available as needed.

A designer that considers how track inspectors process information and develops interfaces that are consistent with that process will enable more effective system performance. For example, an image highlighting a missing bolt may be immediately understood by a track inspector, while a strip chart that presents gage measurements and other variables may not be as easily understood because it is a more abstract representation of the track condition. For outputs that do not immediately match the users' mental model, additional training may be required.

If an interface is confusing or poorly understood, this makes it more difficult for analysts to access the information they need and ensure that they make appropriate safety decisions. Clear, easily understood data outputs and user interfaces, particularly those that are consistent with how the user thinks about the track condition, can help analysts and track inspectors perform their jobs effectively and readily act on the information generated by automated technologies. Good interface design can also support efficiency.

9.3.2 Training and Experience

Maintaining skills and expertise among human team members is critical to effective system performance. Track safety standards mandate regular visual inspections and track inspectors maintain their skills through routine inspections. The introduction of additional automation may change the role of human inspectors and introduce new training requirements.

- **Mitigating deskilling due to increased automation.** Increasing the use of automation could potentially change the role of inspectors in a range of ways. Inspectors may:
    - Continue to inspect regularly, but with increased support from automation to help them focus on suspected problem areas and identify defects
    - Perform inspections in a way that is more heavily focused on verification of what the technology finds, rather than bearing the primary responsibility for detection
    - See larger changes, such as inspecting at a reduced frequency or reviewing track condition data remotely (*if permitted by waivers or future regulation*)

  Any of these changes could make it more difficult for inspectors to maintain the skills and expertise that they need to find and recognize defects. If inspectors no longer use these skills at the same frequency, they may come to rely more heavily on the technology. For certain challenging or uncommon types of defects, it may be especially difficult to maintain the skills necessary to make final decisions about required actions.

Additionally, changes to the inspectors' role or frequency of inspection could lead to other changes, such as increased territory sizes, which could further increase the challenges of maintaining expertise. Inspectors currently maintain comprehensive knowledge of their territories through regular inspections. This domain expertise may be compromised if inspection frequency drops below the level required to build and maintain that knowledge because of increased automation.

Skill degradation may make it more difficult for the human to assess whether the automation is performing properly. If track inspectors, data analysts and others serve as a backstop to the automation, they require the skills to determine when the automation fails.

One way to mitigate deskilling is through refresher training. If railroads shift the role of inspectors and deskilling becomes a risk, they may have to spend more time and money on refresher training to maintain inspector skills. Railroads depend on current or former craft employees to act as trainers, Employees learn job skills through on-the-job training that are difficult to teach in a classroom. If track inspectors take on a reduced role in data collection, it may become more challenging to pass on requisite knowledge and skills as the number of experienced "expert" inspectors who can act as trainers and mentors decreases.

It is important for railroads to recognize where human expertise and hands-on knowledge play essential roles in the inspection process. If they plan to rely on such expertise to validate the data provided by automation, it will be essential to find ways to keep inspectors' skills up to date.

- **New training requirements created by automation.** Railroads may also need to foster new capabilities related specifically to the use of automation. For example, if track inspectors are required to read and interpret the outputs of automated technologies to verify defects, they may need training to understand outputs such as strip charts that may not be part of their existing skillsets. Technology operators and data analysts also require training for their respective roles and will need to understand the technologies they work with, especially when a railroad adopts a new technology.

  Some of this training requirement may be mitigated by clear, easily understood displays and interfaces, but training could be beneficial for complex data outputs or those that do not match what employees are used to. Training on interpreting data may also help inspectors find and verify exceptions more quickly, so that railroads can make sure they are getting the most out of the automated technology.

  Training can also help inspectors and other employees understand the capabilities and limitations of the automation, helping them develop an appropriate level of trust and understand why what they are seeing does or does not align with the automated output. For example, it may be useful for inspectors to understand whether data was collected in a loaded or unloaded condition, as wide gage readings collected under a loaded condition may not be detectable during a walking inspection. Railroads should update their training to reflect evolving technologies.

9.3.3 Communication and Coordination Between Team Members

The third theme that emerges across the four track inspection tasks is communication and coordination. Coordination is central to HAT, and includes coordination between humans and automation, as well as human-to-human and automation-to-automation coordination.

Some examples of communication coordination include:

- Communication between automated data collection systems and remote data analysts, and between data analysts and field staff sent to verify any exceptions
- Communication and coordination between multiple automated systems to integrate data outputs to make those outputs more efficient for field staff to verify
- Communication and coordination between technology operators or track inspectors and dispatchers to obtain track time or to impose speed restrictions

Communication and coordination are essential to making sure that from data collection to data analysis to decision making and action, each human or automated technology in the track inspection system has sufficient context to perform their tasks, particularly if that task is occurring in a different location or at a different time than the previous task.

As railroads increase the use of automation, particularly forms of automation that collect or analyze data without humans present, the need to coordinate across locations will only grow. This will require communication infrastructure to transmit data, as well as to facilitate communication between employees by various methods (e.g., phone, radio, email).

9.3.4 Supporting Both Short-Term and Long-Term Goals

When considering where automation can add value, assess which maintenance decisions are the most difficult for humans and whether automation could provide better data or analytical support to aid in those decisions.

Similarly, railroads, technology manufacturers, and researchers may prioritize automation development for the defects that are the most difficult for humans to detect, or that are not readily detected by other, existing technologies. In some cases, it may be valuable to develop automation to detect and classify conditions that require qualitative assessments to reduce the need for humans to make subjective judgements. However, such capabilities may be challenging to automate and require the use of more sophisticated analysis tools like machine learning and AI.

Prior to adopting any new track inspection technology, railroads should consider the challenges and constraints associated with that technology and how they want to use it. Understanding these challenges and constraints can support system designers in exploring roles for humans and automation, selecting appropriate teaming structures, and developing detailed design requirements for technologies and HATs. This will increase the likelihood that such HATs will perform effectively and efficiently and address the railroad's goals.

10. Conclusion

The railroad industry seeks to adopt emerging and evolving automated technologies to improve the efficiency and effectiveness of the track inspection process. In addition to improving railroads' ability to detect and address defects in the short-term, automation and increased availability of track condition data may also facilitate more effective long-term planning for proactive maintenance activities.

The use of automation introduces new challenges alongside these benefits. The most prevalent approach to adopting new automation is technology centered. It focuses on automating as much as possible, based on the capabilities of the technology, and relies on humans to manage any limitations of the technology. This approach can introduce problems, such as skill degradation, erratic workload, and reduced efficiency.

This report explored how a human-centered approach to design can mitigate many of the challenges associated with automation. This means carefully considering how tasks and responsibilities should be allocated to humans and automation and taking into consideration the roles, abilities, and limitations of each, as well as the possible interactions between them.

Research on function allocation (i.e., the process of examining how tasks should be divided and shared between humans and automation) has evolved over time, with more recent approaches emphasizing HAT and seeking to optimize the joint performance of all humans and technologies involved in a process.

This report described a process for designing HATs that can be summarized as follows:

1. Step 1: Examine process requirements and system-level goals
2. Step 2: Explore possible roles for humans and automation
3. Step 3: Consider tradeoffs and challenges
4. Step 4: Develop detailed design requirements
5. Step 5: Implement and assess HATs

Researchers explored how this process could be used at a high level. The authors of this report encourage system designers to develop context-specific HAT requirements for any technologies or applications of automation they wish to use.

A four-stage information processing model was used to examine the requirements for four tasks in the track inspection process, including data collection, data analysis, decision making, and action. It examined the complexities of each of these tasks, including possible roles for humans and automation, teaming possibilities using LOA models and systems perspectives, and design considerations and tradeoffs.

It is important to understand and carefully weigh the strengths and weaknesses associated with each design choice being considered, which is one of the steps needed to design HAT structures that will work for the desired environment.

Overall, important areas to address in track inspection HAT include:

- Technology and interface design
- Training and experience

- Communication and coordination between team members
- Supporting both short-term and long-term goals

Railroads, technology manufacturers, and track inspection researchers may use the information contained in this report to adopt a more human-centered approach to the use of automation in track inspection.

11. References

Al-Nazer, L., Raslear, T., Wilson, R., & Kidd, J. (2017). *Quantification of the sensitivity of two prevalent track inspection systems.* (Report No. DOT/FRA/ORD-17/03). Federal Railroad Administration.

Baillargeon, J. P. (2020, October). FRA RD&T: Using AI to improve safety. *Railway Age.*

Bainbridge, L. (1983). Ironies of automation. *Automata, 19*, 775-779.

Battiste, V., Lachter, J., Brandt, S., Alvarez, A., Strybel, T. Z., & Vu, K. L. (2018). Human-automation teaming: lessons learned and future directions. *Human Interface and the Management of Information. Information in Applications and Services. HIMI 2018* (pp. 479-493). Las Vegas, NV: Springer.

Billings, C. (1997). *Aviation Automation: The Search for a Human Centered Approach.* Mahwah, NJ: Erlbaum.

Cognitive work analysis. (2021, August 13). Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Cognitive_work_analysis

Cooke, N., Demir, M., & Huang, L. (2020). A framework for human-autonomy team research. *Human Computer Interaction International Conference.* Copenhagen, Denmark.

Cummings, M. L., Huang, L., Zhu, H., Finkelstein, D., & Wei, R. (2019). The impact of increasing autonomy on training requirements in a UAV supervisory control task. *Journal of Cognitive Engineering and Decision Making, 13(4)*, 295-309.

Dadashi, N., Wilson, J. R., Golightly, D., & Sharples, S. (2014). Framework to support human factors of automation in railway intelligent infrastructure. *Ergonomics, 57*(3), 387-402.

Department of Defense. (2020). *Design criteria standard: Human engineering.*

Ehrlich, K., & Rohn, J. (1994). Cost justification of usability engineering. In R. Bias, & D. Mayhew, *Cost-justying usability* (pp. 73-110). New York: Academic Press.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors Journal, 37*(1), 32-64.

Federal Aviation Administration. (2016). *Requirements for a human factors program.*

Fitts, P. M., Viteles, M. S., Barr, N. L., Brimhall, D. R., Finch, G., Gardner, E., & Stevens, S. S. (1951). *Human engineering for an effective air-navigation and traffic-control system.* Ohio State University Research Foundation.

Fleming, C. H. (2015). *Safety-driven early concept analysis and development.* Massachusetts Institute of Technology.

France, M., Melnik, G., Safar, H., & Multer, J. (2020). *Reducing hazards associated with visual and automation-aided track inspections* (Report No. DOT/FRA/ORD-21/18). Federal Railroad Administration.

Golightly, D., Kefalidou, G., & Sharples, S. (2017). A cross-sector analysis of human and organisational factros in the deployment of data-driven predictive maintenance. *Information Systems and E-Business Management, 16*, 627-648.

Grimes, E., Wright, K., & Hillier, G. (2009). *Integration in design project: Delivering the return on investment.* London: Mott MacDonald.

Gutzwiller, R. S., Espinosa, S. H., Kenny, C., & Lange, D. S. (2018). A Design pattern for working agreements in human-autonomy teaming. *Advances in Intelligent Systems and Computing, 591*.

Hollnagel, E. (2001). Human-Oriented Automation Strategies. *Proceedings of the World Congress on Safety of Modern Technical Systems*, (pp. 443-452).

Inagaki, T. (2001). Adaptive Automation: Sharing and Tradingof Control. *The Proceedings of the Transportation and Logistics Conference*, (pp. 79-84).

Joe, J., O'Hara, J., Hugo, J. V., & Oxtrand, J. H. (2015). Function allocation for humans and automation in the context of team dynamics. *Procedia Manufacturing, 3*, 1225-1232.

Johnson, M., Vignatti, M., & Duran, D. (2020). Understanding human-machine teaming through interdependence analysis. In *Contemporary research.* CRC Press.

Kefalidou, G., Golightly, D., & Sharples, S. (2018). Identifying rail asset maintenance processes: A human-centric and sensemaking approach. *Cognition, Technology and Work, 20*, 73-92.

Li, Y., & Burns, C. M. (2017). Modeling automation with cognitive work analysis to support human-automation coordination. *Journal of Cognitive Engineering and Decision Making, 11*, 299-322.

Melnik, G., Roth, E., Multer, J., Safar, H., & Isaacs, M. (2018). *An Acquisition Approach to Adopting Human Systems Integration in the Railroad Industry* (Report No. DOT/FRA/ORD-18/05). Federal Railroad Administration.

National Aeronautics and Space Administration. (2014). *Human integration design handbook.*

National Aeronautics and Space Administration. (2014). *Human integration design process.*

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A Model for Types and Levels of Human Interaction with Automation. *IEEE Transactions on Systems, Man and Cybernetics: Part A Systems and Humans, 30*, 286-297.

Pritchett, A. R., Kim, S. Y., & Feigh, K. M. (2014). Measuring human-automation function allocation. *Journal of Cognitive Engineering and Decision Making, 8*, 52-77.

Roberts, C. (2003). Predicting track inspection maintenance and renewal requirements using decision support systems. *World Congress on Railway Research (WCRR)* (pp. 678-688). Edinburgh: The World Railway Conference.

Roth, E. M., Shushereba, C., Militello, L. G., Diiulio, J., & Ernst, K. (2019, December). Function allocation considerations in the era of human autonomy teaming. *Journal of Cognitive Engineering and Decision Makig, 13*(4), 199-220.

Saadat, S., & Sherrock, E. (2018). *Autonomous Track Geometry Measurement Technology Design, Development, and Testing* (Report No. DOT/FRA/ORD-18/06). Federal Railroad Administration.

Save, L., Feuerberg, B. E., & Avia, E. (2012). Designing Human-Automation Interaction: A New Level of Automation Taxonomy. In D. De Waard, K. Brookhuis, F. Dehais, C.

Weikert, S. Rottger, D. Manzey, . . . P. Terrier (Ed.), *Proceedings Human Factors of Systems and Technology.* Toulouse, France: Human Factors and Ergonomic Society.

Sheridan, T. B., & Verplanck, W. L. (1978). *Human and computer control of undersea teleoparators.* MIT Man-Machine Systems Lab.

Shukla, D. (2021, July 7). *Industry 4.0 Solutions for New-Age Railways and Airways.* Retrieved from https://www.electronicsforu.com/technology-trends/tech-focus/industry-4-solutions-railways-airways

Singh, R. R. (2018). Designing for multi-agent collaboration: a shared mental model perspective (doctoral dissertation).

Small Unmanned Aircraft Systems 14 C.F.R. Part 107. (2016).

Track Safety Standards 49 C.F.R Part 213. (2019).

Appendix A. Glossary of Terms

TERM	DEFINITION
Abstraction hierarchy	A multi-level knowledge representation framework for describing the functional structure of a particular work domain or system.
aTGMS	“Autonomous” Track Geometry Measurement System. A type of TGMS (an automated inspection technology) that is not staffed, and instead consists of sensors and computers mounted on a locomotive or freight car. This automated technology runs continuously, and data is analyzed (typically after the fact) by employees not on board the vehicle.
Automated technology; automated inspection technology	A technology which uses sensors to collect track condition data (rather than human senses) and uses computers to perform some interpretation or analysis of that data. May include some human analysis and decision making. Typically used in addition to visual inspection.
Automation-aided inspection	Track inspection that uses automated technologies, such as TGMS and aTGMS, to assist humans with the work of finding, making decisions about, and recording track defects and maintenance conditions.
Class-limiting defect	A track problem that exceeds thresholds or falls outside track-class based requirements set by an FRA regulation. This type of defect can be addressed by reducing track class (i.e., imposing speed restrictions) so that the track no longer exceeds thresholds for the new, reduced track class.
Cognitive task analysis	A research method for uncovering and representing what people know and how they think.
Cognitive work analysis	A framework to model complex sociotechnical work systems. The framework models different types of constraints, building a model of how work could proceed within a given work system. The focus on constraints separates the technique from other approaches to analysis that aim to describe how work is conducted or prescribe how it should be conducted (Cognitive work analysis, 2021).
Conventional TGMS	A type of TGMS that operates using staffed vehicles (often referred to as “geometry cars”) equipped with sensors and computers. Railroads must schedule conventional TGMS inspections as they require track time and dedicated operators, as well as a locomotive and train crew to pull the geometry car. The operators are onboard and can analyze data in real time.
Data analysis	A process of inspecting, cleansing, transforming, and modelling data with the goal of discovering useful information and informing conclusions.
Data analyst	The person responsible for examining an automated system’s output.

TERM	DEFINITION
Degraded track condition	An identified track problem that may or may not exceed an FRA-regulated threshold. May be classified as either a safety defect or a maintenance condition.
Exception	A potential defect or maintenance condition identified by an automated inspection technology such as conventional TGMS or aTGMS.
Function allocation	A method of determining how to allocate functions to humans and machines (automation). Though this has historically been treated as an either/or decision, modern approaches to function allocation include more nuanced relationships, such as dynamic task allocation and shared task performance.
Human-automation teaming (HAT)	Broadly defined for the purposes of this report as humans and technology working together toward the same goals.
Hi-rail vehicle	A dual-mode vehicle which can operate both on rail tracks and a conventional road.
Level(s) of automation (LOA; LOAs)	Refers to a continuum from manual to fully automatic operations for systems involving humans and machines.
MABA-MABA	Men are better at – Machines are better at
Maintenance condition	A track problem that does not exceed FRA regulation thresholds but does exceed thresholds voluntarily set by the railroad.
Maintenance standard	Internal standards set by the railroad and used to hold track to stricter standards than FRA regulation and typically based on FRA regulations for the next highest class of track.
Maintenance planning	The process of determining what longer-term maintenance work should be done and how it should be carried out. This is one aspect of a railroads' capital planning, i.e., deciding how to allocate resources over time.
Safety defect ("defect")	A track problem that exceeds thresholds or falls outside requirements set by an FRA regulation.
Safety issue	A track problem that poses safety concerns, including safety defects or problems comprised of multiple degraded track conditions below regulatory thresholds.
Staffed inspection vehicle	Refers to automated inspection technologies with onboard staff. May also be referred to as "manned."
Systems perspective	A perspective which considers emergent properties, such as safety, which result from interactions between people, technology, and processes.
TGMS	Track Geometry Measurement Systems: A category of automated inspection technology used to examine track geometry. Includes both

TERM	DEFINITION
	conventional track geometry measurement systems (TGMS) and autonomous track geometry measurement systems (aTGMS).
Technology operator(s)	Humans involved in operating an automated inspection technology. They may be responsible for inputting parameters and monitoring the data output, as well as screening the data output and determining what actions are needed (e.g., dismissing exceptions, contacting the engineering department, or reaching out to dispatchers).
Track safety standards	Regulations set forth in 49 CFR Part 213 regarding railroad track classes, allowable speeds, and track inspection and maintenance.
Verification	The process of following up on potential track safety issues, either identified by an automated inspection technology or by other track users and inspectors (e.g., bridge and building inspectors).
Visual inspection	Inspections that a human inspector performs either while walking the track or riding in a hi-rail vehicle. These inspections are primarily but not exclusively visual. Inspectors may also use auditory cues (e.g., rattling noises) and kinesthetic or motion cues (e.g., bumps and vibrations) in addition to visual cues to detect track conditions and identify issues.

Appendix B. Expanded Discussion of the General Process for Designing HATs

This Appendix covers the general process for designing HATs that is discussed in Section 4. This is an expanded version of that discussion that includes additional information and references human factors literature and methods that may be useful for system designers and human factors professionals. Consult Appendix C for a list of guiding questions that correspond to each step in the general process for designing HATs.

Step 1: Examining Process Requirements and System-Level Goals

One of the first activities essential to designing effective HATs is to examine the work that the team will perform. It is not possible to allocate work effectively unless the system designer is clear about all the work that must be done. Examining the work that the team will perform requires thinking beyond what the technology itself can do and thinking about all the work that must be done by the users of the technology and what must be accomplished by the joint human-automation team more broadly.

Roth et al. (2019) suggest two questions that system designers must ask at the beginning of the function allocation process: "what is the nature of the work to be done?" and "what makes it challenging?" Understanding the answers to these questions should precede asking "who does what?"

Roth et al. (2019) note that this examination of operational demands and challenges should include not only the typical functions to be performed, but also functions that may occur in non-routine situations. For example, in track inspection, it is important to consider not only routine inspections but also those that may be required after severe weather events or other circumstances that may result in damage to the track.

This examination of process requirements may require eliciting knowledge from subject matter experts or those who normally perform the task. Cognitive Task Analysis and Cognitive Work Analysis are two useful methods for eliciting such information and examining task requirements and system goals (Li & Burns, 2017).

Two important considerations when examining process requirements and system-level goals are task decomposition and task interdependencies.

- **Task Decomposition**. In addition to recognizing what makes the work challenging, it is also important to understand how work can be broken down into tasks and subtasks. This helps ensure each individual part of the process is considered and allows decision makers to consider how each task can be divided or shared among humans and automation.

  An abstraction hierarchy is one way of examining a work process and its associated tasks that can be used in function allocation (Li & Burns, 2017; Roth et al., 2019). This involves representing the work at increasing degrees of specificity, beginning from the very general and narrowing to reflect specific ways that those functions can be physically implemented. As a simple example of an abstraction hierarchy, consider "illuminating a room" as a system goal or functional purpose. Next, the abstraction hierarchy examines values and priorities, as well as constraints and restraints, that shape how this functional purpose can be achieved. For example, perhaps the room does not have any natural light sources and the designer would like to prioritize accessibility. Following this, the hierarchy specifies functions that could meet the goal, such as turning on a light. However, simply stating the general function does not require specifying how the function is performed. The next level of abstraction identifies more specific physical functions that may accomplish the desired function, such as flipping a switch, pressing a button, or entering the room. Lastly, switches, buttons, and motion sensors are all physical forms or objects that may be used to perform the function.

This report uses a very simple task decomposition for track inspection based on a four-stage cognitive information processing model (Parasuraman et al., 2000), which breaks the track inspection process into four tasks: data collection, data analysis, decision making, and action. Task decomposition is detailed in Section 5.1.1. These tasks and subtasks may have interdependencies that must be considered as well.

- **Task Interdependencies.** When tasks are decomposed into subtasks, there may still be critical interactions between them that are necessary for the success of the track inspection process. Johnson et al. (2020) described several types of interdependent relationships between tasks.

  Activities may occur sequentially or in parallel. Many track inspection activities take place sequentially. For example, data collection occurs before data analysis and decision making. However, some activities can be performed in parallel, such as inspecting for different types of defects.

  Activities can also have "and" relationships, meaning that all of the activities must be completed, or "or" relationships, meaning that one of the activities must complete successfully. Both data collection and data analysis must occur before decision making. However, data collection by a human or data collection by an automated technology must identify a defect to ensure it can be appropriately addressed.

  These relationships can be summarized as four types of task relationships in joint activity (Johnson et al., 2020):

  1. **Sequential-And**: All activities need to be completed successfully in sequence
  2. **Sequential-Or**: Activities executed in sequence until one is completed successfully
  3. **Parallel-And**: All activities can be executed in parallel and need to be completed successfully
  4. **Parallel-Or**: All activities can be executed in parallel but are racing each other until one completes successfully

These relationships between tasks have implications for how to structure the HAT and the design requirements. For example, sequential activities distributed among different humans and technologies mean that those actors must coordinate to hand off each task. Using more "or" relationships can add redundancy to the inspection process and create opportunistic relationships that, while perhaps not strictly required for the process to function, can use teaming and coordination to increase system resilience (Johnson et al., 2020).

Step 2: Exploring Possible Roles for Humans and Automation

Once system designers have examined the process requirements and system level goals, the next step is to explore possible roles for humans and automation for each task.

This includes examining possible roles for humans and automation, examining various ways that they could work together to perform the task and their capabilities to perform or support various activities, and noting interdependencies between team members.

- **Exploring Roles Using LOAs.** LOA models provide a common way to describe and categorize human-automation interactions (e.g., sharing or trading off work, using automation as a backup to human actions, or using automation to replace some subset of the tasks a human normally performs.

  Early research in human-automation interaction noted that humans and automation may interact by sharing work, or trading off work, using automation as a backup to human actions, or using automation to replace some subset of the tasks a human normally performs (Sheridan & Verplanck, 1978). LOA models provide one way of categorizing and describing these interactions.

  Parasuraman et al. (2000) described LOAs using a four-stage information processing model. This is the model used in this report. For track inspection, the four tasks of data collection, data analysis, decision making, and action are detailed in Section 5.1.1. Other researchers, including Save et al. (2012) built upon this LOA framework, providing a detailed classification system for LOAs across each of these four tasks.

  Although such frameworks are not comprehensive, they are nonetheless helpful because they are a widely used and easily understood way of describing human-machine interactions. This report includes examples within a LOA framework for each task in the inspection process.

- **Exploring Roles Using Teaming Models.** In addition to using LOA models, it is often valuable to examine possible roles for humans and automation that do not fit neatly into such a framework. As noted in Section 3.3, more recent developments in the HAT literature provide more holistic perspectives on the ways that humans and automation can perform work.

  In examining the roles for humans and automation within a system, it may be useful to consider systems where work may be distributed across a range of people and automated technologies (Singh, 2018). Where appropriate, designers may also examine ways that the workflow can be altered dynamically (e.g., adaptive automation). This ability to adjust who does the work based on the work context represents a more recent consideration in the design of HATs (Gutzwiller, Caitlin, & Lange, 2018; Inagaki, 2001). These approaches call for examining the system in a more holistic fashion to address the interdependent nature of the work.

- **Assessing Capabilities of Humans and Automation for Each Role.** As noted in Section 3.1, the capabilities of humans and automation were some of the earliest criteria used in function allocations. Technology has evolved significantly since these early capability-based function allocation methods were developed, and automation can now perform many tasks that were previously only possible for humans. However, it remains true that humans and automation each have unique strengths and weaknesses that allow them to contribute to systems in different ways. Therefore, examining the capabilities of humans and automation remains an important aspect of system design.

Rather than treating the capabilities of automation as the primary function allocation criteria, modern techniques take a more nuanced approach to integrating human and machine capabilities into function allocation to capitalize on the strengths of each. One example of how capabilities can be integrated into modern function allocation approaches is included in the interdependence analysis proposed by Johnson et al. (2020). The capability analysis they propose involves examining the capability of humans and automation to either perform or support various tasks. Table 8 lists the capability assessment categories used in interdependence analysis.

Table 8. Capability Analysis Categories, adapted from Johnson et al. (2020)

Performer	Supporting Team Member(s)
A: I can do it all	A: My assistance could improve efficiency
B: I can do it all, but my reliability is <100%	B: My assistance could improve reliability
C: I can contribute but need assistance	C: My assistance is required
D: I cannot do it	D: I cannot provide assistance
E: Not applicable/not significant	E: Not applicable/not significant

The inclusion of supporting roles in capability assessments is an important shift from traditional capability-based function allocation methods that allocate tasks primarily according to the capability of automation to perform a task under normal conditions. This analysis considers that both humans and automation can play supporting roles, providing a better picture of the range of possibilities for HATs. Examining such possibilities also allows system designers to consider whether they are creating supporting roles that would be difficult for humans to perform, a common pitfall of technology-centered design approaches.

- **Teaming Interdependencies.** Once designers identify possible roles for humans and automation and examine their ability to successfully serve in performing or supporting roles, determine where teaming interactions or coordination between humans and automation are necessary, as well as where coordination is optional but could improve efficiency or reliability. Johnson et al. (2020) discuss the importance of interdependence and present a method for analyzing interdependencies, as well as a set of principles that can be used in designing HATs.

  If the human or automation performing the task has limited capability or reliability, a supporter may be important to make sure the task is successful. For example, a technology with low reliability in identifying defects may require a human to carefully review its outputs. Similarly, a human with low ability to detect internal rail flaws may require an automated technology to perform data collection to effectively detect such defects.

In cases where a supporting actor is not required, "opportunistic teaming" can contribute to system resilience. For example, following a visual inspection with the use of an automated inspection technology may increase the likelihood of detecting defects. Often, simply allocating a task to the human *or* automation would be limiting the success of the joint HAT and using a combination of human and technology to perform a task may be beneficial or even necessary for optimal task performance.

Step 3: Consider Tradeoffs and Challenges

After identifying the possible roles of humans and automation, the relative capabilities of each to perform those roles, and any associated interdependencies, consider tradeoffs and challenges associated with the desired teaming structure. Each option has benefits and challenges. System designers must determine which option(s) are best suited to meet the system goals and requirements (as discussed in Section 4.3). Assess human factors impacts of teaming structures and select options that meet the system requirements.

- **Assessing Human Factors Impacts of Possible Function Allocations.** In assessing tradeoffs and challenges associated with possible roles for humans and automation, it is important to consider how these decisions affect the humans in the system.

  For example, one possible consequence of increasing the use of automation is loss of situation awareness (Endsley, 1995). As organizations introduce more automation, the potential exists for humans to become more removed from the details of their work. These changes contribute to loss of situation awareness and skill loss that makes it challenging to take over when the automation fails. Billings (1997) identified four characteristics that contribute to problems with automation, shown in Table 9. These are examples of considerations that should be examined when assessing the viability of design decisions related to HATs.

Table 9. Characteristics Associated with Evolution of Automation, adapted from Billings (1997)

Automation characteristics	How it impacts the human
Complexity	Makes details more difficult to understand and remember
Coupling	Relationships or interdependencies among or between automation functions may be hidden from view and surprise the human by their behavior.
Autonomy	Self-initiated behavior by the automation may take place which was not expected by the human operator. Deciding whether this behavior is appropriate can be a challenge because of tight coupling.
Inadequate feedback	Automation may not communicate or communicates poorly what it is doing and why.

- **Selecting Options that Meet System Requirements.** Roth et al. (2019) suggest that tradeoffs may take into consideration several factors such as maturity of the required technology, human-system integration considerations (e.g., personnel selection and training costs), and operational considerations (e.g., ability for humans and/or automation to perform under unusual circumstances).

For track inspection, system designers, researchers and railroads will need to consider which combinations of automated technologies and human actors best meet their needs. This may be partly shaped by constraining factors such as the costs of inspection technologies, compliance with regulations, requirements for track time, and staging and training requirements.

Pritchett et al (2014) identified five requirements for effective function allocations, as listed below.

- Each agent must be allocated functions that it is capable of performing
- Each agent must be capable of performing its collective set of functions
- The function allocation must be realizable with reasonable teamwork
- The function allocation must support the dynamics of the work
- The function allocation should be the result of deliberate design decisions

These requirements can be used to examine whether the function allocation or HAT structure is likely to result in effective work or likely to encounter problems.

Step 4: Developing Detailed Design Requirements

Once system designers determine which HAT structures are best suited to meet the overall system goals and high-level requirements, based on the tradeoffs and considerations discussed in Section 4.3, they must then determine specific design requirements for the system.

Designers may derive some design requirements from the interdependencies identified in prior steps, as well as the expectations for what tasks the HAT will perform. Some requirements may address activities that occur outside the inspection process itself, such as training. These design requirements can help address any limitations associated with the selected HAT structure.

- **Develop requirements based on task interdependencies and expectations for the HAT's performance.** For tasks that are shared among humans and automation, interdependencies of the process may shape design requirements. For example, Johnson et al. (2020) cite requirements for observability, predictability, and directability as important parameters to consider. These factors describe "who needs to observe what from whom, who needs to be able to predict what, and how members need to be able to direct each other for a given aspect of the work." These requirements may dictate design decisions such as the locations of the humans and automated technologies, the timing of their activities, and the design of human-machine interfaces.

  There are many examples in the research literature that can be used to help develop detailed design requirements. For example, researchers have developed criteria for HATs based on what is known about human-human teams (Joe et al., 2015; Johnson et al., 2020). Designers may establish work agreements to describe the roles of humans and automation, the tasks they will perform, and establish shared expectations (Gutzwiller et al., 2018).

- **Consider requirements beyond the inspection process itself (e.g., training requirements).** Requirements may also extend beyond the four inspection tasks discussed in this report. For example, researchers have examined how requirements for training may change as the use of autonomous systems increases. One study of UAV operators found that humans may require less skill-based training to achieve the same performance when supported by automation. However, the study also found that novice operators may have a greater bias toward trusting the automation and overlooking when it makes errors, whereas experienced operators may be better able to catch these errors (Cummings et al., 2019). New forms of training may be required to develop appropriate levels of trust in automation and maintain the expertise needed to identify automation errors.

Step 5: Implementing and Assessing Human-Automation Teams

After design decisions have been made, it is important to assess whether the HAT is functioning as intended. This involves selecting assessment criteria and determining whether to adjust the design or operation of the HAT.

- **Choosing assessment criteria.** Pritchett et al. (2014) developed a set of eight criteria for evaluating HATs, which can be used to assess proposed HAT structures before they are implemented, as well as to assess existing HAT structures. These criteria are listed in Table 10. Designers should select assessment criteria that consider the impacts on human team members as well as the overall performance of the HAT.

Table 10. Criteria for Evaluating HATs, adapted from Pritchett et al. (2014)

Teaming Metric	Description
Workload/Taskload	Workload includes both cognitive and physical workload stemming from the task itself, as well as from required teamwork and coordination. It is important to identify whether the chosen teaming structure will lead to workload spikes
Mismatches between Responsibility and Authority	Mismatches between responsibility and authority are a possible problem that can arise from delegating tasks to automation while assigning final responsibility to humans. This creates additional work for the humans who must monitor the automation. Matching authority and responsibility (i.e., giving humans greater authority over the tasks they bear responsibility for) may reduce coordination and monitoring burdens.
Stability of the Human's Work Environment	A stable work environment allows human team members to predict and plan for upcoming actions. Unpredictability or instability may occur if function allocation changes unexpectedly, automation unexpectedly requires new actions from the human, or outside factors cause unexpected changes to the work.
Coherency of a Function Allocation	Coherent function allocations establish clear roles and work practices for human and automated team members. Overlapping or unclear areas of responsibility may require more significant coordination to prevent conflicts.
Interruptions	Though sometimes necessary, interruptions can be disruptive to human work. Structuring HATs to minimize interruptions except where truly warranted may improve human performance.
Automation Boundary Conditions	Boundary conditions are the limits of automation's abilities to perform its assigned functions. Such boundary conditions may not always be explicitly known; however, it can be valuable to monitor when automation fails to achieve its targets to understand where boundary conditions have been exceeded and design around these limitations.
System Costs and Performance	System performance can include both safety and robustness to non-routine occurrences, while costs may include things like resource requirements. Costs and performance can be simulated or measured during system operations to determine whether they meet the desired targets.
Human's Ability to Adapt to Context	Function allocations should support humans' ability to dynamically adapt to the context. This may include using automation in flexible ways, monitoring more or less closely depending on the situation and type of work. If the human role is over constrained, the loss of flexibility may have negative impacts on system performance.

- **Making changes to the design of new or existing HATs.** Though integrating HAT considerations into the systems engineering process early on is most effective, designers can still make improvements after a particular technology is implemented, particularly if unforeseen challenges related to human-automation interaction arise. However, when assessing existing systems, the ability to make improvements in system performance may be more limited and/or more costly than if designers had conducted such an assessment earlier in the design process.

Appendix C. Questions to Guide Design

This appendix consolidates the questions included in the body of this report for convenient reference.

Questions to Guide Function Allocation in Track Inspection

Below are a set of questions for designers at various steps in the function allocation process:

1. Examine task requirements and system goals
    - What are the overall goals of the process?
    - What makes that work process challenging?
    - How can the process be broken into smaller tasks for analysis?
    - In what ways are these subtasks interdependent?
2. Explore possible roles for humans and automation
    - What roles can humans and automation perform or support?
    - What LOAs are possible or desirable?
    - In what ways could humans and automation interact?
    - In what ways does this task depend on or influence others?
3. Consider tradeoffs and challenges
    - What are the benefits and challenges for each option?
    - Do the roles for both humans and automation make sense?
4. Develop detailed design requirements
    - What is necessary to make the HAT successful, particularly given any interdependencies between tasks?
    - What challenges can be addressed through system design?
    - What additional requirements, e.g., training, should be considered?
5. Implement and assess HATs
    - Once implemented, how will system performance be monitored?
    - What unanticipated challenges have been encountered?
    - Where is there room for improvement?

Questions Regarding the Data Collection Task

- Who (or what) is primarily responsible for collecting the data?
- Who (or what) is involved in supporting data collection (e.g., monitoring the quality of data, etc.)?

- Where are human(s) and/or automation involved in data collection located when inspection is taking place?
- When are humans involved in data collection (e.g., throughout the process, periodically when they receive alerts, or after it is completed)?
- How often is data collected (e.g., continuously during revenue service, routinely occurring on a fixed interval, periodically but irregularly, etc.)?
- What type of data is being collected?
- How does the method of data collection affect *analysis* of that data?

Questions Regarding the Data Analysis Task

- What is the purpose of the analysis (e.g., identifying safety concerns, identifying long-term maintenance needs, or both)?
- What types of track conditions or defects are being analyzed? Are they objective or subjective assessments (i.e., is the severity classification based on a single threshold, or more complex pattern recognition)?
- What kind of data is being analyzed (e.g., data from visual inspection, data from a single automated technology, data from multiple automated technologies, data from several inspections using the same technology, etc.)?
- Does the analysis require gathering any additional data for verification purposes? If so, when and how does this occur (e.g., at the time of inspection and analysis, or later)?
- Who (or what) is primarily responsible for analyzing the data (e.g., track inspectors evaluating defects during visual inspection, automated systems doing the bulk of data processing from an automated data collection run, etc.)?
- Who (or what) is involved in supporting data analysis (e.g., data analysts reviewing analysis outputs, field staff verifying presence of defects, etc.)?
- Where does data analysis take place (e.g., at the location of the inspection, or elsewhere)?
- When does data analysis take place (e.g., at the same time as inspection, or later; all at once, or spread out over time)?
- How many humans are involved in data analysis, and how do they coordinate with one another?

Questions Regarding the Decision Making Task

- What type of decision is being made (e.g., determining appropriate short-term actions or planning for long term maintenance needs)?
- What types of track conditions or defects are included in this decision making? Are they objective or subjective assessments?
- What kind of data sources informed the analysis on which these decisions are based (e.g., data from visual inspection, data from a single automated technology, data from multiple automated technologies, data from several inspections using the same technology, etc.)?

- Who (or what) is primarily responsible for decision making (e.g., track inspectors, other field staff, supervisors, data analysts, or an automated system)?
- Who (or what) is involved in supporting data analysis (e.g., automated analysis tools, decision support tools such as automation that narrows or proposes options, or other humans)?
- Where does decision making take place (e.g., at the location of the inspection, or elsewhere)?
- When does decision making take place (e.g., at or near the same time as inspection, or later)?
- How many humans are involved in decision making, and how do they coordinate with one another?

Questions Regarding the Action Task

- What type of action is being taken (e.g., maintenance or repair of track infrastructure vs. communication with dispatcher to impose a speed restriction or remove track from service)?
- Who (or what) is primarily responsible for this action (e.g., track inspectors, other field staff, supervisors, data analysts, or an automated system)?
- Who (or what) is involved in supporting data analysis (e.g., automated technologies for performing maintenance, automated communications with dispatch, or additional humans to support action execution)?
- Where does the action take place (e.g., at the location of the inspection, or elsewhere)?
- When does action take place (e.g., at or near the same time as inspection, or later)?
- How many humans are involved in the action, and how do they coordinate with one another?

Abbreviations and Acronyms

ACRONYM	DEFINITION
AI	Artificial Intelligence
aTGMS	Autonomous Track Geometry Measurement Systems
FAA	Federal Aviation Administration
FRA	Federal Railroad Administration
GoA	Grades of Automation
HAT	Human-Automation Teaming
HATs	Human-Automation Teams
LIDAR	Light detection and ranging
LOA(s)	Level(s) of Automation
TGMS	Track Geometry Measurement Systems
UAS	Unmanned Aircraft System