



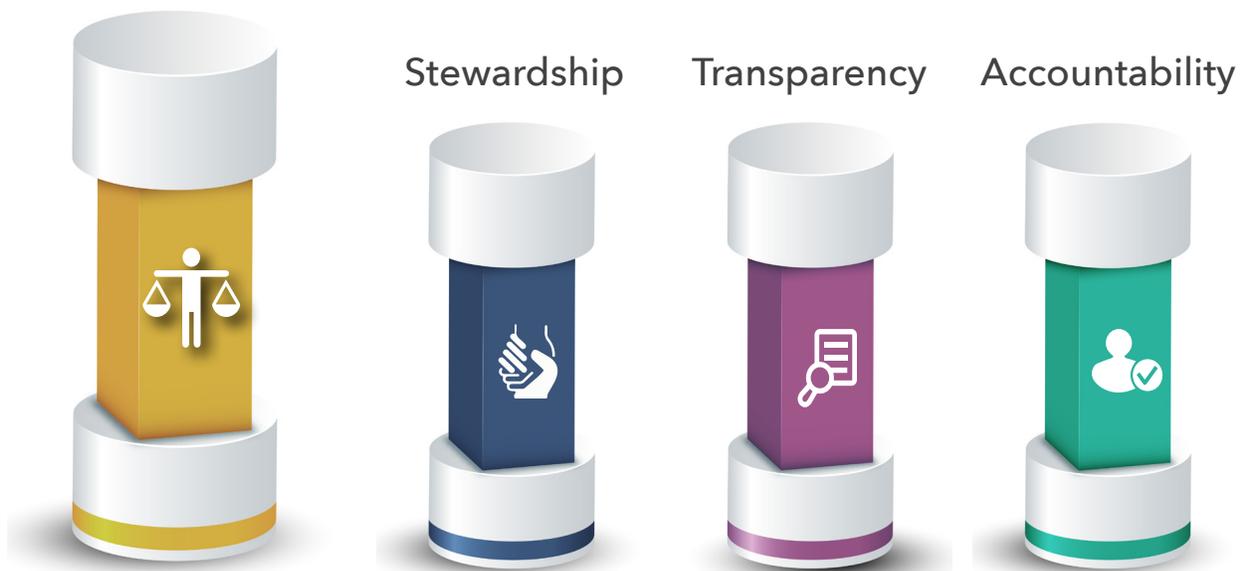
# MHS

Beyond Assessments

## Ethics

At MHS, we are stewards of the personal information our customers provide. We use that data to understand individual behaviour and develop effective interventions. As such we believe that it is imperative that we have a well-articulated and easy to follow Trust Framework that governs how we interact with the data that we are entrusted with. The MHS Trust Framework has 4 pillars, ethics, stewardship, transparency and accountability. This paper delves deeper into the ethics pillar to provide more context around our ethics statement.

## Ethics



**When I first started working on the overall Trust Framework for MHS**, I saw ethics solely through the lens of my experience, training, and background in psychology and as a business leader. I would have defined it as 'a set of guidelines that provide structure and rules to professions and professionals', a moral code that somehow our training and education uniquely prepared us to abide by. After further research, it has become clear to me that I was oversimplifying a deep and complex topic.

Ethics, and in particular the ethical use of data in an age of artificial intelligence, algorithms, and prediction, has generated much discussion recently, yet debates about the ethical use of information and predictive results date back at least to the dawn of the computer age with many 'schools' of thought each attempting to definitively provide a structure or framework that succinctly defines a very messy topic .

I grew up on a farm where things just had to get done. It's in my DNA to work hard and to finish what I start. I was raised in an environment that was far too practical to enjoy lengthy debates that seemingly don't have a right or wrong answer. You ate the food you grew, even if it included animals that you personally cared for. There was no right or wrong, there was just practicalities. As a result, Philosophical topics have, until recently, never been inherently interesting to me. Things changed as I assumed the CEO role for Multi-Health Systems and led the effort to redefine a 35 year young success story for the digital age. It quickly became evident that I needed to further explore and explain what we meant by ethics in our Trust Framework, I had to embark on a journey deeper into territory that was not only unfamiliar but also at times uncomfortable. To be clear, over my career there have been

many times where I have had to invoke my ethical principles to solve a business problem, it is just that until now, I haven't had to examine the foundations of my ethical principles or more importantly how they fit into the greater debate around ethics in a digital and data driven global economy.

Ethics is the first of the four pillars of our MHS Trust Framework , with data stewardship, transparency and accountability being the other three. Our statement on ethics reads:

***"we commit to ensure that our use of data and the outcomes, predictions, prescriptions and actions we take based on data, will always be conducted in an ethical fashion, ensuring rigorous attention is paid to ensure bias (conscious and unconscious) is removed from our products and services, always with the interest of the individual and their community as the guiding factor."***

## Ethical Frameworks

Ethical theories can be organized into three main groupings

- Metaethics
- Normative ethics
- Applied ethics.

Metaethics as its name implies looks at the big picture. Where do our ethical principles come from? What role do reason, the will of god and other factors, play in establishing our ethical principles. Normative ethics takes a more practical approach and examines the moral standards that guide the establishment of right and wrong, while applied ethics concerns itself with examining specific controversies such as capital punishment, environmental concerns and assisted suicide.

When attempting to resolve the issues raised in applied ethics, the frameworks provided by Metaethics and normative ethics are invoked as tools that can be used to provide answers or guide discussions. If we look at the world of Artificial Intelligence as an applied ethical issue, then we can use normative ethics and metaethics to help us provide guidance in resolving the inherent ethical dilemmas that arise when considering when, where, and how to apply AI techniques. For the purpose of this paper, I am going to leave metaethics aside, since it offers little in terms of practical applications in day to day life and focus on normative ethics.

## Ethics in an AI World

If the application of Artificial Intelligence is an applied ethical issue, and normative ethics map to the problem domain, the challenge is to identify which normative ethics principles we choose? Normative ethics can cover everything from rules and regulations provided by the government, to professional standards and codes of conduct developed for specific industries, to our own individual 'moral' standards.

When it comes to the issue of ethics in modern complex topics such as AI, and within the AI field the data intensive Machine Learning techniques, it is only recently that legislative bodies have begun to create comprehensive rules and regulations to govern the use and application of these advanced technologies. As the media reports more on the outcomes that data and technology are creating all levels of government will increasingly follow the lead of the European Union in drafting regulation on the use and transparency of data in decision making.

When attempting to resolve the issues raised in applied ethics, the frameworks provided by Metaethics and normative ethics are invoked as tools that can be used to provide answers or guide discussions. If we look at the world of Artificial Intelligence as an applied ethical issue, then we can use normative ethics and metaethics to help us provide guidance in resolving the inherent ethical dilemmas that arise when considering when, where, and how to apply AI techniques. For the purpose of this paper, I am going to leave metaethics aside, since it offers little in terms of practical applications in day to day life and focus on normative ethics.

As the sea of often conflicting regulations and legal guidelines continues to [d]evolve, perhaps we can turn to professional standards for guidance. MHS' solutions and services are primarily consumed by those in 'helping professions', Psychologists, business and life coaches, Human Resources personnel, Probation and Parole workers, and Social Workers for example. The majority of these professionals have codes of conduct or rules that govern their actions. I don't know of any that yet specifically mention how to work with technologies such as AI, but I do know that they are all based on principles of 'do no harm' with healthy doses of 'respect your professional limits' and 'client confidentiality' thrown in and they allow individuals a great deal of personal and professional judgment in applying the standards.

With these open ended professional standards as guides, we are then left with our own individual standards, and by individual, I don't just mean us as single individuals but also as individual organizations. It is at this point that organizations such as MHS need to develop and implement ethical standards that govern how they operate in this increasingly complex environment.

## Applied Ethics in Practice

How do we put this theory into practice? How does the above discussion on ethics help MHS to interpret or apply the ethics statement component of our trust framework and how do our clients hold us accountable to our words?

Let's look at an example. Company A is hiring for a very technical role. They have a clearly defined set of competencies and job requirements. They are able to track performance of existing people in the role and have performance metrics with solid objective measurements (remember, this is a hypothetical scenario). They have used the historic performance data as well as the requirements for the new role to train an algorithm, allowing it to sort and prioritize applicants. To further expand the scenario, consider also that Company A has been called out in the industry for having very poor representation of women and visible minorities in their technical employee pool, and as a result have set diversity targets they are committed to meeting. In this situation, how do they ensure they get the best qualified person for the job while also meeting their diversity targets? What if two or more applicants are equally qualified, that is they get the exact same result from the algorithm? Which person gets hired?

The reality is that most algorithms are not currently built or trained to handle decisions of this complexity nor to recognize the ethical component of the recommendation. In a hiring scenario, if two candidates are equally qualified according to the selection criteria, how does an algorithm prescribe which candidate to make an offer to? Further, algorithms also struggle with competing objectives such as how to balance qualifications with diversity requirements.

In cases like this it is essential that an algorithm not make the final decision, operate with autonomy in the AI vernacular, but to make a prescriptive recommendation that informs the decision of a subject matter expert. But how do we get people to not rely solely on the recommendations of an algorithm but rather to consider the input of the algorithm in a broader decision making framework? After all, the comment that "because the computer says so" is often heard to justify action. What we seek is to retain decision making by the "human in the loop" or the HIL as it is technically known as.

There are a couple of techniques available in AI systems that can support the retaining of the decision at the human level. These include using competing algorithms that potentially provide different results, or programming algorithms to have probability factors included that introduce uncertainty thereby providing greater diversity and transparency regarding trade-offs that might need to be made to arrive at a final decision. These techniques would not produce a single score but a range of results that would require the human to use judgment in the decision making. There are also organizational values and ethical principles that can be developed and applied, just as we are doing at MHS, to guide the humans away from over relying on computer scores.

## Retaining the Human in an Increasingly Automated World

Humans, have the capacity to understand, internalize and take actions or make decisions based on ethical principles. Whether we act on this capacity or not is a choice. AI systems, at least in their current form and thankfully absent the capacity for artificial general intelligence (AGI), does not have the ability to make decisions on ethical principles. Given this, it is MHS' position that we must default, at least in areas that involve decisions with high stakes outcomes for individuals, to the concept of Intelligent Assistance rather than Artificial Intelligence, ensuring that transparency into how a recommendation was arrived at, with what probability and a measured and re-creatable certainty index is maintained. When we reconceptualise AI as IA, we ensure that we retain a HIL to ensure that decision making is not left entirely to algorithms, rather we leverage algorithms to do the computational work leaving humans to apply judgment. An example of how AI as IA is shown below (Figure 1: IA not AI).

**Figure 1: IA not AI**

AI in an IA capacity creates great advantage to those who utilize it, but with great power comes even greater responsibility which is why we adhere to the MHS Trust Framework in our actions, services, and products.



MHS is a trusted, global, neuroscience and bioinformatics company whose purpose is to provide data driven insights that predict and improve individual and organizational success.