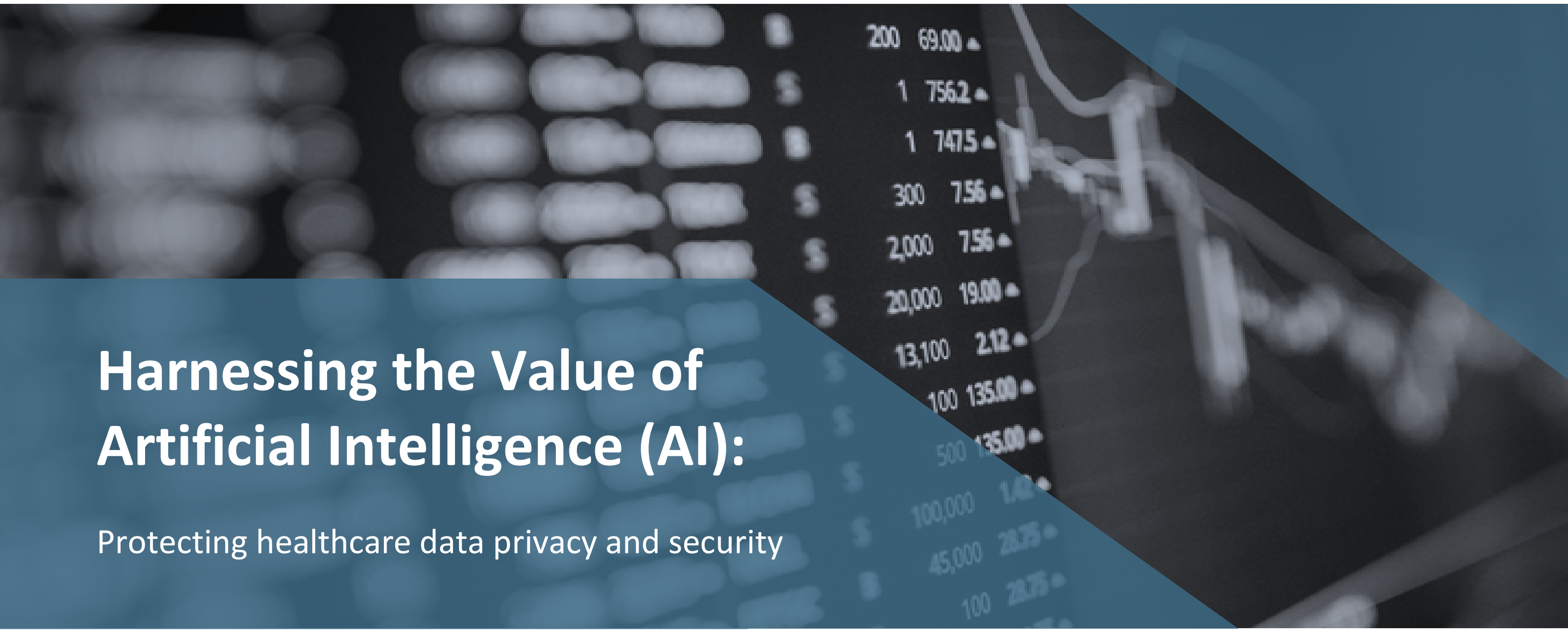


Harnessing the Value of Artificial Intelligence (AI):

Protecting healthcare data privacy and security

AI at its core encompasses taking advantage of the rich and diverse sources of information that individuals and health care providers amass across the continuum of health-related behaviors, interactions, and management. Being able to better leverage these stores of data to enable predictions of health risk and more proactive engagement to promote better health care outcomes via AI tools and applications offers great promise. With this comes a responsibility to safeguard and ensure responsible stewardship of sensitive information. This includes both the use of the data itself as well as the output, predictions and actions that ensue.

Striking the right balance between utilizing patient data to advance healthcare and ensuring robust privacy protections is essential for fostering trust in AI applications.

Addressing these risks in the new world of AI requires a multifaceted approach that involves transparency, responsible data practices, robust security measures, and ongoing evaluation. By acknowledging and proactively mitigating risks, healthcare organizations can work toward successfully harnessing the potential of AI while safeguarding patient welfare and upholding ethical standards.

So, what is it about AI that creates unparalleled risk relative to data security and privacy?

AI (particularly generative AI) introduces several distinct privacy concerns due to its ability to process personal data along with other information in discovery of patterns and insights that may generate sensitive information or novel identification of individuals and their circumstances – that goes beyond what is presented in the data itself.

The difference between what is currently considered "mainstream" and generative AI lies in respective capabilities and application, in this regard:

Mainstream AI –

Capability - to learn from data and make predictions based on that input data. Making smart decisions within a prescribed set of rules.

Application – pattern recognition

Generative AI –

Capability - to learn underlying patterns once trained on a set of data and then generate new data that mirrors the training data set.

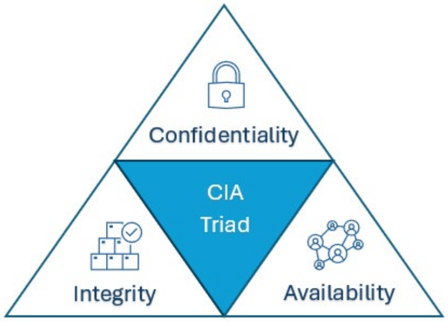
Application – pattern and/or data generation

Mainstream AI systems are primarily used to analyze data and make predictions. Mainstream AI excels in task-specific applications wherein there is a lot of data and/or the need for multiple iterations to complete analysis and generate results.

Generative AI goes a step further by actually creating new data (corresponding to its training dataset), yielding additional data fields not present in the original (training) dataset. Generative AI has the potential to add value where creation of novel insights are desired. Other examples of generative AI include creating never-before seen images, as well as imaginative and/or inventive narratives and stories.

These two types of AI are not mutually exclusive. As a case in point, mainstream AI can analyze consumer behavior data and generative AI can use that analysis to create personalized content or data. This simple example serves to illustrate the need to evaluate risk considerations based upon the AI methodology used. Mainstream AI analysis based on a data set will carry different impact, likelihood and risk levels than a generative AI analysis that creates new data upon which decisions may be made or actions taken.

Traditional Data and Information Security CIA Triad Applied to AI



Let’s take a further look into types of risk that can be associated with AI from a security and privacy perspective and look at a few ways to mitigate that risk. The traditional security CIA (confidentiality, integrity, availability) Triad offers a high-level approach to guide our discussion.

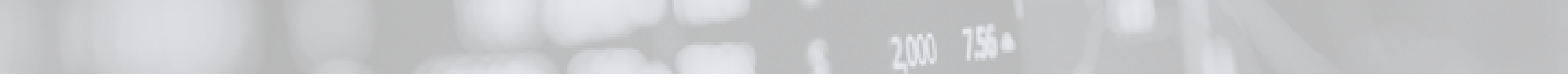
But first, no discussion of security and privacy can ensue without first recognizing the need for company standards and policy to guide the use of AI. While most health care organizations have data security and privacy guidelines in place that can readily be made applicable, don’t assume employees or others accessing data will recognize how these guidelines apply when using AI tools.

Many companies are beginning the exploration of AI and may not feel ready to publish a policy, but user standards should be developed and shared now with staff. These standards or guidelines should remind staff of long-standing policy and process regarding the approved use of data, the responsibility to protect data and tie these elements to examples so the workforce member can associate security practices with the use of AI.

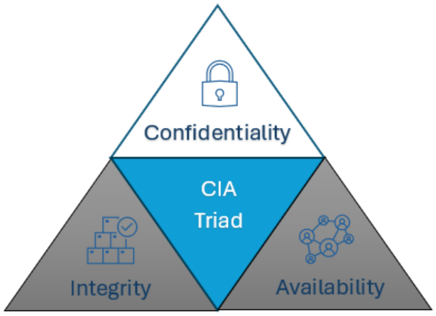
This accessing of AI tools to perform even the most rudimentary tasks is likely already taking place within your organization, so it is worth taking steps right now to clarify that existing data security and privacy guidelines apply. Case in point, an employee is tasked with improving the readability of a procedure document, so they upload it to a free AI tool to obtain suggestions. Seemingly innocuous; however, if that procedure describes a company’s ‘secret sauce’ it has now effectively been shared with all users of that free AI tool. The leap to sharing data and the consequences is easy to make and users should be reminded of the applicability of security measures to the use of AI.

Characteristics associated with responsible use of AI that policies and/or procedures may address include:

- ✓ **Fairness** – always ask the question “will we leave some groups of people worse off because of the algorithm’s design or its unintended consequences”.
- ✓ **Transparency** – disclosures relating to use of algorithm generated decision-making.
- ✓ **Explainability** - while transparency may offer advance notice and/or opportunity to deny access to personal information, explainability involved “after the fact” disclosure about the use of algorithms in specific outcomes/decisions.



Confidentiality



AI is being held up as the next wave of improved accessibility and one can readily understand the applicability. Just as we moved from the notecards of library card catalogs to digital listings and online searches, we can now quickly query the Library of Congress with a few well-developed questions, typed or verbal. This accessibility in terms of healthcare data confidentiality will be challenged by these new AI-driven capabilities.

Questions to be addressed may include:

- ? How are strict data accessibility standards maintained throughout the AI process, including access to generated results.
- ? How do inherent biases in the source data impact accessibility? As we noted earlier in this article, bias in AI-generated data can lead to unintentional exclusion of certain groups from treatment protocols, program participation or other health care related opportunities.

A model focused on data collection and processing may mitigate risk/algorithmic discrimination in several ways:

- ✓ Data stewardship requirements aimed at safeguarding uses of personal information.
- ✓ Data transparency or disclosure rules, as well as individual rights to access information relating to them.
- ✓ Data governance rules that prescribe appointment of a privacy officer, conduct of privacy impact assessments, or product planning ("privacy by design").
- ✓ Rules on data collection and sharing to reduce risk associated with the aggregation of data that enable inferences and predictions but may involve some trade-offs with the benefits of large and diverse datasets.

Two important elements relative to Confidentiality in AI worth mentioning are risk of Data Breach and use of Consumer Generated Identifiers:

Data Breach

AI-powered systems rely on vast amounts of data, which can become vulnerable to data breaches and cyberattacks.

AI algorithms can use seemingly innocuous data to make inferences about people's private information, such as their political views or sexual orientation. This can be a violation of privacy and can lead to discrimination or stigmatization.

Algorithms can also inadvertently lead to discovering the identity of individuals within a source dataset that has been redacted (personal identifiers removed to protect anonymity [reidentification]).

These risks illustrate the need to implement processes to constrain data and generated from: 1) being used to further train the AI model beyond its intended use; or, 2) leaking out to the public.

Use of Consumer Generated (personal information) Identifiers

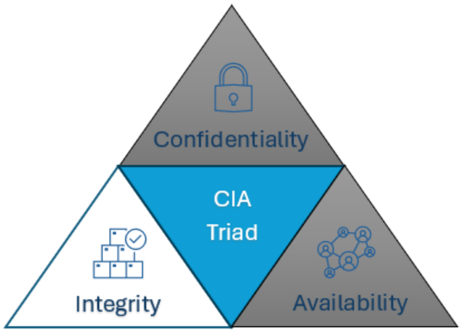
Use of personal data to feed algorithms raises the bar with respect to policies and procedures for data storage, usage, and access.

- ? Where is this data coming from?
- ? Where is the data stored?
- ? Who can access it, and under what circumstances?

Steps must be taken to minimize the data used to that which is absolutely necessary, coupled with consent-based data collection where feasible. Policies and procedures need to clearly specify where a user will have access to what data and for what purpose. Access needs to be limited to those who "need to know" and only for the purpose(s) specified.

A simple example involves the use of human resource information which may include pay levels, hiring salary, history of promotions and/or job classification changes. Salary information is initially “scrubbed” from the data, but AI might be able to “predict” current salaries based on these other inputs, thus revealing information meant to be hidden from view.

Integrity



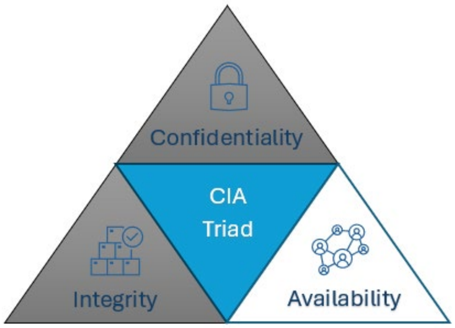
Another of the three central principles of data security is integrity and data integrity requires that we maintain the accuracy, consistency, and reliability of data. Some relevant methods of achieving this include ensuring strict data validation processes are followed; granular access roles limit the altering, deletion, or other manipulation of data; the review of data outputs; as well as standard technical techniques such as encryption in transit.

Extracting the greatest value from AI is supported by data integrity as high-quality input data is required and organizations should re-evaluate their data’s fidelity and accuracy with an eye toward how it will be utilized by, and potentially transformed with, AI. But that is just the start. Steps to ensure ongoing integrity need to encompass everything that happens downstream. Organizations must ensure that designated AI users and tools maintain the integrity of the source data as well as new AI-generated values.

Questions to be addressed may include:

- ? How is data validation maintained throughout the AI process?
- ? How is new AI-generated data monitored on an on-going basis to ensure that data does not evolve or leak outside the parameters defined for that generated data and its use? Simply put, source data accuracy must be maintained by ensuring AI-generated data is not co-mingled in an un-managed fashion.
- ? Is AI-generated output encrypted as it transits systems and comes to rest?
- ? How does a company identify and account for inherent biases in the source data that may impact integrity? If AI systems are “trained” on biased data this can lead to an amplification of those biases, leading to unfair or inaccurate decisions. This could have far-reaching consequences, such as disparities in diagnoses or treatment recommendations across different socio-economic groups.

Availability



Ensuring availability of healthcare data for typical healthcare operations requires the diligent maintenance and management of systems transmitting, holding, processing, storing, and displaying that data.

Closely tied to data integrity, AI requires large, readily available data sets to generate optimal results.

Questions to be addressed may include:

- ? Will healthcare operations be impacted by processing power used by AI?
- ? Will data availability for AI use be hindered by architectures built on the principle of minimum necessary utilizing segmentation or segregation of data?
- ? Conversely, will secure data architectures be challenged by AI-driven data availability requirements?
- ? Privacy concerns must be considered and contractual elements in data use agreements, patient privacy notices, etc. must be reviewed in light of the use of data in the training of AI tools or generation of new potentially identifiable data.

Ethical questions addressing availability of data and by extension healthcare will be debated for some time. For example, does the ability to identify trends and needs resulting in the extension of healthcare services to previously underserved populations offset concerns of massive data stores being made available to AI engines owned by private companies? Such discussions are beyond the scope of this article but offer food for thought.

Weighing the Risk and Benefits of AI in Healthcare, Human Intelligence Will Always Be Important

AI has potential to change the medical industry in the future but there are certain attributes that simply cannot be achieved with “smart” technology. That goes for efforts to protect data privacy and security, but also deriving value overall. This requires thoughtful consideration of new organizational constructs and accountabilities that establish parameters around the differential use of AI as technology versus the human touch.

AI is a technology that is meant to augment, not replace human judgement and critical thinking. In fact, many organizations have established policies that specify AI’s role as an adjunct to human decision making (not a replacement or stand-alone). There are countless benefits of AI in healthcare, but when accuracy matters, there is no replacement for human intelligence. By augmenting human intelligence with technology, healthcare organizations can reach new peaks of efficiency and productivity. Having a “human in the loop” is a critical consideration.

This is the main approach used by the European Union’s General Data Protection Regulation (GDPR) – simply put, “the human impacted has recourse to another human who can review the decision and explain the logic behind it”.

Healthcare professionals will continue to see their work evolve as technology helps streamline some of the more data-intensive and time-consuming processes. However, human judgement and critical thinking will be required to consistently evaluate AI output and be prepared to identify and intervene where and when inevitable errors occur.

Considerations around how to address staffing, skill sets and other potentially new organizational constructs will be discussed further in our follow up article about operational readiness for AI.