Ethics in Prognostics and Health Management

Kai Goebel[1, 3], Brian Smith[2], Anupa Bajwa[2]

[1]*Palo Alto Research Center, Palo Alto, CA 94304, USA*
kai.goebel@parc.com

[2]*NASA Ames Research Center, Moffett Field, CA 94035, USA*
{brian.e.smith, anupa.bajwa}@nasa.gov

[2]*Luleå Technical University, SE-971 87, Luleå, Sweden*
kai.goebel@ltu.se

ABSTRACT

As we enter an era where intelligent systems are omnipresent and where they also permeate Prognostics and Health Management (PHM), the discussion of moral machines or ethics in engineering will inevitably engulf PHM as well. This article explores the topic of ethics within the PHM domain: how it is relevant, and how it may be addressed in a conscientious way. The paper provides a historical perspective on ethics-related developments that resulted in the formulation of engineering ethics codes, regulations, and policies. By virtue of these developments, ethics has already been encapsulated in PHM systems. The specific areas that have traditionally driven ethics considerations include safety and security, and they increasingly include privacy, proprietary considerations (protection of intellectual property), and environmental protection. During the course of future technology development, innovations will increasingly impact all of these topics. It is argued here that consciously embracing these issues will increase the competitive advantage of a PHM technology solution. As a guideline, specific ethics attributes are derived from professional engineering ethics codes, and a path towards insertion into a requirements flowdown is suggested.

1. INTRODUCTION

At first sight, talking about ethics in PHM seems somewhat peculiar because PHM practitioners operate with the implicit assumption that PHM in and of itself provides benefits to its users. For example, PHM seeks to prevent systems from failing and therefore either prevent or diminish economic loss or harm to equipment or even people. As the role of PHM expands from prediction of failure times for hardware components to both predicting events of interest in large connected mixed environments and also to making autonomous decisions, it becomes increasingly evident that questions surrounding ethics need to be answered. Ethics, in a very general context, is concerned with what is good and of value (Murphy, Gardoni, Bashir, Harris, Masad, 2015) and engineering ethics considers these issues in the context of engineering.

Both ethics and morals relate to "right" and "wrong" conduct and they are sometimes used interchangeably. However, they refer to different things: ethics refers to rules provided by an external source – the "shoulds" such as codes of conduct in workplaces or principles in religions (Mizzoni, 2017). Morals refers to an individual's own principles regarding right and wrong (Holmes, 2007). It is important to clarify the distinction at this juncture because the discussion here is not meant to challenge what an individual ought to consider wright or wrong. Rather, it is meant to illuminate codes of conduct in PHM as this field evolves into less charted territory that may face more ethics questions.

PHM, a specialty area of engineering, started out strictly as a component-centric examination of equipment faults and failures. Over the years, those examinations advanced to systems that were more complex and with that, the definition of prognostics has been broadened as well. Recently, PHM has expanded into the realm of equipment, and we now see applications to airspace, human health, portfolio management, power plant performance optimization, and insurance underwriting. Further evolution will take PHM to autonomous decision making for self-driving vehicles, cyber-threat prediction and prevention, and perhaps – with some imagination – intent prediction for various application areas (see Figure 1 for a depiction of that evolution). It is easy to see how PHM has already had (and will have even more so

in the future) an impact on the environment, privacy, safety, and certainly on liability.

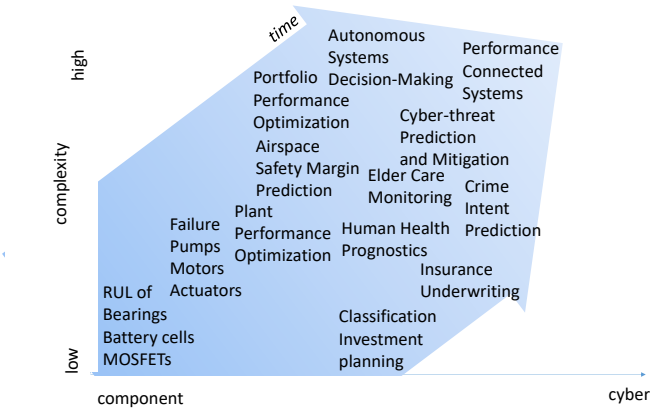

Figure 1. Evolution of PHM

Besides the need to deal with technical issues, such as recognizing patterns upon which a decision is based (a typical PHM function), we may soon need to consider issues such as fairness and bias that a decision-making algorithm may exercise. In addition, ethical considerations come into play in the selection of the datasets to be used as input to train the algorithms. The patterns identified are dependent on the integrity, biases, and completeness of the input datasets. Hidden biases may not be easily measured with common performance metrics. It may also be necessary for PHM analysis tools utilizing black box artificial intelligence approaches to be designed such that the conclusions they draw can be transparently explained to the end user. This is needed not only to satisfy engineers' curiosity who want to understand the model, but also to provide assurances that ethics traits (such as lack of biases), can be traced through the algorithms. As a community, we will have to ascertain that there are guidelines that not only adhere to clear ethics guidelines, but proactively dissolve potential liability issues.

This article seeks to illuminate some of the issues surrounding ethics in PHM. The article suggests how ethics considerations may be incorporated into PHM requirements that are both acceptable and binding. As the engineering discipline of PHM grows in stature and use within our society, engineers will come under increasing scrutiny by the public, the media, the government, and the profession itself on the moral and ethical dilemmas posed by PHM capabilities. Having a thoughtfully developed sense of ethics among practitioners, along with tools that aid in integrating ethics principles into PHM algorithms, will be vital in that process.

2. PHILOSOPHICAL RUMINATIONS

Ethics definitions include notions of right and wrong that have evolved within our cultural context over the last few thousand years. Three major schools of thought are often cited (Tännsjö, 2013): Deontological ethics (law-based ethics: what is my duty?), utilitarianism (what is the greatest possible good for the greatest number?), and virtue ethics (what is the best form/version of this particular thing, in these particular circumstances?). Some ethicists discuss the notion that there is something like a moral perception that tells humans right from wrong which is done not merely using measurable observations but it is processed through our emotional fabric, beliefs, and desires (Frank, 2016) which drive a *motivation* to act morally. Somewhat complicating the matter is that the resulting output may not be absolute: there are many gradations on how moral an action is. We are faced, daily, with making decisions in the midst of moral dilemmas where we typically must trade-off one ethical consideration for another. Nonetheless (and to overcome the question: whose ethics do we need to follow?), there are sets of minimum bounds for ethical behavior that have been encoded in laws and policy for the benefit of society that agreed to these guidelines. An established framework – with precursors in the English Bill of Rights of 1689 (Bill of Rights, Great Britain, 1689), the French Declaration of Rights of Man and Citizen of 1789 (French National Constituent Assembly, 1789), and the US constitution of 1791 (US Congress, 1791) and ultimately the Universal Declaration of Human Rights by the United Nations in 1948 (United Nations, 1948) – is that one may exercise one's liberty – which imposes societal duties on others not to interfere with those liberties.

For algorithms, this poses a number of challenges: although it is relatively easy to encode hard rules, it is a far more challenging task to encapsulate the emotional moral response that we instinctively have in a cultural and societal context. It falls on those who define requirements and, ultimately on the software engineer how the duties and limits of ethics are interpreted. When designing these algorithms it is the responsibility of the engineer to fulfil the "duty of inquiry" (Clifford, 1877), i.e., to actively probe requirements as opposed to assuming that someone else would surely have sorted every ethical underpinning out to the last detail.

3. ENGINEERING ETHICS AND ACCREDITATION

Engineering ethics evolved from a perspective of a personal concern to that of a professional concern as engineering was rising in prominence towards the end of the 19th century (Flavell, 2016; ASME, 2015). Owing to a number of fatalities due to technical failures in bridges, vessels, and other structures, the professions sought to impose both licensing requirements as well as a code of ethics. Engineering ethics codes contain all elements found in corporate and civil

service ethics policy, including integrity, conflict of interest, honesty, confidentiality, etc. Additionally engineering ethics prescribe that an engineer's duties are to protect public health, safety, and welfare. A rather simplistic view of the ethics onion is shown in Figure 2. PHM and Engineering ethics are part of a set of ethics behavior that includes professional ethics and social ethics. Engineering ethics may have a more limited scope to the degree that they should be considered when building a PHM solution. This is in contrast to personal and social ethics where other additional attributes are at play. Values are a superset of ethics that include general notions of right and wrong. And, in this simplified view, morals would be an all-encompassing class that also is concerned with customs and traditions.

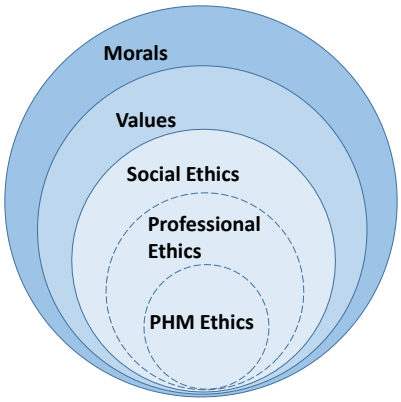

Figure 2: Ethics, Values, and Morals

3.1. Ethics Codes

The National Society of Professional Engineers acknowledges in its preamble to the Code of Ethics that "Engineering has a direct and vital impact on the quality of life for all people. Accordingly, the services provided by engineers require honesty, impartiality, fairness, and equity, and must be dedicated to the protection of the public health, safety, and welfare" (NSPE, 2007). The various engineering societies (ASME, ASCE, IEEE, and others) have adopted similar statements and delineate various instances of proper behavior, including fairness and accountability. Even though some of the codes were developed decades ago, several make reference to more recent insights such as the need to "adhere to the principles of sustainable development in order to protect the environment for future generations" (NSPE, 2008). Missing generally are references to derived ethics concerns such as cybersecurity, privacy protection, intellectual property rights for information technologies or similar "modern" issues. However, these can usually be derived from the generic statement on public health, safety, and welfare.

3.2. Engineering Accreditation

Becoming an engineer is a process that varies widely around the world. In some regions, use of the term "engineer" is actually regulated and there are specific procedures and requirements for obtaining a registration, charter or license to practice the procession. Licenses are obtained from the government or a charter-granting authority, and engineers are subject to regulation by these bodies (Layton, 1986). In addition to licensing, there are voluntary certification programs for various disciplines that involve examinations accredited by the Council of Engineering and Scientific Specialty Boards (Anderson, 1999). At any rate, engineers pledge to adhere to their respective engineering ethics code.

3.3. Engineering Ethics Attributes

Drawing on the common understanding in the engineering ethics codes that public health, safety, and welfare are immutable qualities, the following sections examine the topics safety, (cyber-) security, privacy, and sustainability, as components of these engineering ethics qualities.

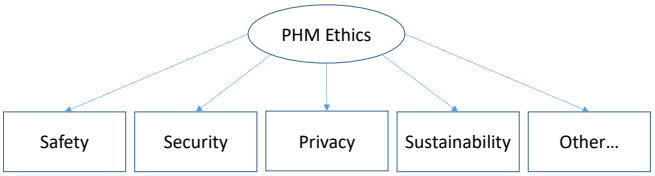

Figure 2: Engineering Ethics Attributes

3.3.1. Safety

Safety is defined as "relative freedom from danger, risk, or threat of harm, injury, or loss to personnel and/or property, whether caused deliberately or by accident" (businessdictionary.com, 2018). A relevant question is: How safe is safe enough? Within the context of PHM, safety is either directly or indirectly impacted by PHM tools. An example of direct impact is through the calculation of a safety margin, or by setting the false positive and false positive rate for safety critical functions (such as rocket abort logic). In the second example, the ethical implication of weighing acceptable loss of life versus economic loss becomes clearly evident. Safety can also be impacted indirectly, for example when alerts are incorrectly interpreted by operators (such as pilots). For example, consider an aircraft cockpit alert that an engine has caught fire. One of the steps in addressing the issue might be to turn the fuel supply and the engine off. If the pilot were to turn off the incorrect engine (which does happen occasionally), the state of safety has just become much worse. While not all events that lead to degraded safety can be prevented by PHM methods, there is a significant proportion of cases where PHM can play an important role in informing about, preventing, or mitigating unsafe conditions and hazards.

Using transportation as an example, safety statistics are compelling evidence that safety is not a solved issue. For example, in aviation, there were between 1989 and 2008 2151 fatalities in 600 accidents. Of those, 109 accidents (with 777 fatalities) were due to equipment malfunction (NTSB, 2011). The UK CAA estimates that around 80% of aircraft mechanical defects (on helicopters) are detectable, so an assumed accident and death/injury prevention rate could be derived using advanced PHM principles

For terrestrial transportation, and more specifically, for cars, the numbers are considerably worse (although typically they do not grab as many headlines as aircraft accidents). The National Highway Traffic Safety Administration (NHTSA) reports that there were more than 5.5 million police-reported motor vehicle crashes in the United States in 2009. A total of 1.52 million of those crashes resulted in an injury, and 30,797 resulted in a death (i.e., several orders of magnitude worse than those in aviation!). The NHTSA report "National Motor Vehicle Crash Causation Survey" (DOT, 2008) cites equipment malfunction as a contributing factor in 6.8% of all cases as being (i.e., for more than 2000 of the deaths). The most cited types of equipment failure are loss of brakes, tire blowouts or tread separation, and steering/suspension failure (DOT, 2008).

The point of citing equipment malfunction in these accidents is that PHM may play an important role in improving and maintaining safety in these transportation cases.

PHM capabilities are used to track data on components and systems including those that have *not* failed. Units whose precise times of failure are unknown are referred to as censored units. Inexperienced analysts frequently do not know how to analyze so-called censored data, and they omit the censored units as a result. This can bias an analysis toward an over-estimation of the rates of failure.

As mentioned previously, moral dilemmas can arise where an organization may trade off one ethical consideration for another. An example of this in the PHM domain is the risk trade between delaying PHM-generated maintenance recommendations of a safety device versus going ahead with them even though such indicated maintenance or corrective action may take the larger system out of service. A more specific example might be deferring maintenance on a brake sub-system to a time of scheduled downtime vs. performing maintenance immediately and using a less capable loaner vehicle as a temporary replacement. The dilemma described here is that there is increased risk exposure in both scenarios with different financial implications. The interplay among the three key system attributes, reliability, availability, and maintainability (RAM), needs to be weighed against business objectives and ethics considerations. Likewise, Failure Reporting and Corrective Action Systems (FRACAS) based on PHM can only operate effectively when integrated with a management philosophy based on the safety of the end users.

3.3.2. Privacy

The Internet of Things presents tremendous opportunity for PHM in areas where monitoring, trending, or prediction is needed. In smart homes one might be interested in tracking power consumption, control heating and cooling, and perform various automated monitoring functions to ensure smooth operation including operation of the internet. One other example is elder care, where the home of a person is equipped with data collection equipment that tracks a person's habits and compares those against baseline behavior with the goal to ascertain that the individual is not in physical harm. This is an example of an extension of PHM to include PHM for human health. The impact on a person's privacy is obvious, depending on how intrusive the sensors are. In particular, the information sought and the sensors considered for collecting this information may include (Townsend, Knoefel & Grouban, 2011) activities of daily living (e.g., door sensors, pressure sensors in chairs or bed, ultrasonic sensors in bathroom and kitchen), location and position of person (e.g., wearable accelerometers, ultrasonic movement detectors, or computer vision for fall detection), sporadic physiological information (e.g., blood pressure monitoring cuff), continuous physiological information (e.g., wearable heart rate monitor), up to complete visual information (e.g., video cameras with image recognition software). Clearly, any such deployment can be controversial and it only works if the independence associated with aging-in-place is valued higher than privacy. This article does not seek to engage in the merits of deployment of these technologies, it merely points out the impact on privacy. Other applications may be less dramatic, such as tracking of energy usage by smart home equipment such as those offered by utility companies, phone companies, or a number of large tracking service providers. In each of those cases, the information – while being made available to the customer – is also being collected by the service provider with certain implications on privacy. Similarly, an airline that engages a company to provide remote monitoring services of its aircraft will have an expectation that airline data are being treated with utmost care. The above cases illustrate potential downsides and unintended consequences of PHM unless ethically reviewed.

3.3.3. Cybersecurity

Cybersecurity and PHM overlap in a number of ways. For one, there is an obvious need to protect the integrity of critical system information or proprietary information. Kwon et al. (Kwon, Hodkiewicz, Fan, Shibutani, PechtK, 2016) point out the need to security in an IoT-enabled PHM environment. In 2016, malware infected hundreds of thousands of connected IoT devices and exploited them to conduct the largest distributed denial-of-service attack seen so far, reaching an offensive capability of about 1.2 terabits per second (De Donno, Dragoni, Giaretta, Spognardi, 2018). Beyond the denial-of-service inconvenience – and perhaps more severe –

information retrieved from systems may, if tempered with, result in changed set-points that in turn result in taking systems off-line, driving systems to unsafe regions, or, more subtly, in preventing needed maintenance. Undeniably, such scenarios may result in situations where the safety of operations is impacted, when a sensitive margin is not acted upon. Even worse, if incorrect decisions are communicated back to the system, assuming that the system has the capability to enact them, at least in theory any arbitrary decision could be imparted. One has to think only of the stuxnet virus that targeted SCADA systems (Siemens control software Step 7) and caused centrifuges at the Tangaz plant to over-speed and self-destruct (Koch & Kuehn, 2017). In general, there are two major vulnerabilities of SCADA systems: unauthorized access to software (virus infections, intentionally induced changes, or other problems that can affect the control host machine); and vulnerability to packet access to network segments that host SCADA devices where little or no security of actual packet control protocol exist. Theoretically, anyone sending packets to a SCADA device could be in a position to control it.

Another example is the jamming and reprogramming of GPS signals that were responsible for diverting a military grade drone over Afghanistan in 2012. The drone was then landed intact in Iran by guiding it. This was done by exploiting a navigational weakness, essentially cutting off communications links of the American bat-wing RQ-170 Sentinel, then reconfiguring the drone's GPS coordinates (Peterson & Faramarzi, 2011). By putting noise (jamming) on the communications, the drone was forced to switch into autopilot and subsequent "spoofing" (i.e., generating and imposing false readings that look real to the system) of fake GPS signals made the drone "land on its own where we wanted it to, without having to crack the remote-control signals and communications" from the US control center (Peterson & Faramarzi, 2011). Clearly, such capability has significant ramifications for in-air on-demand mobility, but also for autonomous terrestrial vehicles and other mobility applications. The relevance to PHM here is that some safety predictions may rely on GPS information

In 2017, hackers used malware dubbed "Triton" to take control of a safety work station at an industrial power plant, then worked their way through the system to reprogram controllers used to identify safety issues. Operators noticed the attack when some controllers inadvertently entered a failsafe mode and caused related processes to shut down (Gibbs, 2017). This attack breached the safety system (which is at the heart of some PHM activities) and as such indicates the potential for other parts of any industrial plant being compromised - while operators may not even initially detect the attack.

PHM systems may also require protections from ransomware. Ransomware is software that infects a computer-controlled system by shutting down vital functions. Typically, the functions cannot be restored until the ransom is paid to the anonymous perpetrator. In 2018, the Boeing plant in North Charleston, S.C. was hit by a ransomware attack with the WannaCry virus that resulted in equipment lockdown and demanding of ransom payment in exchange for release of the computer system. Boeing feared at the time that similar viruses might be directed against equipment used in functional airplane tests, which could lead to it spreading to aircraft flight-critical software (Gates, 2018). While the virus exploited Windows operating system vulnerabilities, the fact that it can impact equipment used in PHM-centric operation or lock down PHM software makes it relevant for the PHM domain.

In 2015 a Ukrainian power plant was hacked, likely using a software "backdoor" to infiltrate the power plant's controls system and to remotely turn off switches that resulted in loss of power to 80,000 customers (Zetter, 2016). Power was restored by manually operating the switches at the substations but it should be noted that fully automated systems would have a harder time recovering from such an attack if a manual restoration function does not exist.

A different type of overlap between PHM and cybersecurity is to use PHM **for** Cybersecurity. The latter has been explored by Evans et al. (Evans, Mishra, Yan, & Bouqata, 2016) where the authors describe how security related protections would be served to integrate fully with Monitoring and Diagnostics systems that assess the health of complex assets and systems and in particular combining system parameters already in use for Prognostics and Health Monitoring (PHM) with security parameters to detect complex cyber threats. Indeed, the idea to use PHM principles for intrusion detection is not new. As summarized in Samrin & Vasumati (Samrin & Vasumati, 2017), the gamut of anomaly detection and classification tools used commonly in PHM (e.g., Naïve Bayes classifier, ANN, Fuzzy clustering, k-means, knn, SVM, random forest, and decision trees, often in combination with some other technique) can be found for network intrusion detection.

3.3.4. Sustainable Development

Any development that a PHM engineer undertakes should ultimately be supportive of sustainability goals, or at least not counteract them. Within the energy domain, sustainability is a goal that is better aligned with the core business than for some other domains because of the direct impact of energy conversion on the environment. In fact, PHM plays a big role in achieving sustainable operations. The functions of monitoring, alerting, and prediction are key enablers in ensuring that environmental goals can be met. Environmental Monitoring and Assessment deals with technical developments and data arising from environmental monitoring and assessment, principles in the design of monitoring systems, and the use of monitoring data in

assessing the consequences of natural resource management and pollution risks. Less clear is where an engineer works on some application that is not obviously associated with sustainability. What is asked through the Engineering Ethics code is that those engineers still keep in mind that their work may ultimately have an impact on the environment – and to design systems appropriately. The application of the on-board diagnostics in vehicles is a good example. The California Air Resources Board (CARB) required that all new vehicles sold in California starting in 1996 have some basic on-board diagnosis (OBD) capability (CARB, 2018). While this capability enhances safety, eases maintainers' troubleshooting, and reduces operational cost, the regulation was motivated by a desire to reduce the exhaust emissions and to institute a state-wide tailpipe emissions testing program.

4. ETHICS IN AI

Artificial Intelligence (AI) is an area that is tightly interwoven with PHM since many realizations of PHM techniques borrow more or less heavily from AI. For example, diagnostics is fundamentally a classification problem which has for several decades been explored using AI pattern matching or other inference tools. Similarly, decision-making (which is related to the "Management" in PHM) is an area where many AI-driven optimization tools have been explored. It is therefore appropriate to glance into the very active discussion on ethics in AI that is taking place in a variety of sectors (Wallach & Allen, 2010; Wallach, 2015) (see also workshops on the topic such as the AAAI Symposia in 2005, 2016, 2018). The debate ranges from trying to understand what ethics means in the context of machines – and how to impart ethics attributes into machines – to the (possibly harder to control) impact of free-roaming artificial ethical entities.

The foundational aspects of that discussion reaches back again into moral philosophy which postulates (in deontological ethics) the ethical position that morality of an action is guided by rules. If it can be interpreted as "obligation or duty", and consequent moral judgment on the actor on whether he or she has complied, this makes implementation of ethics somewhat easier than the fixation often found on pathological (but extremely unlikely) decision cases where a choice is given for an algorithm to choose between killing one or more persons with widely differing ages. Nonetheless, moral actions are often times associated not only with doing the right thing, but having the freedom to choose to do the right thing (Johnson, 2006). This is a critical difference and will ultimately decide whether we will treat AI systems as truly autonomous and allow them to make their ethical choice (Bryson, 2016) (which may then have evolved from our understanding now) or whether we keep a tighter leash on AI systems to conform with our value system – in which case, and perhaps somewhat paradoxically, we may not be able to expect truly ethical behavior from an AI system. In general, one may have to be skeptical about being able to realize satisfactory fully functional ethical behavior in an AI system any time soon (Frank, 2016). This should not be an excuse to skirt the issue of integrating ethics into AI – and PHM algorithms. But it may be a better goal for the time being to integrate ethics attributes into algorithms that possess measurable engineering metrics. A well-designed set of attributes can help engineers think consciously about complex ethical issues that may not have intuitively obvious solutions. It can also prevent suboptimal solutions arising from human limitations such as illusory correlation, overconfidence, limited experience, or cognitive overload.

5. ETHICS AS A DESIGN REQUIREMENT

The above discussion centered mostly on delineating different ways how ethics is part of various engineering functions. The creation of safe systems involves the application of methods from a variety of disciplines, coordinating and controlling the system creation process, and performing these functions under the influence of a number of external factors. Using systems engineering principles can assure the creation of truly safe systems. Creating dependable systems requires that systems engineers develop an ethical awareness of the holistic, interdependent nature of these processes and their effects on the safety of the systems being created. The notion of PHM in complex systems, therefore, transcends "down-and-in" engineering, management, and social processes and can only be achieved as an emergent, ethical property of a system that accounts for all of these domains. This "health" property is enabled by systems engineering practices that examine the dynamic processes within the context of the lifecycle phase in which the projected system will operate, the scale and complexity of the system being created, and the social interactions that take place among the individuals and organizations involved in the overall task of creating the system. During the systems engineering development stage, requirements developed during the conceptual stage are translated into conceptual product architectures, and alternative designs for specific and tangible elements that will execute the system functions are postulated including the overarching ethical implications of the architectures. (Nui, 2017)

A critical question for PHM ethics is how one can actually integrate ethics into the PHM process. One straightforward solution might be to impose ethics as a top-level requirement. The question then becomes how that requirement flows down into more tangible low-level requirements that can be verified and validated. To that end, it is useful to recall that engineering ethics codes articulate different attributes of ethics, in particular safety, privacy, cybersecurity, and sustainability (see section 3.3). No claims are made here that this is an exhaustive (or mandatory) list of ethics attributes, only that it is a good place to start with.

It is important that requirements can be measured with suitable metrics. The next sections explore how the PHM ethics attributes mentioned here can be concretized in measurable metrics that in turn can be incorporated – and sometimes already are part of – design requirements.

5.1. Safety Requirements

Flowdown of a safety requirement might lead to metrics such as tripping some safety threshold. This could be a simple univariate threshold such as an upper pressure setting, temperature limits, current limits, or really a limit to any physical quantity that is deemed critical for safe operation of the asset. These thresholds can also be realizations of multivariate combinations of different physical quantities that together lead to unsafe conditions. Alternatively, the threshold could be realized as a complex set of features that are derived using advanced analytics from measured observations. An example is a set of features calculated from a vibration power spectrum. Another example might be a set of features that has been derived from sensor data via machine learning and has no obvious physical meaning but has instead been found to correlate with unsafe conditions. In addition, time to critical event conditions can also be expressed as a threshold, such as remaining time to battery exhaustion in drones (Saha, Quach & Goebel, 2012).

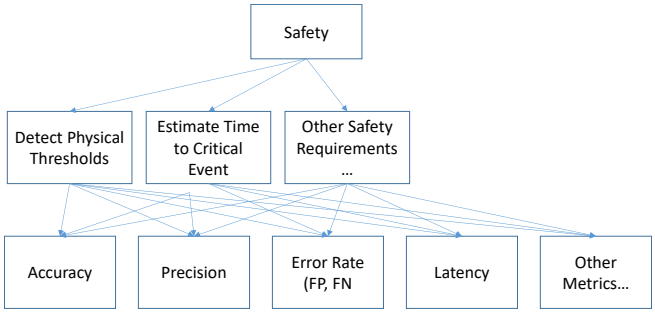

Figure 3: Safety Requirement Flowdown

Another example is time to unsafe event in the airspace as expressed as a combination of anticipated flight route, traffic, proximity to convective weather, controller fatigue, aircraft health and/or energy state, etc. (Roychoudhury, Daigle, Goebel, Spirkovska, Sankararaman, Ossenfort, Kulkarni, McDermott & Poll, 2016). The field for defining safety is probably the most mature of all the attributes considered and the easiest to align with traditional PHM activities. From the physical threshold, or estimate time to safety critical event, metrics such as accuracy, precision, error rate, latency etc. can be used. PHM practicioners are quite familiar with these metrics. Figure 3 shows some of the safety requirements and associated metrics.

5.2. Privacy Requirements

Privacy is an attribute that is often not immediately considered in PHM development. However, collection, processing, and storage of proprietary or sensitive operational data is frequently an integral part of a PHM solution. This requires designers to consider confidentiality and data protection, at the minimum (de KKlerk, 2017). It may also be required to define disclosure and consent as well as level of user control. Those requirements then can be flowed down into the need to provide access safeguards as well as protection against data leakage. In addition, it may be necessary to provide a sufficient level of anonymization, for example when PHM analysis data are shared between different competitors such as pooling best practices of successful repair amongst different airlines (Maggiore & Kinney, 2009) in a secure offering where the identity of the airlines and other proprietary information are hidden but beneficial statistical information on successful repair is shared. In that context, pseudonymity can also be used as an aid in providing a desired level of privacy. PHM designers may also have to think about PHM analysis data retention limitation as defined by user specification. Finally, "unlinkability" metrics describe the inability of an observer to decide whether certain items of interest are related or not. This attribute is meant to ensure that "a user may make multiple uses of resources or services without others being able to link these uses together" (ISO, 1999). It should be noted that many of the metrics listed here are not easily quantifiable on a continuous scale. Instead, they are often evaluated with binary fulfilment levels. Figure 4 shows privacy requirements and associated metrics described above.

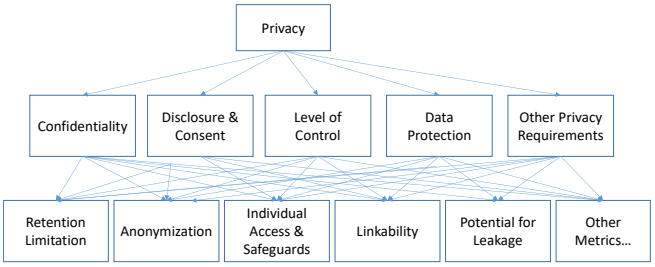

Figure 4: Privacy Requirements

5.3. Cybersecurity Requirements

Cybersecurity is of importance not only to PHM and has received a lot of attention in many other application areas as well. Indeed, hundreds of measures (Black, Scarfone, & Souppaya, 2008) have been developed to address cybersecurity concerns. Cybersecurity is of increasing importance in PHM as well as discussed in section 3.3.3. While complete coverage exceeds the scope of this paper, some representative requirements and metrics are listed here. These include low vulnerability, change protection, high level of control, malicious code detection capabilities, and

incident prevention and handling capabilities. From those requirements, metrics can be derived that include Vulnerability Scanning Coverage, Percent of Systems with No Known Severe Vulnerabilities, Number of Known Vulnerabilities, Percent of Changes with Security Exceptions, Security Testing Coverage, (Payne, 2006), Security audit logs of individual systems, Number of systems within an organization that were tested over the course of a year, Number of port scans detected, System patch status, Presence and strength of intrusion detection system, etc. A sample mapping is shown in Figure 5. While this section gives only a brief glimpse into the complexity of this field (which in itself has spawned a whole industry with associated research, conferences, etc.), it is meant to point out the needs to consider these requirements as they touch on critical capability that is vital for judicious deployment of PHM services.

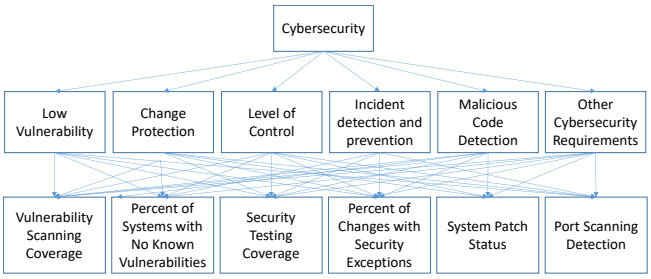

Figure 5: Cybersecurity Requirement Flowdown

As mentioned previously (Evans 2016), Security Prognostics (SP) for critical systems should fully integrate with Monitoring and Diagnostics (M&D) systems. To detect complex Cyber threats it may be possible to combine system parameters already in use by M&D systems for PHM with customary security parameters. Combining system parameters used by M&D to detect non-malicious faults with the system parameters used by security schemes to detect complex Cyber threats will improve: (a) accuracy of critical PHM (b) security of M&D, and (c) availability and safety of critical systems. Evans (2016) introduces the notion of Remaining Secure Life (RSL), assessed based on the propagation of "security damage," to create the prospect for Security Prognostics. RSL will assist in the selection of appropriate response(s), based on breach or compromise to security component's and potential impact on system operation. Defining RSL may become increasingly important for critical cyber-physical systems that support high-energy processes or healthcare IT or patient monitoring systems dependent on reliable PHM.

Interestingly, with the increasing digitization of PHM data in the healthcare industry, a wide range of devices (including traditionally non-networked medical devices) are now Internet- and inter-connected. Mobile devices (e.g. smartphones) are one common device used in the healthcare industry to improve the quality of service and experience for both patients and healthcare workers, and the underlying network architecture to support exchange of PHM data among other things is also referred to as medical smartphone networks (MSNs). MSNs, similar to other networks, are subject to a wide range of attacks (e.g. leakage of sensitive patient information that is derived from hospital patient-monitoring PHM systems by a malicious insider using a smartphone). This is a risk that needs continuous monitoring. (Meng, Li, Xiang & Choo, 2017)

5.4. Sustainability Requirements

Section 3.3.4 argued for the relevance of sustainability for PHM. Generally, sustainability requirements can be organized into different indicators, namely environmental, economic, and social (see Figure 7). The metrics for these indicators include, like any other metric, target thresholds. For example for sustainability requirement of atmospheric impact one might be able to define a specific emissions threshold. For sustainability requirement resource usage, one might be able to define thresholds for consumption of energy, water, or fuel. These metrics should be used to both assess how PHM technology stresses sustainability requirements (perhaps caused by using sensors with rare materials or using methods that have high energy usage) but also how PHM helps to meet sustainability requirements in alerting and preventing negative impact on atmospheric and resource threshold violations. ASTM (formerly known as American Society for Testing and Materials, an international standards organization) describes in documents coming from subcommittee E60.13 on Sustainable Manufacturing how some PHM principles (without using the term PHM, though) are being borrowed to achieve sustainability goals.

A key element of sustainability is the success of PHM technology in creating wealth. The economic indicators go somewhat further than conventional financial reporting in describing the creation of wealth or value, and in reporting its distribution and reinvestment for future growth. Where that is not explicitly expressed elsewhere, an example metrics might include return on investment (ROI) or similar.

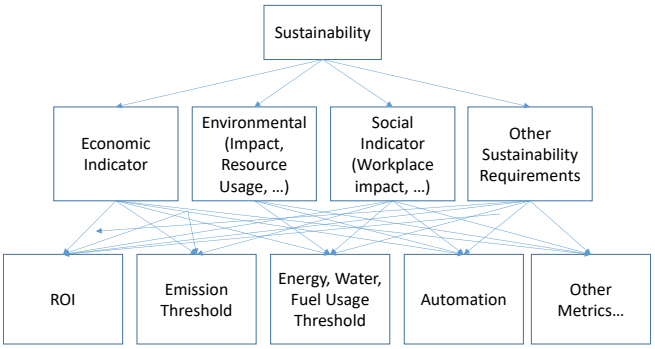

Figure 6: Sustainability Requirements

Social indicators are a bit more difficult to assess. They are the soft indicators that are not typically included as an engineering requirement. Social indicators reflect the company's attitude to treatment of its own employees, suppliers, contractors and customers, and also its impact on society at large. Impact on workplace could be measured by degree of automation or by degree of autonomous decision systems. Impact on customers might include reduction of delays, cancellations. As with all earlier requirements, specific applicability has to be judged on a case by case basis.

6. CONCLUSIONS

This article explored the topic of ethics within the PHM domain: how it is relevant, and how it may be dealt with in a conscientious way. The paper provides a historical perspective on ethics-related developments that resulted in the formulation of engineering ethics codes, regulations, and policies. By virtue of these developments, ethics has already been encapsulated in PHM systems. The specific areas that have traditionally driven ethics considerations include safety and security, and they increasingly include privacy, and environmental protection. During the course of future technology development, innovations will increasingly impact all of these topics and how they may interact – in both knowable and unanticipated ways - from a systems engineering perspective. It is argued that consciously embracing these issues will increase the competitive advantage of a PHM technology solution. Industry and government will inevitably face choices between the rising demands on operators to achieve profitability and minimal out-of-service times that can be achieved using PHM tools *vis a vis* the ethical constraints of safety. Having awareness of these dilemmas in a consolidated way, as well as a mechanism to address these key domains, will help to ensure that PHM systems will not only enjoy maximum impact and least pushback but also provide a competitive advantage.

As a guideline, specific ethics attributes were derived from professional engineering ethics codes, and a path towards insertion into a requirements flowdown was suggested. The way to ensure that capabilities become part of an end-user product in the intended way is through the verification and validation (V&V) of requirements. Similarly, "ethical engineering design" elements need to be incorporated at the earliest stages of research into concepts of operations and enabling technologies. V&V is not a one-shot activity and instead is performed as the capability is being developed and matured. To that end metrics are measured to satisfy the requirements. Ethics requirements should not be treated any different and they should organically satisfy high level goals.

> *PHM "ethics is an issue that goes to the heart of engineering practice. It reflects the customs, habits, and values of engineering as a profession and reflects the time-tested experience, seasoning and training of practicing engineers. In some senses, a code is a "timeline" for the profession because it mirrors the conventions, routines and patterns of the profession but shifts as those conventions, routines and patterns change.*
>
> *As the profession of engineering [in Prognostic Health Management] grows in stature within our society, the engineering and engineers will be increasingly examined and scrutinized by the public, the media, the government and the profession itself on moral and ethical questions. Having a thoughtfully-developed code of ethics along with [PHM practitioners] that adhere to that code will be vitally useful in that process."*[1]

ACKNOWLEDGEMENT

This work [2] was in part supported by the NASA ARMD/AOSP/SWS project.

REFERENCES

Anderson, W. (1999). A Primer on Credentials for Engineering and Related Fields, *The Council of Engineering and Scientific Specialty Boards*, April 2, 1999.

ASME. (2015). Unwritten Laws of Ethics and Change in Engineering, *ASME Press*.

ASTM International (2018). Subcommittee E60.13 on Sustainable Manufacturing, E3012-16 Standard Guide for Characterizing Environmental Aspects of Manufacturing Processes.

Black, P., Scarfone, K., & Souppaya, M. (2008). *Cyber Security Metrics and Measures*, Wiley Handbook of Science and Technology for Homeland Security, (Ed.) John G. Voeller, Wiley, 2008.

Bryson, J., (2016). Patiency is not a Virtue: AI and the Design of Ethical Systems, *AAAI Workshop Notes of 2016 Spring Symposium*, pp. 202-207.

businessdictionary.com, (2018). *http://www.businessdictionary.com/definition/safety.html*, accessed 3/13/2018

California Air Resources Board (2018). On-Board Diagnostics (OBD) Program. *Arb.ca.gov*.

[1] Adopted from A. Schwartz, National Society of Professional Engineers, https://www.nspe.org/resources/ethics/ethics-resources/other-resources/engineering-ethics-search-solutions

[2] Kai Goebel performed most of this work while employed at NASA Ames Research Center

https://www.arb.ca.gov/msprog/obdprog/obdprog.htm, Retrieved 10/17/2018.

Clifford, W., (1877) The Ethics of Belief, Originally published in Contemporary Review.

De Donno, M., Dragoni, N., Giaretta, A., & Spognardi, A. (2018). *Security and Communication Networks*, Vol. 2018, Article ID 7178164.

Evans, S., Mishra, P., Yan, W., & Bouqata, B. (2016). Security Prognostics: Cyber meets PHM", *Proceedings IEEE PHM Conference*.

Flavell, E. (2013). *The ASCE Code of Ethics: PRINCIPLES, STUDY, AND APPLICATION*. ASCE. Archived from the original on 2013-12-03

Frank, L. (2016). Metaethics in Context of Engineering Ethical and Moral Systems, *Workshop Notes, AIAA Spring Symposium*, pp. 208 – 213.

French National Constituent Assembly. (1789). *Déclaration des droits de l'homme et du citoyen de 1789*.

Gates, D., (2018). Boeing hit by WannaCry virus, but says attack caused little damage, *Seattle Times*, *https://www.seattletimes.com/business/boeing-aerospace/boeing-hit-by-wannacry-virus-fears-it-could-cripple-some-jet-production/*, last accessed 4/10/2018.

Gibbs, S. (2017) Triton: hackers take out safety systems in 'watershed' attack on energy plant, *The Guardian*, *https://www.theguardian.com/technology/2017/dec/15/triton-hackers-malware-attack-safety-systems-energy-plant*, last accessed 4/10/2018

Great Britain Parliament (1689), *Bill of Rights.*

Holmes, R., (2007). *Basic Moral Philosophy*, Wadsworth.

ISO (1999). Common Criteria for Information Technology Security Evaluation, *ISO/IEC 15408*.

Johnson, D. (2006) Computer Systems: Moral Entities but not Moral Agents, Ethics and Information Technology, vol. 8, pp 195-204.

de Klerk, T. (2017). How to Measure Privacy: a Structured Overview of Analysis Methods for Privacy-Enhancing Technologies, *Notes from Twente Student Conference on IT*.

Koch, R., & Kuehn, T. (2017). Defending the Grid: Backfitting Non-Expandable Control Systems, *Proceedings 9th International Conference on Cyber Conflict*

Kwon, D., Hodkiewicz, M., Fan, J., Shibutani, T., & Pecht, M. (2016). IoT-Based Prognostics and Systems Health Management for Industrial Applications, *Special Section on Trends and Advances for Ambient Intelligence with Internet of Things (IoT) Systems*, IEEE Access.

Layton, E. (1986). *The Revolt of the Engineers: Social Responsibility and the American Engineering Profession*. Baltimore, Maryland, USA: The Johns Hopkins University Press.

Maggiore, J. & Kinney, D. (2009). "Monitoring Real-Time Environmental Performance". *Aeromagazine*, Qtr_03 09

Meng, Weizhi, Li, Wenjuan, Xiang, Yang, Choo, Kim-Kwang Raymond, A bayesian inference-based detection mechanism to defend medical smartphone networks against insider attacks, *Journal of Network and Computer Applications, Volume 78, 15 January 2017, pp. 162-169.*

Mizzoni, J. (2017). *Ethics – the basics*, J., Wiley Blackwell.

Murphy, C., Gardoni, P., Bashir, H., Harris, Jr., C.E., & Masad, E. (Eds.). (2015). *Engineering Ethics for a Globalized World*, Springer.

National Motor Vehicle Crash Causation Survey (2008). *Report to Congress*, DOT HS 811 059, July.

NSPE - National Society of Professional Engineers. (2007). Code of Ethics for Engineers, *Publication #1102*, revised July 2007.

NSPE Board of Ethical Review. (2008). Sustainable Development – Threatened Species, *Case No. 07-6*, *http://ncees.org/wp-content/uploads/2014/08/NSPE-Board-of-Ethical-Review-Sustainable-Development-Threatened-Species.pdf*, last accessed 4/10/2018.

NTSB (2011). Aviation Accident and Incident Data System, NTSB, online database *http://www.ntsb.gov/aviationquery/index.aspx*, last accessed 3/10/2011

Nui, G., (2017) Data-Driven Technology for Engineering Systems Health Management: Design Approach, Feature Construction, Fault Diagnosis, Prognosis, Fusion and Decisions, Chapter 2 - Design Approach for Systems Health Management,, ISBN: 978-981-10-2031-5

Payne, S. (2006). A Guide to Security Metrics. *NIST 800-55 Rev 1, Sections 5.0-6.0 NIST 800-100, Section 7.0*.

Peterson, S., Faramarzi, P. (2011). Iran hijacked US drone, says Iranian engineer", *Christian Science Monitor, https://www.csmonitor.com/World/Middle-East/2011/1215/Exclusive-Iran-hijacked-US-drone-says-Iranian-engineer* , last accessed 4/10/2018.

Roychoudhury, I., Daigle, M., Goebel, K., Spirkovska, L., Sankararaman, S., Ossenfort, J., Kulkarni, C., McDermott, W., & Poll, S. (2016). Initial Demonstration of the Real-Time Safety Monitoring Framework for the National Airspace System Using Flight Data, *Proceedings of AIAA AVIATION 2016*, AIAA-2016-4216.

Saha, B., Quach, P., & Goebel, K. (2012). Optimizing Battery Life for Electric UAVs using a Bayesian Framework", *Proceedings of 2012 IEEE Aerospace Conference*.

Samrin, R., & Vasumati, D. (2017). Review on Anomaly based Network Intrusion Detection System, *Proceedings of 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT).*

Tännsjö, T. (2013). *Understanding Ethics*, Edinburgh University Press, 3rd Ed.

Townsend, D., Knoefel, F., & Grouban, R. (2011). Privacy versus autonomy: A tradeoff model for smart home monitoring technologies, *Proceedings Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*.

United Nations. (1948). *Universal Declaration of Human Rights*.

US Congress (1791) *United States Constitution*.

Wallach, W., & Allen, C. (2010). *Moral Machines: Teaching Robots Right from Wrong*, Oxford University Press.

Wallach, W. (2015). *A Dangerous Master: How to Keep Technology from Slipping Beyond Our Control*, Basic Books.

Zetter, K. (2016). Everything we know about Ukraine's Power Plant Hack, *Wired*, *https://www.wired.com/2016/01/everything-we-know-about-ukraines-power-plant-hack/*, last accessed 4/10/2018

BIOGRAPHY

Kai Goebel is a Principal Scientist at Palo Alto Research Center where he explores the intersection of AI with cyberphysical systems. Prior to joining PARC, he worked at NASA Ames Research Center where he was the Area Lead for Discovery and Systems Health. His research interest is in the areas of machine learning, real time monitoring for safety, diagnostics, and prognostics. He has fielded numerous applications for manufacturing systems, aircraft engines, unmanned aerial systems, space systems, transportation systems, energy applications, and medical systems. He holds 18 patents and has published more than 350 papers in the field. He received the degree of Diplom-Ingenieur from Technische Universitaet Muenchen in 1990 and the Ph.D. from the University of California at Berkeley in 1996. Dr. Goebel worked between 1997 and 2006 at General Electric's Corporate Research Center in upstate New York where he was also an adjunct professor at Rensselaer Polytechnic Institute. Dr. Goebel is now an adjunct professor at Lulea Technical University. He is a co-founder of the Prognostics and Health Management Society and he is currently associate editor of the International Journal of PHM. While pursuing the Ph.D. at UC Berkeley, he was member of the student Pugwash group at Cal where he helped to craft the Engineering Ethics pledge that was listed on commencement brochures.

Brian Smith is an aerospace engineer in the Human Systems Integration Division at NASA Ames. He is a graduate of Occidental College, and has a Master of Engineering from Cal Poly San Luis Obispo. He holds Private Pilot and Part 107 Unmanned Aircraft General (UAG) Pilot Certifications from the FAA. Mr. Smith serves on the Risk Management Working Group (RMWG) in the NASA Aeronautics Research Mission Directorate (ARMD). The role of the RMWG is to identify and analyze risks to ARMD mission and management objectives. Mr. Smith is Ames' representative on the Cyber Security Engineering Group within the ARMD Airspace Operations and Safety Program (AOSP). This group studies and analyzes project technologies with a focus on the operational security environment. Most recently, Mr. Smith was the moderator of the Security and Counter-Drone Perspectives panel at the Urban Air Mobility Symposium held in San Carlos, California, in November 2017. Mr. Smith provides a liaison function between NASA Ames and the Office of Naval Research.

Anupa Bajwa is a researcher in the Intelligent Systems Division at NASA Ames Research Center. She has led the implementation of autonomous capabilities for spacecraft and for unmanned aerial vehicles. She led the Fault Management team on the LADEE mission. She has worked on implementing Integrated Systems Health Management (ISHM) for launch vehicles on projects such as PITEX, Space Launch Initiative, and Constellation. She has extensive training in model-based systems engineering. She holds a doctorate degree in Aerospace Engineering from the Pennsylvania State University.