Using Text Analytics Solutions with Small to Medium Sized Manufacturers: Lessons Learned

Michael Brundage[1*] and Radu Pavel[2]
[1] National Institute of Standards and Technology Gaithersburg, MD, USA
[2] TechSolve Inc. Cincinnati, OH, USA
Michael.brundage@nist.gov, pavel@TechSolve.org

Abstract

As Smart Manufacturing becomes more prevalent throughout industry, manufacturers are continuing to look for ways to more efficiently apply advanced data analysis methods to improve their decision processes. One promising area for improving decision making is through the use of natural language processing (NLP) methods on text-based data in maintenance. Maintenance personnel often capture important information on the problems and repairs throughout the manufacturing facility in informal text. This information is key to improving maintenance decisions, such as scheduling, dispatching, diagnosis, and inventory management, but is difficult to access due to the informal and domain specific nature of the text. Methods are available to aid manufacturers with parsing through this information, however small-to-medium sized manufacturers (SMMs) still have issues in implementing NLP solutions in practice. To this end, this paper discusses lessons learned in applying a NIST developed methodology to SMMs maintenance data.

1 Introduction

Within a manufacturing facility, maintenance logs that capture repair information, e.g., the problem, the solution, or the cause, are often completed by various operators or maintenance technicians. These technicians and operators often do not follow a set terminology or structure when entering this information. These inconsistencies in entering data even occur when only one person captures such information, such as when a manager enters all maintenance logs into a database. Due to such data logging inconsistencies, it is often difficult to observe or discover patterns or actionable information, particularly when a supervisor that is not directly involved with the maintenance process is reviewing the maintenance logs.

NIST researchers have developed technology using text analytics that has the ability to address this deficiency through its ability to assign tags, identify patterns, and extract actionable information from industrial data logs (Sexton Nestor, 2019). This methodology and subsequent analysis techniques have been developed for some time (Sexton, 2018; Sharp, 2019; Brundage, 2018; Sexton, 2017, Sharp, 2017). The software is open-source and available on GitHub[†] for all to use. Currently, the software helps maintainers annotate their Maintenance Work Order (MWO) data through a process called "tagging". The MWOs are inputted as comma-separated variable (.csv) files with UTF-8 encoding into the Nestor GUI and the user goes through the tagging process to create an annotated, tagged MWO dataset.

[*] Corresponding Author: michael.brundage@nist.gov
[†] https://www.nist.gov/services-resources/software/nestor

Considering the potential of the technology, NIST works with industry to further refine and improve their solution through assessment trials with data from manufacturers that can help reveal opportunities for improvement both in technology efficiency and in robustness of its applicability. TechSolve is working with NIST to assess the capabilities of the technology using maintenance data from manufacturing organizations willing to learn more about the potential advantages and suitability of such technology for the annotation, organization and analysis of their maintenance work orders/logs.

2 Data Collection Process

TechSolve leveraged its network of manufacturers to identify and recruit companies considered good candidates for this effort. A list of companies was compiled and readied for engagements starting in January 2019. Twenty seven (27) companies were contacted and assessed. Due to confidentiality constraints, the name of the companies cannot be disclosed. However, a list of their NACIS (North American Industry Classification System) codes and main characteristics is provided in Table 1, below.

Table 1. NAICS code, approximate number of employees, and the annual revenue for the companies contacted during this project (companies listed in random order)

No	NAICS Code	Employees	Annual Sales	Notes
Company 1	332119 - Metal Crown, Closure, and Other Metal Stamping (except Automotive)	60	$19M	Provided data
Company 2	336350 - Motor Vehicle Transmission and Power Train Parts Mfg	189	$37M	Provided data
Company 3	326199 - All Other Plastics Product Mfg 333514 - Special Die and Tool, Die Set, Jig, and Fixture Mfg	42	$10M	Declined to provide data

Company 4	332111 - Iron and Steel Forging	16	$4.2M	Declined to provide data
Company 5	332710 - Machine Shops	30	$629K	Declined to provide data
Company 6	333111 - Farm Machinery and Equipment Mfg	72	$42M	Provided data
Company 7	442299 - All Other Home Furnishings Stores	10	$1.5M	No electronic files
Company 8	334413 - Semiconductor and Related Device Mfg	142	$48M	Concerned with trade secrets/confidentiality
Company 9	311612 - Meat Processed from Carcasses	360	$25M	Expressed interest but no follow-up from company
Company 10	332710 - Machine Shops (Primary)	30	$6.1M	No follow-up from company
Company 11	322211 - Corrugated and Solid Fiber Box Manufacturing (Primary)	15	$5.9M	Declined due to limited availability of data
Company 12	333413 - Industrial and Commercial Fan and Blower and Air Purification Equipment Manufacturing (Primary)	31	N/A	Did not express interest in the opportunity

Company 13	335999 - All Other Miscellaneous Electrical Equipment and Component Manufacturing (Primary)	127	$25M	Management did not consider they have significant equipment and associated maintenance to qualify for this project
Company 14	423830 - Industrial Machinery and Equipment Merchant Wholesalers (Primary)	200	$1.7M	Management did not want to pursue opportunity
Company 15	333618 - Other Engine Equipment Manufacturing (Primary)	13	N/A	Did not express interest in the opportunity
Company 16	333249 - Other Industrial Machinery Manufacturing (Primary)	6	N/A	Management admitted they do not yet collect data in electronic format
Company 17	336390 - Other Motor Vehicle Parts Manufacturing (Primary)	2	N/A	Too small; limited maintenance necessary
Company 18	332710 - Machine Shops (Primary)	235	$39M	Very slow to reply. Too busy to commit for opportunity
Company 19	333922 - Conveyor and Conveying Equipment Manufacturing (Primary)	800	$800 M	Did not express interest in the opportunity

Company 20	331524 - Aluminum Foundries (except Die-Casting) (Primary)	18	$2M	Committed to send data but stopped short of sending a file
Company 21	332812 - Metal Coating, Engraving (except Jewelry and Silverware), and Allied Services to Manufacturers (Primary)	21	$2.8M	Did not express interest in the opportunity
Company 22	811310 - Commercial and Industrial Machinery and Equipment (except Automotive and Electronic) Repair and Maintenance (Primary) 336390 - Other Motor Vehicle Parts Manufacturing (Secondary)	366	$76M	Did not express interest in the opportunity
Company 23	336412 - Aircraft Engine and Engine Parts Manufacturing (Primary)	100	$16M	Did not express interest in the opportunity
Company 24	333511 - Industrial Mold Manufacturing (Primary)	183	N/A	Expressed interest but declined sending files
Company 25	334418 - Printed Circuit Assembly (Electronic Assembly) Manufacturing (Primary)	170	$44M	Did not express interest in the opportunity

Company 26	332911 - Industrial Valve Manufacturing (Primary)	150	$50M	Expressed interest but declined sending files
Company 27	333912 - Air and Gas Compressor Manufacturing (Primary)	50	$1.3M	Expressed interest but declined sending files

Although some of the companies expressed interest in the program, they withheld from sharing data over confidentiality and trade secret concerns. Other companies specified that they did not collect data, although they are interested to implement "best practices" and appropriate software solutions, such as computerized maintenance management systems (CMMS) or enterprise resource planning (ERP) systems. Such companies expressed the need for help in identifying those "best practices" and appropriate software, and mentioned that they had difficulties identifying a solution suited for them due to lack of knowledge in the field. In other cases, the companies were collecting maintenance information but could only output it in a printed form and were unable to export files in excel or .csv file format.

3 Lessons Learned

The companies collecting data in electronic format typically used three types of software: 1) a non-maintenance specific database (e.g., access or excel), 2) a computerized maintenance management system (CMMS) (e.g., Fiix), or 3) more generic planning system, such as the Enterprise Resource Planning (ERP) system (e.g., Plex). The companies looking to upgrade their maintenance work order capturing routine to an electronic platform, expressed interest in best practices and available/recommended solutions on the market – e.g. what would be the criteria to choose a good system for us? What system would be best for us? The companies that were already in possession of a software platform were interested to know if their practices and the way they are collecting the information are aligned with best practices. In addition, the manufacturers were interested in what would be more efficient and relevant analytics and charting for the maintenance work order they collect.

Concerning the engagement with industry, it was found that manufacturers recognize that the health and maintenance of the manufacturing assets represent an important area of their operations. The importance of maintenance appeared to be directly proportional with the size of the company and the cost of the product being manufactured. Nevertheless, from the contacted manufacturers that were engaged in communications with TechSolve, approximately 30% did not seem to have computerized means of capturing the maintenance work order data. This has been justified either from the perspective of the size of the company (too small), or the limited complexity of the equipment (e.g. conveyors or welding equipment). A limited number of companies, approximately three, confirmed they are collecting maintenance work order data and expressed interest in providing data; however, they were unable to export the data in csv or excel format. The main findings of the interactions with the companies that were contacted are summarized below:

- Approximately 75% of the companies that were contacted for this initiative expressed interest to learn more. However, only 50% of the interested companies moved forward with phone conversations or in-person visits. Eventually, from all contacted companies, only

seven expressed intent to provide files, of which only three provided files eventually. From the three organizations that provided files, one was very concerned with the confidentiality of the data to the point that all operator names, asset names, and their locations had to be coded/changed.

- The majority of the companies expressed concern with and asked for maintaining the confidentiality of the information. If data sharing would be desired and further publication of the results, then the data should be stripped of identifiers and the provider should approve its release before the publication of the data and/or of the results.
- The companies compliant with ISO 9001 and AS9100 were more likely to have maintenance work order data.
- The companies that have maintenance records typically use a CMMS or ERP system to capture the information, and the work orders are logged in a database.
- Concerning the use of maintenance work order data, some companies appeared to have software with various capabilities to generate graphs or run statistics. However, the full functionality of the software, or the actual use of the software capabilities was not presented to the TechSolve team. Nevertheless, all companies expressed the desire to get better analytics and ways of visualizing data that would allow them to better understand the maintenance activities and extract actionable information.
- Long term, the companies expressed interest in implementing monitoring systems that would enable condition-based maintenance approach, versus reactive or preventative approaches.
- With regard to the maintenance work order data, all companies expressed interest in a solution that would help them better organize that data, and were interested to learn more about NIST's efforts on guidelines, standardization, and technology development addressing the manufacturing assets maintenance
- Due to the variety of systems used to collect data, the files shown had various column headers. Although only a limited number of files have been provided, the sample covers the typical scenarios discussed with industrials that span from custom made spreadsheets with small number of columns in an Excel file or Access database or using CMMS files with very large number of columns
- The variety of data collection format or the confidentiality restrictions imposed initial organization and filtering of the data files to enable proper processing with the NESTOR software and sharing the information with the NIST team.

After the examination of the maintenance work order files provided by the manufacturers, the following observations became apparent:

- Each company seem to collect data in its own, custom way, based on internal needs and guidance from the software provider; however, no particular "best practices" were pointed out or noticed.
- The names of the columns describing the maintenance task and/or resolution was different across the processed files.
- There often are no accurate records of the actual time it took to repair one item.
- There is limited information of who noticed the fault and who repaired the fault.
- Some descriptions are too simplistic, others may only be understood by someone that is very familiar with the manufacturing asset.
- It is needs to be clear to what extent the information is used for analysis and potential improvement opportunities. In general, the users would like to be able to derive (with simplicity) additional analytics/charting facilitating actionable information.

4 Conclusions and Future Work

This paper discusses lessons learned with SMMs for implementing a text analytics solution for analyzing maintenance work orders. The biggest concern of the manufacturers was providing proprietary information for analysis, thus, anonymization methods are important to improve the overall text analytics process. Most manufacturers that collected data had analytics and visualizations, but wanted more intuitive tools. Lastly, these manufacturers expressed interest at more predictive capabilities for discovering maintenance needs in the future and overall, the manufacturers involved in this initiative expressed interest in efforts associated with PHM for manufacturing assets.

The manufacturers involved in this study agreed that it would be very helpful to have guidance on best practices and product selection criteria with regard to capturing and processing maintenance work order data. Standards in the space of text analytics for manufacturers are needed to aid manufacturers in performing this analysis themselves (Weiss, 2019; Sexton Standards, 2019). The natural language processing concept and the availability of a technology to be used for organizing and annotating their data was regarded positively.

NIST Disclaimer

The use of any products described in this paper does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that products are necessarily the best available for the purpose.

References

Sexton, Thurston B., and Michael P. Brundage. Nestor: A Tool for Natural Language Annotation of Short Texts. No. Journal of Research of the National Institute of Standards and Technology. 2019.

Sexton, Thurston, et al. "Benchmarking for keyword extraction methodologies in maintenance work orders." Proceedings of the Annual Conference of the PHM Society. Vol. 10. No. 1. 2018.

Sharp, Michael, et al. "Selecting Optimal Data for Creating Informed Maintenance Decisions in a Manufacturing Environment." Model-Based Enterprise Summit 2019. 2019.

Brundage, Michael P., et al. "Developing maintenance key performance indicators from maintenance work order data." ASME 2018 13th International Manufacturing Science and Engineering Conference. American Society of Mechanical Engineers Digital Collection, 2018.

Sexton, Thurston, et al. "Hybrid datafication of maintenance logs from ai-assisted human tags." 2017 IEEE International Conference on Big Data (Big Data). IEEE, 2017.

Sharp, Michael, Thurston Sexton, and Michael P. Brundage. "Toward semi-autonomous information." IFIP International Conference on Advances in Production Management Systems. Springer, Cham, 2017.

Weiss, Brian A., et al. Summary Report on the Industry Forum for Monitoring, Diagnostics, and Prognostics for Manufacturing Operations. No. Advanced Manufacturing Series (NIST AMS)-100-23. 2019.

Sexton, Thurston, and Michael P. Brundage. "Standards Needs for Maintenance Work Order Analysis in Manufacturing." Model-Based Enterprise Summit 2019. 2019.