**A Deontological Ethics Analysis of the IBM Watson Health Failure**

STS Research Paper

Presented to the Faculty of the

School of Engineering and Applied Science

University of Virginia

By

Devin Gardner

March 4th, 2022

Signed: _____

Approved: _____ Date: _____

Benjamin J. Laugelli, Assistant Professor, Department of Engineering and Society

**Introduction**

According to a market analysis report published by Grand View Research, the global artificial intelligence (AI) market was valued at $62.35 billion USD in 2020, and is expected to expand at a compound annual growth rate of 40.2% from 2021 to 2028 (Grand View Research, 2021). The increasing economic growth of the AI market is a strong indication of the recent rise of the field, as many industry executives agree that AI has entered a "golden age" and become an emerging technology of the 21$^{st}$ century (McKendrick, 2019). In the last decade alone, breakthrough developments in AI technology such as long short-term memory networks, backpropagation training, and deep learning have greatly contributed to the popularity of AI in practical applications (Schmidhuber, 2020). Given all of this, it is no wonder why computing powerhouse IBM decided to invest in the AI market early on: after developing the technology for seven years, the famous supercomputer Watson made history in 2011 when it defeated two former *Jeopardy!* champions on national television (IBM, 2011). Riding the wave of Watson's sensational victory, IBM decided to announce the very next day that the next big challenge for Watson would be the healthcare industry. The company was confident Watson would prove to be a useful tool for the field of medicine with its natural language processing (NLP) technology, and they promised the product would be commercially available within eighteen to twenty-four months (Strickland, 2019). This new application of Watson, coined Watson Health, seemed like a good idea on paper: but in reality, IBM would lose billions of dollars on this failed project over the next decade. The IBM Watson Health failure has been documented thoroughly over the years, and many scholars have constructed detailed case studies examining both the financial and technological faults of the product. However, what scholars fail to acknowledge is whether the actions taken by IBM over the course of the Watson Health project were morally virtuous or not.

In this paper, I will investigate the morality of actions taken by IBM employees during the engineering and marketing of Watson Health products. By applying the theory of deontological ethics to this case study, I will demonstrate that these actions were in fact morally responsible, as they followed two of the fundamental rules from the code of ethics proposed by the National Society of Professional Engineers (NSPE): hold paramount the safety of the public, and perform services only in areas of competence. In the end, I will prove that even though IBM overpromised on what they could deliver, ultimately the company handled the failure of the project in a morally acceptable way.

**Background**

According to IBM, Watson is defined as an efficient real-time analytical engine that runs on over ninety servers for a total of 2,880 processor cores (IBM, 2011). The broader goal of the company in creating Watson was to introduce a new generation of technology that could find answers in unstructured data more efficiently than standard search technology. IBM reached a milestone in this mission in late 2010 when Watson became good enough at *Jeopardy!* to win 70% of its games against former champions, and once again in 2011 when Watson made its famous debut on the show (IBM, 2011). Following this, the company decided to pursue applications for Watson in healthcare for three distinct reasons: the market size of the healthcare industry, the promising potential of Watson for AI applications, and recent losses in other computing fields. IBM believed if they could make their mark in the business of healthcare, an industry that has consistently been valued at $3 trillion in the US economy over the last decade, they would be able to compensate for their failing hardware ventures and become a competitor in the artificial intelligence realm – and what better way to do so than with their new flagship product, Watson. IBM established the Watson Health division in 2011, building it up to an

organization of over 7,000 employees and investing over $5 billion in it over the years, only to end up selling a majority of its data and products to a private equity firm for around $1 billion in 2022 (O'Leary, 2022). This sale officially labeled Watson Health as a failed project for IBM, as the company finally removed their stake in healthcare after losing billions of dollars betting on Watson. After countless case studies on the downfall of the division, it was revealed that the failure of Watson Health was primarily attributed to unnecessary complexity and overhead, difficulty of access to healthcare data, an abundance of hype around AI at the time, and a gap between R&D and commercialization (Yang, 2019).

**Literature Review**

Numerous scholars have documented the failure of the Watson Health division, often using it as a case study for investigating the factors that cause the collapse of businesses centered around emerging technologies. These scholarly sources tend to use a variety of evidence, such as verbal accounts from IBM employees or records of business interactions between Watson Health and their clients, to highlight the economic and technical shortcomings of the project. However, what these sources fail to mention is how IBM managed to act morally responsible during the design and testing of Watson Health products, despite the utter financial disaster that ensued.

In *IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care*, Strickland starts out by outlining the dreams IBM initially had for Watson in the field of medicine, with the hope being that Watson "could reduce diagnosis errors, optimize treatments, and even alleviate doctor shortages – not by replacing doctors but by helping them do their jobs faster and better." Strickland then goes on to discuss the shortcomings of the Watson Health division, ultimately concluding that "the Watson Health story is a cautionary tale of hubris and hype". By consulting scholars in the field of AI and former employees of IBM, Strickland was

3

able to gather enough evidence in order to paint a clear picture of the company's goals for Watson Health, as well as the faults that prevented them from meeting those goals. While Strickland does summarize the rise and fall of Watson Health and analyze the factors that led it to fail, she does not once consider the morality of the actions IBM took when conducting the development of this project.

Yang evaluates the actors and agents that led IBM to fail to generate significant profits from Watson Health in his paper *The Rise, Fall, and Resurrection of IBM Watson Health*. In this scholarly source, Yang compiles a list of direct quotes from IBM executives and employees in order to illustrate the strengths and weaknesses of the Watson Health division, such as how IBM is "early in the market of AI" but still "lacks the technological improvements" to make Watson practical. For example, Yang took the words of one employee regarding how some products coming out of Watson Health "weren't even utilizing that much AI or very simple AI models" and used them to demonstrate how one shortcoming of Watson Health was the abuse of the term "artificial intelligence" leading to decreasing market trust. In short, the overall aim of this scholarly source is to answer "why IBM fell short in appropriating value from its innovation – against expectations?" Yang utilizes an abundance of direct evidence from IBM itself to highlight the reasons why IBM did not financially gain from Watson Health, but he does not once question whether or not the IBM employees he interviewed were acting morally virtuous or not.

In terms of providing comprehensive investigations into the factors that led Watson Health to fail economically and technologically, Yang and Strickland did a thorough job. However, both scholarly sources failed to consider the ethical implications of the actions IBM undertook, moreover whether or not IBM was acting morally righteous in the first place. This

paper will use the evidence provided by these scholarly sources in order to expand upon their arguments, exploring how IBM demonstrated ethically responsible behavior despite the economic and technological failures surrounding Watson Health.

**Conceptual Framework**

The ethical framework this paper will be using to evaluate the morality of the actions taken by IBM regarding the Watson Health division is the theory of deontology. The basis of deontological ethics is that "all humans have universal rational duties to one another, centering on their duty to respect the other's humanity" (Misselbrook, 2013). Popularized by German philosopher Immanuel Kant in the late 18th century, the central idea of deontology is that we are all morally obligated to act in accordance with a set of principles and rules regardless of their outcomes ("Kantian Duty Based (Deontological) Ethics", 2013). For instance, if one were to bite into an apple, a deontologist would determine if eating that apple was morally responsible by checking if this action broke any laws or rules, whereas a consequentialist would decide whether eating the apple was right or wrong by weighing the outcomes of the action.

In terms of the Watson Health case, deontological ethics can be applied to evaluate if the actions taken by IBM employees were morally correct or not by checking if the actions followed an accepted set of rules or principles. Since the Watson Health division was focused on engineering innovations in the field of AI, it would be appropriate to use the standards of ethics proposed by the NSPE to determine if IBM followed the proper rules for ethical engineering practices. The six fundamental canons of ethical engineering proposed by the NSPE can be seen below, in Figure 1 (NSPE, 2019).

> **I. Fundamental Canons**
> Engineers, in the fulfillment of their professional duties, shall:
>
> 1. Hold paramount the safety, health, and welfare of the public.
> 2. Perform services only in areas of their competence.
> 3. Issue public statements only in an objective and truthful manner.
> 4. Act for each employer or client as faithful agents or trustees.
> 5. Avoid deceptive acts.
> 6. Conduct themselves honorably, responsibly, ethically, and lawfully so as to enhance the honor, reputation, and usefulness of the profession.

*Figure 1: Code of Ethics for Engineers*

From looking at Figure 1, it can be seen that some of the principles proposed by the NSPE include issuing statements only in the most truthful manner, avoiding deceptive acts, and conducting themselves responsibly and honorably. For this paper however, we will be focused on making sure IBM acted in accordance with the following two rules: protect the safety of the public, and perform only in areas of competency. In my analysis, I will assume each employee of IBM represents the interests of the company, and as such each action of an employee is indicative of the moral standing of IBM itself.

**Analysis**

The executives and employees at IBM should have followed two essential rules during the course of the Watson Health project: protect the safety of the public, and perform only in areas of competency. By making sure IBM acted in accordance with these two rules proposed by the National Society of Professional Engineers, I will prove that even though the Watson Health division was a financial and technological disaster for them, the company still acted morally responsibly nonetheless. The following two sections will closely examine the actions IBM took during the course of the Watson Health project, where each section will concentrate on one of the two rules, and consider if each action taken is in line with said rule.

<u>Protecting the Safety of the Public</u>

First, I shall consider if IBM held paramount the health and safety of the public when acting during the Watson Health project. This begs the question though: how can I differentiate between an action that serves the interest of public safety, and an action that does not? Public safety violations are defined as actions that compromise the safety or overall peace of the public, which can be anything from disorderly conduct to an unauthorized use of hazardous materials ("Public Safety Violations", 2022). By this definition, I can consider any action taken by IBM that poses harm to public safety to be in violation of this NSPE rule. However, this definition leaves too broad of an interpretation for what can be deemed as harmful: this is why the NSPE provides a set of guidelines for each rule in their code of ethics, to help narrow down what explicitly is within bounds of each rule and what is not. The guidelines for the rule regarding public health and safety can be seen below, in Figure 2.
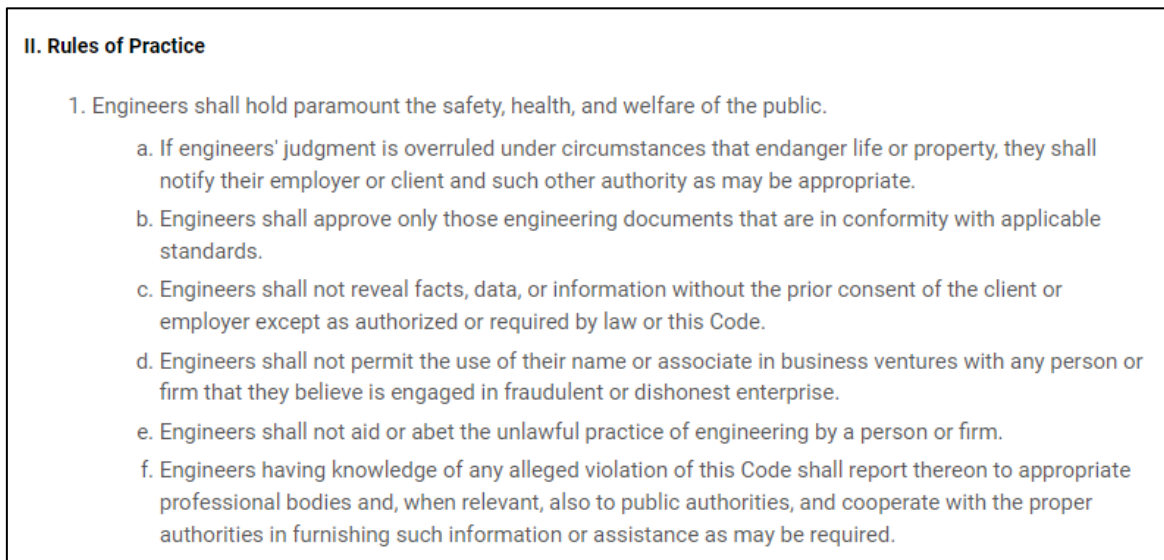


**II. Rules of Practice**

1. Engineers shall hold paramount the safety, health, and welfare of the public.
    a. If engineers' judgment is overruled under circumstances that endanger life or property, they shall notify their employer or client and such other authority as may be appropriate.
    b. Engineers shall approve only those engineering documents that are in conformity with applicable standards.
    c. Engineers shall not reveal facts, data, or information without the prior consent of the client or employer except as authorized or required by law or this Code.
    d. Engineers shall not permit the use of their name or associate in business ventures with any person or firm that they believe is engaged in fraudulent or dishonest enterprise.
    e. Engineers shall not aid or abet the unlawful practice of engineering by a person or firm.
    f. Engineers having knowledge of any alleged violation of this Code shall report thereon to appropriate professional bodies and, when relevant, also to public authorities, and cooperate with the proper authorities in furnishing such information or assistance as may be required.

*Figure 2: Guidelines for Public Safety for Engineers*

Using the guidelines provided in Figure 2, an action can now be deemed as serving the interest of public safety with greater clarity than before. Now, we can accurately determine whether or not

IBM acted morally responsibly in accordance with the public safety rule, and whether or not they kept the health and safety of the public in mind when they took action. As such, it would now be appropriate to begin addressing the actions IBM took during the course of the Watson Health project.

In order to function as an AI doctor, Watson needed an extensive amount of data access to patient records, diagnostic charts, and medical journals. In order to gain access to these resources, IBM spent over $5 billion in buying a series of health companies, including organizations such as Truven, Phytel, Explorys, and Merge (O'Leary, 2022). This gave Watson Health access to a wealth of information for training Watson, including the biggest insurance database in the nation covering over 300 million lives, and a clinical dataset of electronic health records representing over 50 million patients. With this in mind, attention should be drawn to Guideline C of the public health and safety rule: "Engineers shall not reveal facts, data, or information without the prior consent of the client or employer except as authorized or required by law or this code." According to Yang, when Watson Health was asked by developers to make their software tools and datasets open-source, they refused on the basis that they knew they were working with sensitive data protected by privacy laws. Furthermore, they knew that their clients, such as Sloan Kettering and the MD Anderson Cancer Center, would not want them to release the proprietary tools Watson Health was developing for them (Strickland, 2019). Even though making the Watson Health software open-source would actively allow developers around to world to engage with these IBM products, increasing their popularity and leading to the company turning a profit, IBM still chose to prioritize the interests of the public and their clients. In doing so, IBM actively acted in the best interest of public health and safety by not only respecting the rights to data of their clients, but by protecting the privacy of millions of patients' data. Simply put, it can be concluded that IBM acted in

accordance with the public safety rule, and as such acted morally responsibly by choosing to actively protect the health and safety of the public during the course of the Watson Health project.

Performing Only in Areas of Competency

Next, I shall consider if IBM only performed services in their areas of competence during the Watson Health project. Unlike previously, determining whether or not IBM is acting in accordance with this rule is a straightforward approach. A competency area is defined as an area in which a person possesses knowledge, skill, and the ability to apply them in a professional setting ("Competency area definition", 2013). Given this definition, the competency areas of IBM are artificial intelligence, business operations, quantum computing, and IT infrastructure, historically speaking (IBM, 2022). Furthermore, IBM has a rich history in the industries of aerospace and defense, electronics, and healthcare. As such, this means that implementing AI solutions in the healthcare industry with Watson Health should be well within their areas of competency. For further clarification, the guidelines for the rule regarding areas of competency can be consulted, as seen below in Figure 3.
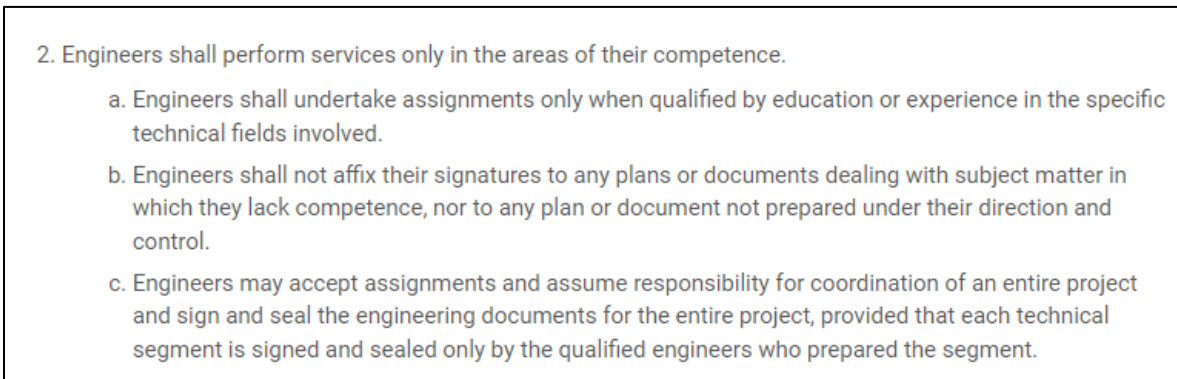
2. Engineers shall perform services only in the areas of their competence.
    a. Engineers shall undertake assignments only when qualified by education or experience in the specific technical fields involved.
    b. Engineers shall not affix their signatures to any plans or documents dealing with subject matter in which they lack competence, nor to any plan or document not prepared under their direction and control.
    c. Engineers may accept assignments and assume responsibility for coordination of an entire project and sign and seal the engineering documents for the entire project, provided that each technical segment is signed and sealed only by the qualified engineers who prepared the segment.

*Figure 3: Guidelines for Areas of Competency for Engineers*

Using the guidelines provided in Figure 3, as well as knowledge regarding the fields and industries IBM is proficient in, it can now be deemed with confidence whether or not the company only acted

in areas of competency during the Watson Health project. More importantly, we can now determine whether or not IBM acted morally responsibly in accordance with the areas of competency rule. As such, it would now be appropriate to begin addressing the professional areas and fields IBM worked in during the course of the Watson Health project.

According to Strickland, one of the primary goals of Watson Health was to bring the Watson AI into medical practices to act as a tool for "clinical decision support". Martin Kohn, the chief medical scientist for IBM Research at the time of the famed *Jeopardy!* match, stated the visionary idea was that "Watson could read patients' health records as well as the entire corpus of medical literature: textbooks, peer-reviewed articles, lists of approved drugs, and so on … with access to all this data, Watson might become a superdoctor, discerning patterns that no human could ever spot" (Strickland, 2019). With this in mind, Watson Health aimed to implement a diagnostic tool that would allow doctors to "give Watson a patient's case history and ask for a diagnosis or optimal treatment plan." However, after working on a prototype for the diagnostic tool for two years, IBM realized that AI could not replace the diagnosis abilities of doctors at this point, and voluntarily gave up on commercializing the product in 2014 (Strickland, 2019). Even IBM's Vice President of Health Care and Life Sciences Research, Ajay Royyuru, admitted that their technological expertise was not best suited for this project, saying "Diagnosis is not the place to go. That's something the experts do pretty well. It's a hard task, and no matter how well you do it with AI, it's not going to replace the expert practitioner" (Strickland, 2019). At this point, attention should be given to Guideline A of the areas of competency rule: "Engineers shall undertake assignments only when qualified by education or experience in the specific technical fields involved." When presented with the chance to continue working on the diagnostic tool, to push it to market and attempt to profit off of it despite the harmful disadvantages of it, IBM decided

against it. IBM knew that taking the diagnostic tool to market would mean competing with the medical expertise of experienced doctors, so they chose to exit an area they felt they were not competent in, knowing that commercializing a poor diagnostic tool would be hazardous to the integrity of medical practices. Moreover, IBM could have attempted to cover up their mistakes with the diagnostic tool and continue working in this area nonetheless, but I believe their choice to make a public statement about moving away from an area they feel as though they aren't proficient in speaks volumes about their moral character. In short, it can be concluded that IBM acted in accordance with the areas of competency rule, and as such acted morally responsibly by choosing to actively protect the integrity of medical practices by only acting in areas they could perform knowledgably in during the course of the Watson Health project.

**Conclusion**

Overall, by applying deontological ethics to the case study of the IBM Watson Health failure, it can be seen that IBM did in fact act morally responsible despite the drastic failures the Watson Health division underwent. With the technology not being there and financial acquisitions not making returns, one would think that would turn a company to unethical practices in order to make a profit; but IBM chose the morally righteous path, chose to protect the privacy of its clients, and step away from areas it felt it was not competent in. It is by no means easy to undergo a massive technological and economic failure like IBM did with Watson Health, but for this company, it seems like it was easy to make the right choices at the end of the day.

<div align="center">**References**</div>

Grand View Research. (2021, June). *Artificial Intelligence Market Size, Share & Trends Analysis*

    *Report By Solution, By Technology (Deep Learning, Machine Learning, Natural*

    *Language Processing, Machine Vision), By End Use, By Region, And Segment Forecasts,*

    *2021 – 2028.* https://www.grandviewresearch.com/industry-analysis/artificial-

    intelligence-ai-market.

IBM. (2011). *A Computer Called Watson*. https://www.ibm.com/ibm/history/ibm100/us/en/

    icons/watson/.

IBM. (2022). *Products & Solutions*. https://www.ibm.com/us-en?utm_content=SRCWW&p1=

    Search&p4=43700050478421002&p5=e&gclid=Cj0KCQjw5-WRBhCKARIsAAId9F

    mt6X6YPHeDAXbDSEsNVLfi78V6Br0K2keH3TWmgMUdeEl5zFjX7IaAln7EALw_w

    cB&gclsrc=aw.ds.

*Kantian Duty Based (Deontological) Ethics*. (2013, January 29). Seven Pillars Institute.

    https://sevenpillarsinstitute.org/kantian-duty-based-deontological-ethics/.

*Competency area definition*. (2013). Law Insider. https://www.lawinsider.com/dictionary/

    competency-area#:~:text=Competency%20area%20means%20an%20area,them%20in

    %20the%20clinical%20setting.

McKendrick, Joe. (2019, October 23). *Artificial Intelligence Enters Its Golden Age*. Forbes.

    https://www.forbes.com/sites/joemckendrick/2019/10/23/artificial-intelligence-enters-its-

    golden-age/?sh=21c93a65734e.

Misselbrook, David. (2013, April). *Duty, Kant, and Deontology*. National Library of Medicine. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3609464/.

NSPE. (2019, July). *NSPE Code of Ethics for Engineers*. https://www.nspe.org/resources /ethics/code-ethics#:~:text=Engineers%20shall%20be%20guided%20in,project% 20will%20not%20be%20successful.

O'Leary, Lizzie. (2022, January 31). *How IBM's Watson Went From the Future of Health Care to Sold Off for Parts*. Slate. https://slate.com/technology/2022/01/ibm-watson-health-failure-artificial-intelligence.html.

*Public Safety Violations*. (2022). FindLaw. https://www.findlaw.com/criminal/criminal-charges/public-safety-violations.html#:~:text=Public%20safety%20violations%20 include%20disorderly,varying%20from%20state%20to%20state.

Schmidhuber, Jurgen. (2020, February 20). *The 2010s: Our Decade of Deep Learning*. AI Blog. https://people.idsia.ch/~juergen/2010s-our-decade-of-deep-learning.html.

Strickland, Eliza. (2019, April 1). *IBM Watson, heal thyself: How IBM overpromised and underdelivered on AI health care*. IEEE Xplore. https://ieeexplore.ieee.org/document/ 8678513.

Yang, Jialei. (2019). *The Rise, Fall, and Resurrection of IBM Watson Health*. University of California Berkeley. http://jultika.oulu.fi/files/nbnfi-fe2020050424858.pdf.