Available online at www.sciencedirect.com

ScienceDirect

Procedia CIRP 00 (2019) 000–000

www.elsevier.com/locate/procedia

53rd CIRP Conference on Manufacturing Systems

The Ethical Use of Human Data for Smart Manufacturing: An Analysis and Discussion

Harley Oliff[a], Ying Liu[a*], Maneesh Kumar[b], Michael Williams[c]

[a]School of Engineering, Cardiff University, Queens Buildings, CF22 3AA, UK
[b]Cardiff Business School, 3 Colum Drive, Cardiff, CF10 3EU
[c]Olympus Surgical Technologies Europe, Cardiff, CF3 OLT, UK

* Corresponding author. Tel.: +44(0)29-20874696. *E-mail address:* LiuY81@cardiff.ac.uk

Abstract

Much of the recent work in the field of smart manufacturing is dependent on a data-driven approach, with an increasing number of data-hungry techniques being introduced to improve the adaptability and productivity of manufacturing systems.
More recently, the collection and use of data generated by employees and individuals has become the source of controversy at both a corporate and societal level. This work presents an analysis of current thoughts and considerations on the use and misuse of data from a number of standpoints and discusses a methodology that enables appropriate parameterization of human performance for use in modelling and simulation for smart manufacturing.

Keywords: *Intelligent Manufacturing; Data Ethics; Human-Robot-Interaction*

1. Introduction

The motivation for this work stems for attempting to utilise machine learning within manufacturing to predict and consequently adapt, to variations in human task performance. During this process, capturing the intricacies and influence of dependent factors on this task performance was a critical issue, to provide a dataset upon which a variety of learning algorithms could be trained, to identify the patterns of variation. This is no easy task, as human performance variation is subject to an almost limitless number of external and internal factors, and the degree of susceptibility to these factors varies between individuals. The constant development of ever more detailed datasets and the continuous monitoring of manufacturing tasks invites with it many questions about the fair and ethical use of this data; and importantly, the acceptable limits after which such monitoring may be considered an invasion of privacy. Such concerns are amplified, such as in the authors' case, where the data collection is directly personal and identifiable, as opposed to monitoring average outputs and statistics for whole teams, as it introduces additional responsibility and culpability for less than ideal performance, which may not be attributable to the individual. There are two conflicting arguments. On the one hand, the employer is entitled to collect and process whatever data they deem relevant to their process and workings. This is necessary not just for the benefit of the employers and the process, but to ensure the employees are being enabled to perform their work successfully also.

The collection and compilation of the human task performance data at the individual level highlights the other side of this argument, however, as these datasets enable employees to be profiled and accountable for their actions at any point, potentially retroactively, in light of changing rules, context or circumstance. It introduces the potential for biases, and decisions made on incomplete or incorrect knowledge, in cases where the data is used partially, or out of context.

The degree to which employers have a right to monitor their employees vs the right employees have to privacy is an ongoing argument, as evidenced by an extensive body of literature[1, 2]. Such a conflict is heightened further, however, by the pervasiveness of data-driven systems evidenced by the conversations occurring regarding the current global state of data collection and personal privacy in society, bringing attention to the ethical and legal responsibilities of corporations and governments with respect to these points. These questions themselves raise further questions when the scenario becomes extrapolated to a future where they remain unaddressed.

This work aims to highlight the potential issues that the collection and analysis of such data presents and to encourage discussion of these points amongst researchers and interested parties moving forward; so that contingencies may be made to prevent problems such as those identified.

2. Literature Analysis

The following section discusses the current state of data in the modern world. The information age has, as all technological revolutions before it, brought with it several issues alongside the capacity to achieve previously impossible things. Data use has become so pervasive, and essential to the function of modern society, that it is inconceivable to imagine a world without it, and yet the social, political and economic effects and opportunities of these emergent domains remain poorly understood and thus prime for exploitation.

2.1. Ethical Data use

The information age has opened up a Pandora's box of opportunity for the creation of wealth based on data. In an age where everything is instantaneous and accessible, people have only just begun to realise the fact that: if you're not paying for the product, you are the product. Understanding the fact that personal data exists as a product in such a way enables parallels to be drawn with traditional capitalist models, to examine how to move forward. Whilst far from the only voice on the matter, In her comprehensive work 'The Age of *Surveillance Capitalism'*[3, 4], Dr.Shosanna Zuboff defines this current climate of information capitalism as*: "A new economic order that claims human experience as free-market raw materials for hidden commercial practice of extradition, prediction, and sales."*. The work discusses at length the pervasive nature of surveillance in the 21st century, and how the collection and analysis of personal data enables new economic logic and opportunity based on behavioural prediction and modification. The work ultimately suggests that these practices present as great a threat to human nature itself, as industrial capitalism does to the environment.

As always, however, there are two sides to the story. This paradigm has improved individual lives in many ways, be it via real-time traffic data and understanding of human dynamics extracted from mobile phone locations[5, 6]), or algorithmically generated shopping and media recommendations based on individual behaviours and preferences [7], the use of human-data as a resource has its unarguable benefits. It has also been leveraged to a huge extent by industry, not only through the refinement of existing processes, but through the creation of entirely new, innovative revenue streams[8]. The issue is that industry and consumer behaviour is now governed by these paradigms. If Netflix was unable to recommend to you content that you were likely to enjoy based on what you have previously enjoyed, it stands to reason that its success would have been much more modest. Such an effect would have reduced the consequent cultural impact, for better or worse, and demonstrates how in many ways, society is defined at this point in time by such practices.

The use of individual data to provide these services has, however, reached a tipping point, and a level of invasion that feels noticeably uncomfortable to many. Companies now analyze data picked up by the microphone inside mobile phones, and through the use of Natural-Language-Processing (NLP), extract key information which is then sold to third parties for advertising purposes. This has been commented on increasingly in recent months [24, 25], with the main issues raised that, although not malicious, there is something inherently unsettling about constant and invisible monitoring.

Fair and ethical data use, and the rights of individuals to privacy is currently a globally reaching, and controversial topic at the heart of the political and societal debate. Such discussions have been taking place for several years, but it is in light of more recent, large scale events, such as the Cambridge Analytica scandal, brought to light as the result of whistle-blowing from within the company itself [9, 10] The information that emerged brought to the public information sphere the first real notions of how their data may be used for purposes that ultimately have very real, and very serious implications. The investigation into Cambridge's practices have made some associations between their activities and high profile political events, including Trump's 2016 election victory and the highly controversial Brexit Referendum the same year, the effects of which are still being felt by the American, European, and British societies and its economies, at all levels. These cases ultimately developed public awareness of just how pervasive the monitoring and profiling of individuals for commercial use has become, and importantly how targeting individuals based on susceptibility to advertising, enabling behaviours to be manipulated at scale.

The issues with personal monitoring extend beyond the social sphere into almost all areas of life, and increasingly into areas of industry. Issues with the way employers monitor their employees can already be seen to be emerging, one notable example being the US trucking industry. The work of Dr.Levy [1, 11] into the use of Electronic Logging Devices (ELD's) in place of conventional methods based on an analogue tachometer and logbook, to monitor the work schedule and task completion.

In brief, truckers in the US, and indeed many countries around the world, are governed by strict working regulations regarding time spend driving versus sleeping, although such regulations are particularly acute in the US, where distances are vast, and shifts often last several days. American truckers often live within the confines of their vehicles, travelling with family and pets, and as a consequence, are affected to a far greater degree by workplace surveillance than most employees. Truckers to whom these new regulations apply have been very vocal regarding the changes, and they raise some interesting points regarding the efficacy of monitoring and data-based enforcement. Firstly, they argue that the devices restrict their autonomy, and in many cases introduce inefficiencies into their working patterns, and all parties involved are worse off, the driver, their employer, and the end consumer of the goods. Dr Levy's work and the vocal criticisms of the truck drivers themselves illustrate another clear issue with job surveillance and the issues with making decisions using this information. It is clear that the ELD's, whilst well-motivated, fall victim to the same trap that most data-based systems do; which is to fail to account for the situational context [12].

The issues presented are a small selection of examples, and the collision of these ideologies is increasingly pervasive. What it is intended to highlight, is the division (and strong arguments which exist on both sides) between the human drive to improve and the human desire for privacy which exists in the information age. It is perhaps ironic to present this as a human dilemma, but it stands to reason that both homo sapiens have a natural proficiency with abstraction and analysis of patterns within data, which, through the development of numbers is responsible for the evolution and function of modern society. Consequently, it could be said that the natural progression of society further hinges on data and the insights it provides[13].

2.2. Data Within manufacturing

The concepts of performance monitoring to improve productivity within manufacturing stem from Frederick Taylor's investigations into manufacturing processes at the turn of the 20th century. This work led to the timings and empirical studies of thousands of manual operations, across a number of tasks. Analysis of the data collected from this study ultimately enabled Taylor to formulate many of the early ideas of maximising productivity, and analysis of manufacturing operations as a process, formulates as his theory of *Scientific Management* [14, 15].

The principles of Scientific Management became pervasive throughout manufacturing and introduced monitoring, and abstraction of these processes into models and databases. As a consequence, the monitoring of manufacturing processes has continually increased, as insights have been sought at increasingly high resolutions. This has particular relevance in the age of big data [16-18], which for the past decade has been a focus of how many manufacturing operations view data capture and storage. The idea that data will enable the generation of insight is now pervasive within modern manufacturing processes and forms the basis of intelligent manufacturing to make use of this accumulated information. Intelligent manufacturing encompasses a collection of techniques and methodologies to leverage the knowledge which may be obtained from the processing of this data, and such techniques have become pervasive within modern applications (Lee, Bagheri et al. 2016). The use of human data is critical in enabling many emerging technologies, in particular, Cyber-Physical-Systems (CPS's) based approaches to digitalization. CPS's are reliant on modelling and monitoring of real-world systems for prediction and analysis, and consequently, capture and profile human performance data [19-21].

These systems are beholden unto this new paradigm of constant surveillance and are capable of reading, processing, and storing sensor data with frequencies of less than a second. As such, they are capable of capturing data at almost any desired resolution. Based on this, data for any required temporal scale can be retrieved and processed as required by the application. This data may be captured either directly through deliberate measurement, or indirectly as a result of other sensor data, e.g. difference in timestamps between sensor activations is equivalent to the cycle time of the set of operations causing the activations.

Thus, in light of the justifiable concern over data usage that has been raised in a wider context, appropriate consideration must now be given by the manufacturing industry as a whole, to the role they play, and the appropriate manner in which to handle both the existing data and generated data in the future. A reduction in the scope of the issue to just the manufacturing industry, unfortunately, does not bring with it a simplification of the problem. The issues of influence that emerge at the macro scale can be seen again to resolve themselves within many smaller systems.

2.3. Identifiable problem

The data-driven economy promises to thoroughly transform almost all aspects of our lives. Providing more efficient and sustainable processes, seamless interactions with machines and intelligent systems, improved infrastructure, transport, and medical care, greater choice and customization, and entirely new products and experiences which such data analysis will make possible. All of these benefits rely however on the careful collection curation and analysis of personal data, and as discussed, larger questions are now beginning to emerge for society at large to address. The issue to be overcome is the conflict of two ideals. The arguments for data collection to necessitate efficiency and productivity, and the counterargument regarding the ethical extent of individual surveillance and profiling based on data.

Considering manual operations, the data generated by this monitoring is tied to the individual performing the manual operation and contains a great deal of information. These types of operations are particularly sensitive to time monitoring. At sufficiently high fidelity and frequency, such as those found in CPS's, a continuous profile of uncontextualized data will exist with which to value and evaluate the performance of each identifiable individual.

Minimal efforts are made within these systems to understand the variations between individuals or conditions, the result of multiple *human factors* that influence task performance, and which may propagate errors through simulations and digital models of such processes, consequently affecting the decision-making accuracy and performance of the real-world process.

Whilst there is an argument that this is necessary for certain learning-based and digital modelling-based approaches to operate effectively, it also opens the door to abuse of this information, either deliberately by malicious actors, or through unknown biases. It is also easy to see how scenarios in which decisions may be made inefficiently based on what is not included in the dataset.

Where learning algorithms are employed to make decisions autonomously, such data could be processed in such a way that biases individuals based on metrics it considers important or is otherwise unintentional on the developers' behalf. In the case where cycle time is the only metric for evaluation, the system may decide that one operator's performance is sub-optimal in terms of speed compared to another, and penalize them, without considering that the first individual makes more mistakes, introducing disturbances and consequently has a lower overall productivity.

There is no process to validate such systems or to be sure that they are making evaluations based on models with a degree of accurate representation of the real-world.

The next section outlines the authors plans to effectively capture variation in human performance based on the individuals involved, in such a way that enables simulation to exist as a valid solution, but which is non-reliant on constant monitoring and resulting in no lasting profile.

3. Proposed Methodology for Anonymization

In light of these issues, there remains a question to be overcome. The authors believe that whilst far from comprehensive, a methodology to profile and understand the on-task performance of human beings within a typical manufacturing scenario and enable effective modelling of these individuals, without the dependence on constant surveillance. The methodology seeks to extract a representative set of parameters from an initial monitoring period, which may be used to represent individuals within a simulated process.

As discussed in the previous section, within manufacturing and system analysis, the task performance is frequently profiled using the cycle time. As such, the methodology aims to capture the variation in this metric along several vectors of variation.

Intuitively, it stands to reason that for any well designed and optimized process aiming for a one-piece-flow, there is likely to exist a much harder lower limit for performance measured in terms of duration, as the number of factors which may slow down process far outweighs those which remain to speed it up. As such, it is hypothesized that the probability distribution for a given cycle time being observed will likely be skewed towards the faster end; i.e. more observations will be made of a well-optimised process working efficiently than will be made of events leading to disturbances. This has a significant consequence in terms of how performance for the operators is averaged and modelled, as purely mean values may misrepresent the observed behaviours when simulated.

As such, establishing the appropriate distribution of cycle time variability is critical for effective human modelling. This must be achieved in terms of both 'piece-to-piece' variation, and the variation introduced by changes in context, such as those influencing fatigue.

Once the appropriate parameters have been generated from the collected data, a distribution may be built and sampled to provide estimates for digital models. These parameters may be derived from any number of the data's dimensions, dependent on whether temporal or individual factors provide a greater magnitude in the variability. Analysis of production processes in this way is of particular relevance in the case where one-piece-flow is the aim, as the process is susceptible to variation. Analysis of these systems and understanding variation and its impact is often done through both throughput and WIP, which is ultimately representative of additional cost [23]. The following section presents an application of this methodology for adequately capturing human performance data over a variable set of timescales, in order to demonstrate and enable the adequate parameterization of digital manufacturing models.

4. Application to Real-World example

A set of task observations was made of a set of sequential, manual assembly processes to illustrate how to effectively manipulate such a dataset to preserve information adequately, whilst also providing sufficient anonymization for the employees in question. The production line was designed to operate as close to a condition of One-Piece-Flow (OPF) as possible and provides a source for data with a high number of observation points and frequency. Each sub-task is easily delineated, and as such easily monitored in terms of the cycle time at each position, which in turn may be monitored for each individual product produced. Such data collection may enable very accurate modelling and simulation of the production process, but by consequence generates profiles of employee performance which may be misinterpreted without the relevant context. For instance, it is difficult to determine the causality of any effects which may be observed, as sequential processes will aid in the propagation of disturbance.

The assembly process is similar to many automated layouts, but interactions occur between humans. This both facilitates collaboration between operators, but also makes the introduction of disturbances increasingly likely, due to the semi-predictable human element. Additionally, however, such a process allows many different examples of human performance may be collected, and for the dataset to be processed accordingly. Cycle Time measurements were performed for each of the process positions. Five measurements were taken at each interval with each interval spaced two hours apart. This process was repeated across two days. Additionally, the measurements were performed across two work shifts, effectively doubling the number of operators profiled. The data processing arguably enables further insight into the process's strengths and weaknesses, as trends may be identifiable when the data is aggregated by task, day, time, or individual. Visualization of the dataset by way of a heatmap aids in understanding the distribution, as does plotting the raw data values as a series for each operator, in addition to the Mean, Std Deviation, and Variance values for each operator, overall conditions. These are illustrated in Figure.1.

	PM				AM			
Operator	1	2	3	4	5	6	7	8
WEDNESDAY	16.36	11.75	16.59	15.40	12.93	14.92	10.42	16.02
	14.10	11.81	16.28	15.90	12.79	13.88	8.60	16.63
	15.18	13.07	15.47	15.74	13.10	15.90	11.60	16.01
	14.54	13.00	18.92	15.22	16.60	16.39	9.18	16.59
	13.42	13.97	16.58	14.84	12.91	13.49	8.76	15.12
	12.56	18.78	13.66	14.74	13.62	14.85	18.89	13.47
	11.21	15.45	17.48	15.50	13.57	13.67	21.15	15.19
	11.39	16.69	15.17	21.14	12.97	14.55	15.03	13.48
	11.05	13.37	18.77	14.75	19.33	14.46	16.45	13.82
	11.43	13.22	17.94	14.44	15.68	14.92	17.27	15.74
THURSDAY	13.90	11.92	14.50	18.74	15.12	15.13	13.97	17.93
	11.55	12.38	18.37	15.76	13.97	15.30	18.17	15.38
	13.67	16.17	13.69	15.54	15.03	14.78	10.90	16.19
	13.09	10.79	15.47	15.30	13.92	14.57	12.05	13.77
	17.31	13.17	17.48	16.07	17.39	13.80	12.90	13.73
	12.19	14.22	16.26	14.05	12.05	22.63	14.72	14.02
	12.74	14.74	12.37	14.56	15.75	13.10	15.81	12.87
	14.10	9.44	15.25	15.41	13.79	11.67	15.20	14.24
	14.06	11.87	15.99	15.93	16.60	10.61	14.15	14.50
	13.01	12.89	15.95	16.49	13.70	13.70	12.44	16.00
Mean	13.34	13.44	16.11	15.78	14.54	14.62	13.88	15.04
Std dv	1.70	2.16	1.76	1.60	1.86	2.31	3.48	1.35
Variance	2.88	4.65	3.08	2.55	3.46	5.33	12.14	1.81

Figure.1. Heatmap illustrating the distribution of cycle time, and mean, std dev, and variance, for all operators, across all conditions

Furthermore, the probability distribution can be extrapolated from the dataset by visualizing the data as a histogram to give an indication of its form, to validate the assumptions made in the methodology regarding distribution, and to identify appropriate parameters. The histogram illustrated in Figure.2 enables several conclusions to be drawn about the nature of task-variability, and how it may be parameterized.

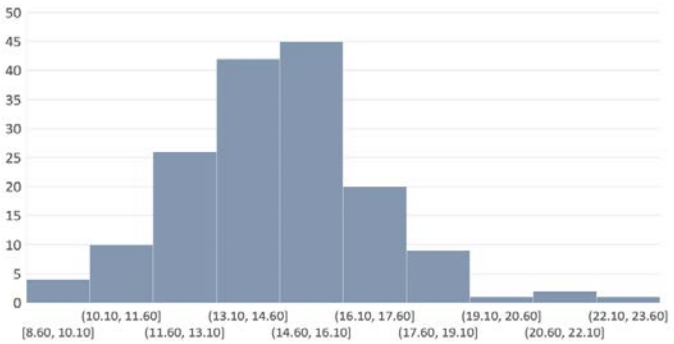

Figure.2. Histogram showing the distribution of cycle times for all operators across all conditions.

The distribution can be seen to follow an approximately normal distribution centered around a mean. This enables samples to be taken from a Gaussian distribution parameterized by a derived mean and standard deviation for each operator. Additionally, these parameters may be derived along whichever vector of variation required to adequately model the influence of human variation. To evaluate the efficacy of this, a simple simulation model was developed of the production line, using the Java-based Anylogic software[22]. The simulation model is typical of a Discrete-Event-Simulation (DES), often used to model and analyze production systems in terms of efficiency and productivity and consists of 4 elements representing different human operators. Each of these elements is characterized by a delay, which is sampled for each product passing through based on the parameterized distribution. The simulation model is illustrated in Figure.3.

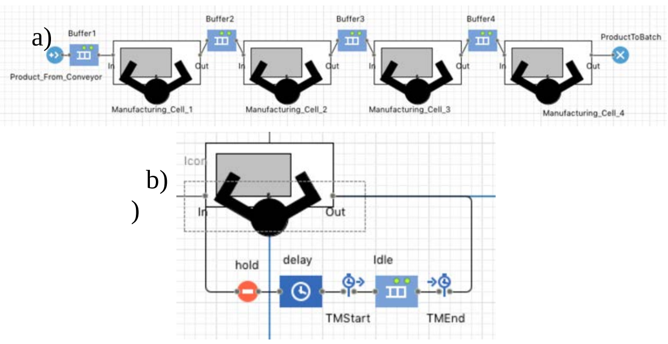

Figure.3 Simulation model developed at a) process level, and b) operator level.

As discussed in the methodology, the next step is to derive parameters from the dataset. This may be done along with any number of variation vectors, and enable the influence of different factors which influence human task performance and different contexts to be compared. The dataset contains four key vectors:

- Time of day AM vs PM
- Day of the Week (Wed vs Thurs)
- Individual Operator
- Individual Task

The dataset can be divided based on these vectors by including or excluding the relevant data points. The mean value and standard deviation are calculated from each subset of data and a separate distribution generated and sampled as discussed, the resulting parameters derived are illustrated in Figure.4.

Operator Position		Variation Vector/Combination of				
		Wednesday Samples	Thursday Samples	All AM Samples	All PM Samples	All Task Samples
1	Mean	13.74	14.15	14.54	13.34	13.94
	Std Dev	2.01	1.63	1.86	1.70	1.82
2	Mean	14.41	13.64	14.62	13.44	14.03
	Std Dev	1.68	2.71	2.31	2.16	2.25
3	Mean	15.21	14.78	13.88	16.11	15.00
	Std Dev	3.52	2.05	3.48	1.76	2.85
4	Mean	15.49	15.32	15.04	15.78	15.41
	Std Dev	1.84	1.00	1.35	1.60	1.50

Figure.4 Derived parameters for each of the variation vectors considered. Means and Std. Deviations used to parameterize a gaussian distribution.

Several iterations of this simulation model were performed, with variation vectors for both the time of day (AM vs PM), the day of the week (Wednesday vs Thursday) and the specific task application (average of all operators and conditions for the position). The results of the simulation runs, in terms of the impact on levels of WIP and overall productivity, are illustrated in Figures.5 and 6 respectively. The final set of iterations used parameterizations for each operator, enabling representative samples to be drawn without the need for constant direct data-driven models. In the example of this study, the manufacturing task was performed by two shifts, one morning, one afternoon. As such, the parameterizations based on the time of day vector, in this instance also represent individual operator parameterizations. Were the case that a single shift covered both the AM and PM samples, the per operation parameters would have been representative.

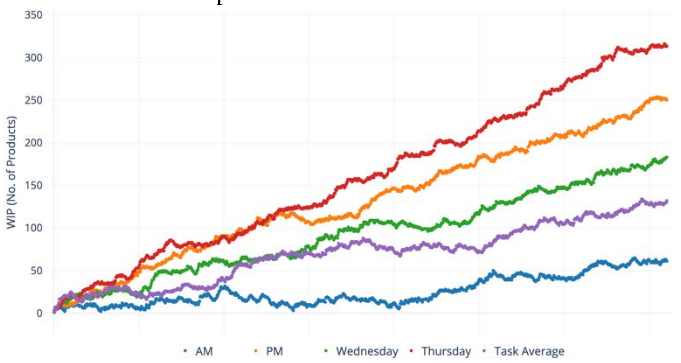

Figure.5 Influence of different human operator parameterizations on WIP levels. Vectors are based on time of day and day of the week and are compared with the average task performance.

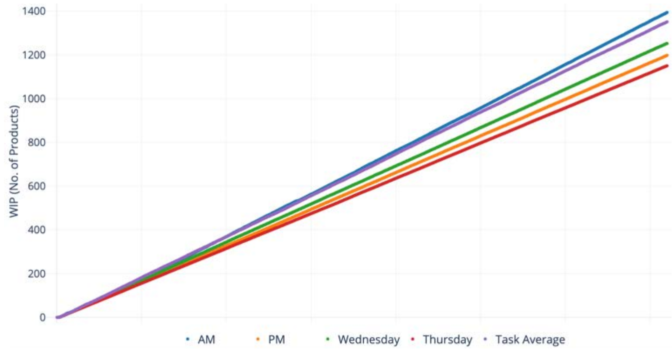

Figure.6 Influence of different human operator parameterizations on overall process productivity. Vectors are based on time of day and day of the week and are compared with the average task performance.

5. DISCUSSION

The presented work aims, to reiterate, not to provide a comprehensive solution to human modelling, but to promote the relevance of the question to the current thinking within intelligent manufacturing; as many of the emerging technologies are dependent on building potentially compromising datasets. The issues center on identifiability, traceability, and crucially, context.

To avoid the dependency on direct monitoring, or stored datasets the authors proposed approach suggests that a small volume of data may provide enough information to build useful profiles to enable appropriate variability within simulation and modelling The importance of this can be seen in Figure.1, which illustrates the variety in repeatability and the variation in process timings that result from human involvement, the result of the influence of human factors, in this case, the time of day, and day of the week. The impact of this variation and its relevance of process understanding and design is clear from Figure.5, which illustrates how small changes to the parameterization of these elements can influence the process.

The approach provides two important avenues to facilitate the application of emerging intelligent technologies, in a minimally invasive manner. Firstly, it enables accurate process models to be built and analyzed according to different factors to understand the overall impact on the process. This is highlighted by Figure.4 which is colour coded according to the magnitude. In all cases, the calculated parameters from all task samples, as expected, represent the average impact of all sources of variation. Secondly, the approach can provide an anonymized profile of each operator for use in these simulations, representing the combined impact of the sources of variation. Furthermore, as the dataset gathered is small, a 'sliding window' approach may be better able to track changes in behaviour temporally, whilst preserving the degree of privacy that aggregated monitoring in such a manner provides; this is certainly an area that should be of focus moving forward.

As the found distribution in human task performance contradicts the authors' logical hypothesis, the work supports the case that initial observations are certainly necessary to accurately establish the parameters of a process or operation. As such, there is a limit to how much cross-application can be achieved with any individual methodology or system for human modelling. This is further illustrated by the non-negligible difference in overall system performance when different parameterizations are used. The overall effect on the process of each variation vector will likely be application-specific, and some initial calibration of the methodology will be necessary in determining the appropriate distribution to sample in each case.

6. Conclusions

To conclude, technology has always threated societal norms and has been almost the sole driver of social change for the last several centuries. The current state of technological advancement, and what it forces us to ask about ourselves and the way we live, and the limits of what is acceptable, is not a new challenge, nor one that will be overcome when these questions are answered. Both sides of this divide have valid arguments, and as such, the next course of action must be to consider the extent to which the needs of business reliant on personal data collection remain legitimate; and to formalize the ethical boundaries which will govern both how this data is collected and processed.

This work, as stated, is intended to provoke these discussions, and outlines what the authors believe, is a step towards enabling the benefits of these advancements to be realized, in a way that preserves individuality and privacy. The authors appreciate the time and attention you have paid to this work and hope that it serves to plant the questions that it asks in your mind.

REFERENCES

[1] K. Martin and R. E. Freeman, "Some problems with employee monitoring," Journal of Business Ethics, vol. 43, no. 4, pp. 353-361, 2003.

[2] C. A. Ciocchetti, "The eavesdropping employer: a twenty‐first century framework for employee monitoring," American Business Law Journal, vol. 48, no. 2, pp. 285-369, 2011.

[3] S. Zuboff, "Big other: surveillance capitalism and the prospects of an information civilization," Journal of Information Technology, vol. 30, no. 1, pp. 75-89, 2015.

[4] S. Zuboff, The age of surveillance capitalism: The fight for a human future at the new frontier of power. Profile Books, 2019.

[5] S.-L. Shaw, M.-H. Tsou, and X. Ye, "Human dynamics in the mobile and big data era," International Journal of Geographical Information Science, vol. 30, no. 9, pp. 1687-1693, 2016.

[6] G. Guido, D. Rogano, A. Vitale, V. Astarita, and D. Festa, "Big data for public transportation: A DSS framework," in 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2017: IEEE, pp. 872-877.

[7] B. Hallinan and T. Striphas, "Recommended for you: The Netflix Prize and the production of algorithmic culture," New media & society, vol. 18, no. 1, pp. 117-137, 2016.

[8] N. Srnicek, Platform capitalism. John Wiley & Sons, 2017.

[9] C. Cadwalladr and E. Graham-Harrison, "The Cambridge analytica files," The Guardian, vol. 21, pp. 6-7, 2018.

[10] J. Isaak and M. J. Hanna, "User data privacy: Facebook, Cambridge Analytica, and privacy protection," Computer, vol. 51, no. 8, pp. 56-59, 2018.

[11] K. E. Levy, "The contexts of control: Information, power, and truck-driving work," The Information Society, vol. 31, no. 2, pp. 160-174, 2015.

[12] Vox, "<yt-formatted-string class="style-scope ytd-video-primary-info-renderer" force-default-style="" style="word-break: break-word;">How job surveillance is changing trucking in America</yt-formatted-string>," in Shift Change, ed. Youtube, 2017.

[13] Y. N. Harari, Sapiens: A brief history of humankind. Random House, 2014.

[14] F. W. Taylor, "The principles of scientific management New York," NY: Harper & Brothers, 1911.

[15] F. W. Taylor, Scientific management. Routledge, 2004.

[16] E. Auschitzky, M. Hammer, and A. Rajagopaul, "How big data can improve manufacturing," McKinsey & Company, 2014.

[17] R. F. Babiceanu and R. Seker, "Big Data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook," Computers in Industry, vol. 81, pp. 128-137, 2016.

[18] J. Gantz and D. Reinsel, "The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east," IDC iView: IDC Analyze the future, vol. 2007, no. 2012, pp. 1-16, 2012.

[19] J. Lee, B. Bagheri, and H.-A. Kao, "A cyber-physical systems architecture for industry 4.0-based manufacturing systems," Manufacturing Letters, vol. 3, pp. 18-23, 2015.

[20] L. Monostori et al., "Cyber-physical systems in manufacturing," CIRP Annals-Manufacturing Technology, vol. 65, no. 2, pp. 621-641, 2016.

[21] L. Wang and G. Wang, "Big Data in Cyber-Physical Systems, Digital Manufacturing and Industry 4.0," IJEM-International Journal of Engineering and Manufacturing (IJEM), vol. 6, no. 4, p. 1, 2016.

[22] Anylogic. (2018). [Online]. Available: https://www.anylogic.com/

[23] Hopp, W.J. and Spearman, M.L., 2011. Factory physics.

[24] Kleinman, Z. Is your smartphone listening to you? 2016.[Online] Available: https://www.bbc.co.uk/news/technology-35639549

[25] Nichols, S., Your Phone Is Listening and it's Not Paranoia. 2018, Vice.[Online] Available: https://www.vice.com/en_au/article/wjbzzy/your-phone-is-listening-and-its-not-paranoia.