Public Quarterly Report

Date of Report: 5th Quarterly Report-December 29, 2020

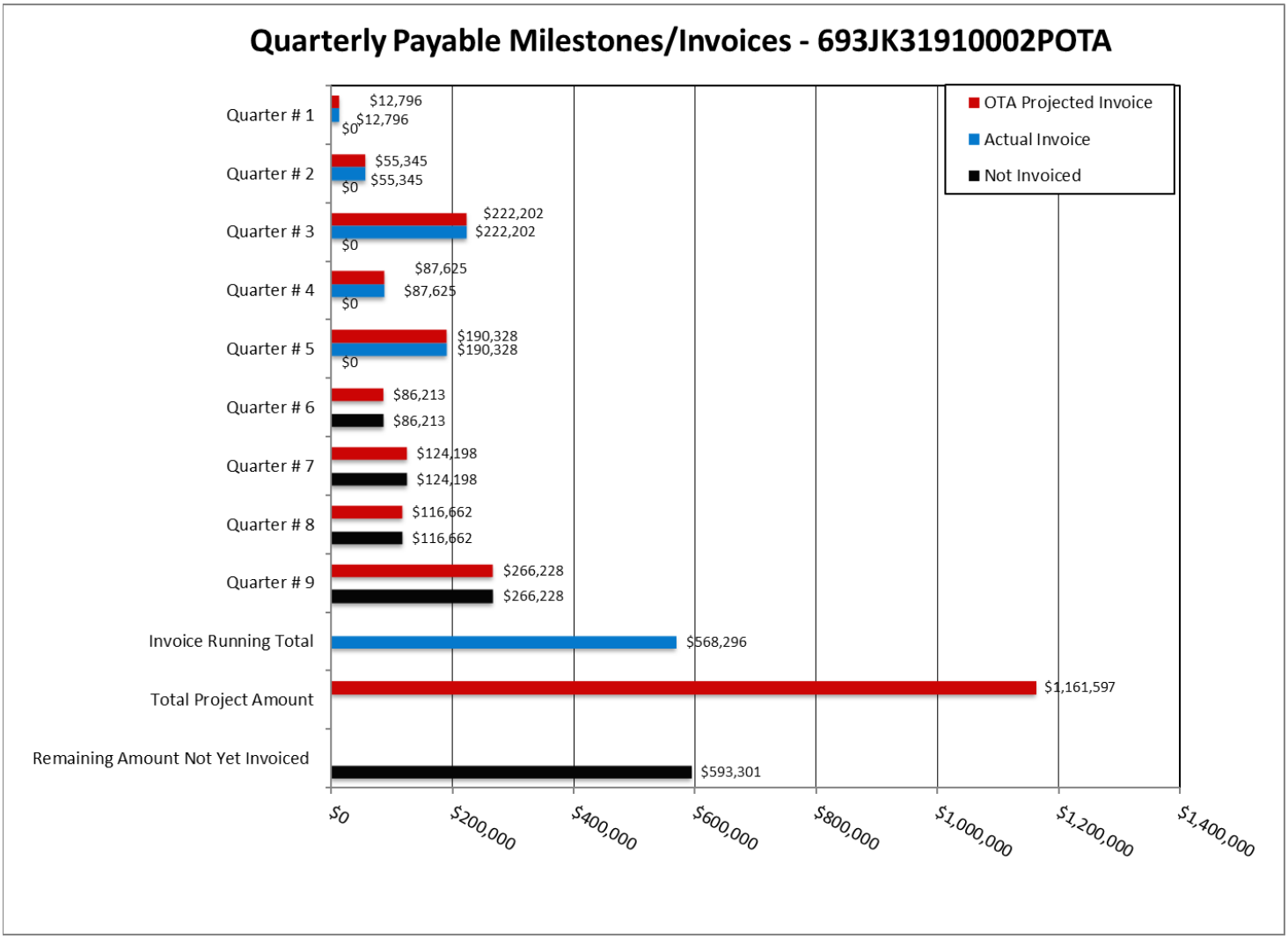
Contract Number: 693JK31910002POTA
Prepared for: DOT PHMSA
Project Title: Data Collection, Normalization and Integration Methods to Enhance Risk Assessment Tools for Decision-Making
Prepared by: GTI (Gas Technology Institute)
Contact Information: Ernest Lever, elever@gti.energy , 847-544-3415
For quarterly period ending: December 29th, 2020

1: Items Completed During this Quarterly Period:

Item #	***Task #***	***Activity/Deliverable***	***Title***	***Federal Cost***	***Cost Share***
19	3.2	Decision Support System	Working prototype decision-support tool with initial documentation	31,929.00	41,525.00
20	4.1.3	Data Integration, Normalization and Uncertainty Tools	Bayesian Network: Methodology for post processing data for BN	49,271.00	0.00
21	4.1.4	Data Integration, Normalization and Uncertainty Tools	Bayesian Network: BN model parameterized by using selected data sources	49,271.00	0.00
22	4.2.4	Data Integration, Normalization and Uncertainty Tools	Data Driven Insights:Causal Inference – identify data combinations that are best dealt with by causal modeling	11,287.00	21,746.00
23	4.3.2	Data Integration, Normalization and Uncertainty Tools	Causal Modeling and Situational Awareness:Situational awareness framework	43,580.00	0.00
24	5.5	Project Management	Fifth Quarterly Report	4,990.00	0.00

A contract modification to correct task numbering and reporting schedule to match the extended timeline was submitted and executed.

3: Project Financial Tracking During this Quarterly Period:



4: Project Technical Status –

The technical work for this quarter progressed as planned and all deliverables have been completed. Detailed documentation for each deliverable is provided in attachments. A short description of project activity is provided in the following numbered points:

1. The goal of this project is to create methods and software tools to help operators of natural gas pipelines better estimate risks and improve risk-mitigation decisions. Lumina Decision Systems, Inc (Lumina) has the primary responsibility to design and develop a decision-support tool called the Analytica Natural Gas Risk Assessment Model (ANAGRAM). The tool will incorporate the new methods and models for risk estimation developed by the team and be made accessible to risk managers at gas utilities. During Q2 2020, Lumina's primary focus Task 2.1, was to design a prototype decision-support tool to collect, normalize, and integrate risks from models, and assist in decisions on risk mitigation projects, and develop initial mock-ups of user-interface screens. During Q3 and Q4 2020, Lumina's primary deliverable was

a. Task 3.2: Decision Support System: Working prototype decision-support tool with initial documentation
b. Key elements of this task were:
c. Extend ANAGRAM to use data on gas distribution assets from PG&E DIMP.
d. Extend the existing ANAGRAM tool based on the revised user-interface design and mock-ups from task 3.1. See the user-interface screenshots with call-outs documenting key features and functions below.
e. Upgrade ANAGRAM to use the new Analytica Cloud Platform 3.0.

2. Within the overall project framework, UMD is responsible for research activities focused on third-party excavation damage (TPD). According to the U.S. DOT's Pipeline and Hazardous Materials Safety Administration (PHMSA), pipeline incidents have led to 281 fatalities, 1,183 injuries, and over $10B in damage over the past 20 years. In addition, excavation damage causes 15% of pipeline failures and is the leading cause of fatal pipeline incidents, making the study of TPD highly relevant into both industry and PHMSA.
   This report details UMD progress over the project quarters Q4 and Q5. The primary task for these project periods was to complete the formation of a Bayesian Network (BN) model that represents TPD in natural gas pipelines. This entailed the definition and parameterization of node states within a modified version of the BN presented in the Q3 quarterly report. The UMD team used a wide range of data sources and parameterization techniques to determine the conditional probability tables (CPTs) and functions that define the causal relationships of the BN model's nodes and their states. The selection of data sources, including damage incident reports and expert-elicited likelihoods, was performed based on specific criteria including data completeness and relevance to TPD. This data was then processed for input into the model. The parameterization of the nodes fell into two overarching methods: parameterization by elicitation or by a Bayesian updating procedure. The specific parameterization approach used varied based on whether the node's value was from a continuous or a discrete value set, the node's parents, and availability of datasets capturing the specific conditional relationships between the node and its parents.
   The result of the past two quarter's work is a functional comprehensive BN model that captures the causal relationships present between excavation practices and pipeline puncture failure resulting from hitting the pipeline. Due to the causal reasoning embedded in the model, situational awareness questions about the connection between certain preventive measures and pipeline damage can now be asked in order to inform decision-makers. This model will be the basis of a case study to further dive into questions of interest related to TPD.

3. This quarter, the Stanford team was tasked with identifying appropriate techniques and preliminary results using machine learning, data discovery, econometrics, or causal inference with the rich collection of utility and third party data we assembled in previous quarters.
   This quarter's analysis has focused on characterizing the potential for decision-relevant insights from satellite-based interferometric synthetic aperture radar (InSAR) data on ground movement for natural gas distribution integrity management. We have received InSAR ground movement data for the entire San Francisco Bay Area and are

combining it with the entire asset and leak database from Pacific Gas & Electric Co. (PG&E).

Our initial analysis last quarter found no clear evidence of a direct spatial correlation between even quite pronounced ground movement and natural gas distribution leaks within a four-year period. There have also been significant changes in PG&E's leak categorization and leak survey frequency, particularly since 2008, limiting the utility of the leak database, which goes back as far as 1970, for detailed statistical analysis of longer-term trends in leak incidence rates. Thus, traditional machine learning and econometric tools are of limited utility in this instance.

This quarter's report focuses on physical causal models as an alternative mechanism through which InSAR data can inform distribution integrity management. Even modest ground movement can induce bending moments in service lines, which can dramatically shorten the lifetimes of plastic pipes, particularly through slow crack growth.

This process is well understood, but calibrating the lifetime prediction equations requires significant information about both the material characteristics and baseline stress state of the pipeline network as well as a sufficiently resolved soil-pipeline interaction model.

This modeling approach would likely require additional data collection for calibration before it can produce actionable insights. We will further characterize precise data requirements in future quarters, but a modest number of in situ pipeline material samples, of order 20 or more, would likely be necessary.

Such a component-specific physical causal model can explain only a subset of all leaks. In this case, cracks in the body of plastic pipe represent 3% of all Grade 1 leaks. However, an ensemble of such models, many of which could likely be calibrated using much of the same data, could dramatically improve the situational awareness of distribution integrity managers, enabling proactive maintenance decisions that improve safety and lower emissions while reducing cost.

4. The causal modeling and situational awareness frameworks developed in previous quarters are shown to be fully capable of addressing nuanced aspects of decision support. This conclusion was reached by reviewing core constructs of situational awareness and critiques of the carious approaches in the literature. The framework developed in this project appears to be robust in the face of these elements and critiques. The causal modeling and situational awareness frameworks are now complete and ready to be applied to case studies that will be developed over the remaining four quarters of the project in consultation with the technical advisory committee,

5: Project Schedule

- Project is on schedule and budget.
- An update meeting with the TAP will be scheduled on the first convenient date in January 2021.

<u>End of report summary – Attachments to follow</u>

Attachments

1. Attachment 1: Future Task Work Overview
2. Attachment 2: Task 3.2 - Decision Support System: Working prototype decision-support tool with initial documentation
3. Attachment 3: Tasks 4.1.3 and 4.1.4 - Data Integration, Normalization and Uncertainty Tools: Bayesian Network Methodology for post processing data for BN, and Methodology for post processing data for BN
4. Attachment 4: Task 4.2.4 - Data Integration, Normalization and Uncertainty Tools: Data Driven Insights: Causal Inference – identify data combinations that are best dealt with by causal modeling
5. Attachment 5: Task 4.3.2- Data Integration, Normalization and Uncertainty Tools: Causal Modeling and Situational Awareness: Situational awareness framework

Attachment 1: Future Task Work Overview

The work scheduled for the 6th Quarter is presented below:

Deliverable	Complete	Responsible	Quarter
Summary report from case study application of the decision support tool, including a list of proposed enhancements	No	Lumina	Q6
Data Driven Insights:Econometric Inference – apply causal models to econometric inference	No	Stanford	Q6
Causal Modeling and Situational Awareness: SA Case study progress report 1	No	GTI	Q6
Sixth Quarterly Report	No	GTI	Q6

End of Attachment 1