

# Ethical Use of Electronic Health Record Data and Artificial Intelligence: Recommendations of the Primary Care Informatics Working Group of the International Medical Informatics Association

Siaw-Teng Liaw<sup>1</sup>, Harshana Liyanage<sup>2</sup>, Craig Kuziemy<sup>3</sup>, Amanda L. Terry<sup>4</sup>, Richard Schreiber<sup>5</sup>, Jitendra Jonnagaddala<sup>1</sup>, Simon de Lusignan<sup>2</sup>

<sup>1</sup> WHO Collaborating Centre on eHealth, School of Public Health & Community Medicine, UNSW Sydney, Botany Road, Kensington, NSW 2033, Australia

<sup>2</sup> Clinical Informatics and Health Outcomes Research Group, Nuffield Department of Primary Care Health Sciences, University of Oxford, Eagle House, 7 Walton Well Road, Oxford, OX2 6ED, UK

<sup>3</sup> Office of Research Services, MacEwan University, Edmonton, Alberta, Canada

<sup>4</sup> Centre for Studies in Family Medicine, Department of Family Medicine, Department of Epidemiology & Biostatistics, Schulich Interfaculty Program in Public Health, Schulich School of Medicine & Dentistry, Western University, Canada

<sup>5</sup> Internal Medicine and Informatics, Geisinger Health System and Geisinger Commonwealth School of Medicine, Camp Hill, PA, United States

## Summary

**Objective:** To create practical recommendations for the curation of routinely collected health data and artificial intelligence (AI) in primary care with a focus on ensuring their ethical use.

**Methods:** We defined data curation as the process of management of data throughout its lifecycle to ensure it can be used into the future. We used a literature review and Delphi exercises to capture insights from the Primary Care Informatics Working Group (PCIWG) of the International Medical Informatics Association (IMIA).

**Results:** We created six recommendations: (1) Ensure consent and formal process to govern access and sharing throughout the data life cycle; (2) Sustainable data creation/collection requires

trust and permission; (3) Pay attention to Extract-Transform-Load (ETL) processes as they may have unrecognised risks; (4) Integrate data governance and data quality management to support clinical practice in integrated care systems; (5) Recognise the need for new processes to address the ethical issues arising from AI in primary care; (6) Apply an ethical framework mapped to the data life cycle, including an assessment of data quality to achieve effective data curation.

**Conclusions:** The ethical use of data needs to be integrated within the curation process, hence running throughout the data lifecycle. Current information systems may not fully detect the risks associated with ETL and AI; they need careful scrutiny. With distributed

integrated care systems where data are often used remote from documentation, harmonised data quality assessment, management, and governance is important. These recommendations should help maintain trust and connectedness in contemporary information systems and planned developments.

## Keywords

Ethics, clinical; data accuracy; information systems; medical record systems, computerized; Artificial Intelligence

Yearb Med Inform 2020

<http://dx.doi.org/10.1055/s-0040-1701980>

## Introduction

The 2018 WHO-UNICEF Global Primary Health Care (PHC) Conference in Astana, Kazakhstan, marked the 40<sup>th</sup> anniversary of the Declaration of Alma Ata and stressed the need to “*better use current systems and data to improve the science and to innovate with artificial intelligence (AI) techniques for electronic decision support and data analytics to achieve Universal Health Coverage and better and fairer health care for individuals and populations.*”

The Astana Declaration highlighted the universal aspiration for interoperability of health information systems (HIS) to enable data sharing to support integrated person-centred health services [1] as well as to link HIS data with national public health datasets, disease and health services registries, death registry, genomic databases and biobanks for research, evaluation and quality improvement. Digital health tools are essential to deliver personalised medicine [2]. AI techniques such as machine learning and algorithmic categorisa-

tion can make sense of complex big data emerging from sensors, wearable devices, clinical observations, clinical trials, social and online platforms, providing insights into the behaviours and physiology of individuals [3].

However, health-related data generated and used outside of clinical settings are not protected by privacy regulations. This enables commercial data aggregators to legally combine individual behavioural and social data from multiple sources for health and other purposes. When incorporated

into HIS, these data blur the distinctions between different categories of protected health data, and between protected data and data collected via commercial apps and services. This dynamic landscape presents significant ethical, technical, and information/data governance challenges to the global health information ecosystem.

Obvious challenges are compromises to individual privacy, including identity theft, and the opaqueness of data analytics. Data repositories are difficult to access to verify and validate the quality of the data and algorithms used [4, 5]. A lack of community engagement, trust, and ethical understanding of AI may distort legislation and policies and be rejected by the community, which could hamper the acceptance and advancement of data science and informatics [6]. In addition, most in the health care software industry have difficulty complying with robust cybersecurity standards and mandatory certification for information security management and personal health information protection, despite strict regulations such as Meaningful Use/Promoting Interoperability in USA [7].

Our objective is to create practical recommendations for the curation of routinely collected health data and AI in primary care, with a focus on ensuring their ethical use for contemporary health practice. This is important with integrated health practice where data are shared and analysed for use distant from the point of data creation and documentation. We defined data curation as the management of data throughout its lifecycle, extending from its initial collection and storage for use by data custodians to support care delivery and for analyses and cataloguing for secondary use [8].

## Methods

The IMIA PCIWG has conducted a series of Delphi groups to research the ethical dimension of informatics and data curation initiatives in primary care since 2011 [9–12]. Our 2016 Privacy, Ethics, and Data Access Framework for Real World EHR Data was the starting point for this work.

This framework lists 14 ethical principles to guide data custodianship and appropriate access to big data by various stakeholders [10]. The 14 ethical principles include: Autonomy, Respect rights and dignity of patients, Respect clinical judgment, Duty to provide care, Protection of the public from harm\*, Beneficence, Justice, Non-maleficence (obligation to not inflict harm intentionally), Reciprocity, Solidarity\*, Stewardship, Trust, Lawfulness\*, and Transparent project approval process. The principles combine principlist and communitarian (denoted with an \*) approaches [13].

We incorporated the FAIR guiding principles for scientific data management: *Findability, Accessibility, Interoperability and Re-usability* [14]. The FAIR principles work within the broader five fair information practice principles (FIPPS): *transparency, use limitation, access and correction, data quality, and security*. The FIPPS are the basis for many privacy regulations [15].

While recognised and enshrined in many international conventions, legislations, regulations, and guidelines, the combined principles are unevenly implemented across countries due to the different country contexts [16]. The context includes cultural bias or differences in demographics or professional groups; a paternalistic perspective, societal issues such as equity and legality, or information sharing as they relate to individual citizens and communities, health professionals, health organisations, and health systems.

Peer-reviewed electronic databases were searched using the terms “ethics” and “health informatics” to support a focused literature review of relevant papers since 2016. After consensus on relevance, ethical challenges were extracted and categorised using the conceptual framework across the entire data life cycle [12]. The challenges at each stage of the data life cycle were categorised as technical, management, organisational, governance, data quality and accessibility. The findings were contextualised to emphasise a patient- and clinician-centred approach to health data, informatics, and care. Recommendations were then developed through an iterative process.

## Results

Despite significant computerisation in practice, there were numerous gaps and challenges in the data life cycle: collection, documentation, storage, management, sharing and use [12, 17]. The health information ecosystem is highly fragmented in Australia [18], Asia [19], Canada [20], and USA [21] with poorly defined protocols for information sharing among providers, settings, and jurisdictions. Many data repositories did not meet the basic legal, technical, and organisational principles for sharing data across settings. Solutions may include a Health Data Research body in the UK to address the jurisdictional fragmentation caused by the separation of the NHS across the four devolved nations (England, Wales, Scotland, and Northern Ireland).

### Recommendation 1: Ensure consent or formal processes to govern access and sharing throughout the data life cycle

Good governance should ensure there is consent or other legal processes (e.g., opt out or use for public health purposes) governing access and sharing. Figure 1 includes the FAIR principles, data quality categories, and information governance, and where they are relevant in the data life cycle. New and often disruptive technologies such as blockchain are being used to accelerate the consent process for clinical trials [22], potentially addressing personal data privacy concerns associated with data sharing in integrated service delivery.

General Practitioners (GPs) have significant roles as creators/collectors, managers, and users of observational health data. Do we need new models of consent from GPs, patients, and other health care providers in the referral and integrated care network at different points in the data life cycle? How do we decide if consent is informed and relevant, and at which point in the life cycle?

### Recommendation 2: Implement sustainable data creation/collection to build and maintain trust as well as permission

Community-driven health data repositories may not be as private as citizens assume [23]. This is particularly true of free online

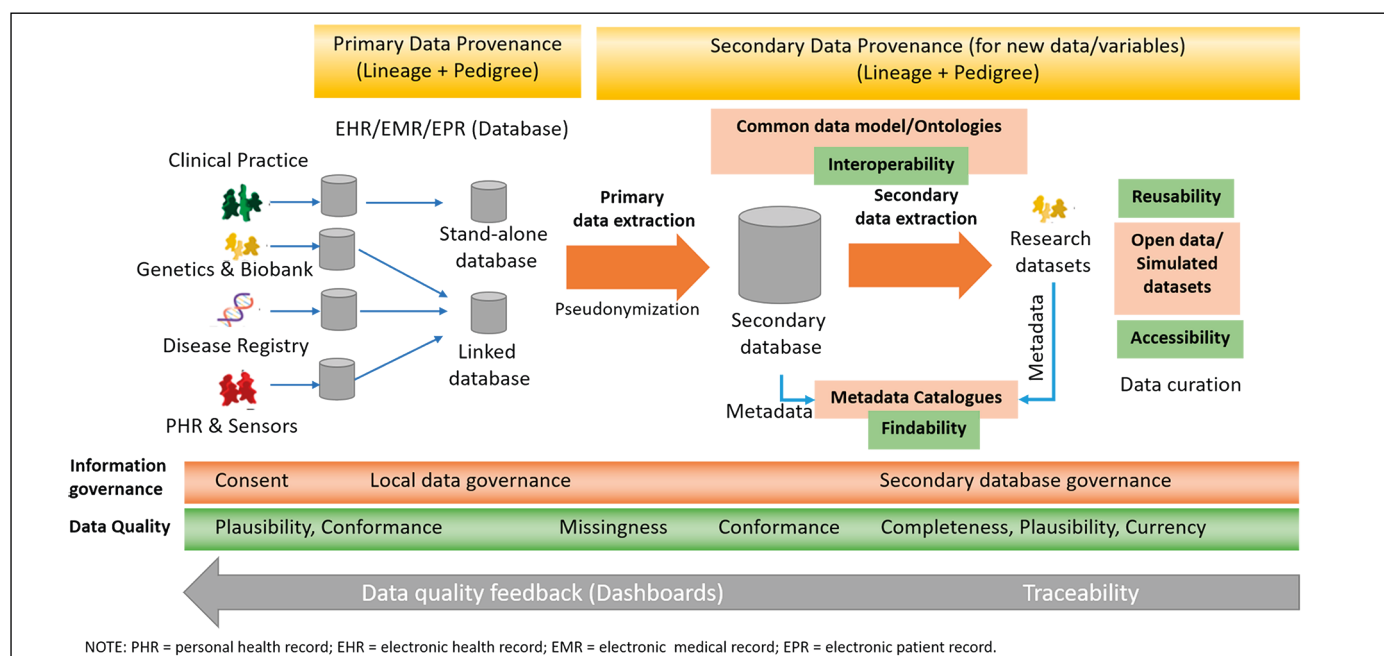


Fig. 1 Data life cycle with FAIR principles, data quality, and governance

services where user agreements state that the owner of the service can use the data collected from the application. Deployment of “apps” and systems that enable unethical and unlawful exchange of digital information undermine trust [24]. Reciprocity, transparency, and mutual trust among actors in the health information exchange network and between data custodians and data providers are essential to ensure willingness to share information. The trustworthiness of data custodians depends on their competence, commitment, and motives.

Patients and service users have high levels of trust in the professionalism of clinical teams and health services [25]. Unnecessary tests and procedures based on unclear data or misinterpretations of unstructured data can compromise patient autonomy, which also interferes with the important task of making well-informed decisions [26].

Legislation in the USA such as HIPAA (*Health Insurance Portability and Accountability Act*) specifically excludes protections once data leave a healthcare provider, a health plan, or clearinghouse (“covered entity”), unless the receiving agent is itself a “covered entity” or has a business associate

agreement. There has been much controversy regarding transmission of these data to large data companies for data mining, mostly focused on the lack of explicit consent from the patients, resulting in control remaining under the responsibility and ethics of system designers and owners. The European Union GDPR (*General Data Protection Regulation (EU) 2016*) gives control to individuals over their personal data.

### Recommendation 3: Pay attention to Extract-Transform-Load (ETL) processes as they may have risks that go unrecognised

ETL from information systems into data repositories must be secure, safe, and accurate [27]. Privacy is important and privacy-preserving linkage techniques to integrate observational data from information systems can be used but they are not always accurate or secure [28]. In addition to the risk of re-identification during the ETL process, there can also be loss of data and compromise of data integrity. For example, a recent study highlighted inaccurate cohort identification where vocabulary mappings of a common data model were

used [29]. The mappings were part of the ETL process and inaccuracies could be due to ETL programming bugs and errors not captured during the quality assurance stages. All risks associated during the ETL process need to be thoroughly identified, assessed, and contingency plans to mitigate these risks should be in place.

### Recommendation 4: Integrate data governance and data quality management to support clinical practice in integrated care delivery and systems

There are ethical issues around information sharing, an essential component of integrated health service delivery, e.g., unauthorised and inappropriate access and use of information. Is patient autonomy more important than access in terms of potential for privacy breaches? Should patients always have full access to their personal health data? Or their children’s? Challenges to data usage to support the delivery of integrated care include: data quality and fitness for purpose; common data models and interoperability; fragmented data governance; proprietary systems and trans-

parency; business model and sustainability of data linkage projects; differing ethical perspectives to processes along the data life cycle; and cognitive load on patients and clinicians. It is often difficult to tease out the quality management and governance of these data challenges. What are the ethical issues that can occur at each point in the data life cycle? Is there systematic bias in the data collected or research questions posed due to poor access to and inequity of care? Are there affordability or commercial biases? A potential solution is equitable AI [30], an equity-focused capacity building strategy by the Canadian Institutes of Health Research [31].

There are good theoretical and ethical reasons for integrating data quality assessment and management (DQAM) with information governance (IG) in an increasingly “big real world data” environment. However, alignment of DQAM and IG across the health enterprise and along the data life cycle [32] is very variable. For ethical reasons, health information ecosystems must process data in ways that are aligned with improving health and system efficiency and ensuring patient safety.

#### **Recommendation 5: Recognise that additional processes are needed to address the ethical issues arising from AI in primary care**

AI and deep machine learning methods have made significant progress, particularly in image processing, e.g., tumour diagnoses and natural language processing to extract complex information from electronic health records. While recommended, scientific review of AI algorithms for reproducibility through sharing of protocols, raw data, and programming codes, introduces risk to patient privacy. It also raises questions about ownership and financial value of large real world datasets and systems [33]. The transformation to AI-enhanced health care needs to be judicious, informed, and systematic to ensure safety and quality of data and care during the transition.

Despite vigorous debate, legislation to regulate the AI industry in the USA has stalled [34] and unlikely to resume soon [35]. The European Union aims to foster

“trustworthy, ethical, and human-centric AI”, emphasising that processes to address high-risk AI systems should be explicit, transparent, traceable, and guarantee human oversight [36]. The requirement for testing and certifying AI systems, including the use of facial recognition, resonates well with the Delphi studies conducted by the IMIA PCIWG, which noted that AI and unsupervised deep machine learning is currently not sufficiently mature, robust, or clinically rigorous to be confidently used without checks in place. The primary care community needs to be proactive and guide the ethical and rigorous development, implementation, and evaluation of AI applications to ensure safety and effectiveness [11].

The tendency for a ‘black-box’ paradigm with AI systems, especially those where intrinsic estimations are biased or not clinically interpretable in biological terms [11], should be strongly discouraged. There is a need for compliance with standards for AI applications and transparency of data processing. In addition, there is a need to share research data, methodologies, and algorithms ethically and securely to ensure reproducibility and generalisability of findings while protecting patient privacy, financial investments, and intellectual property. This sharing can be achieved through the use of videos showing screenshots with outputs based on specific queries or allowing access to “data enclaves” where the tester can drive the AI-based system. Whether this balanced approach is achievable will require a degree of trust within the scientific community that appropriate development and evaluation methods were used [33].

The growing complexity of AI applications has led to the development of algorithms known as *Explainable AI* (XAI) [37]. These are more acceptable to clinicians who need to understand the algorithm and verify the results, especially when things go wrong. Applications based on XAI are also more favourable with medical regulators. Thus, software is lightly regulated where doctors can verify and, preferably, validate the algorithms’ answers to a benchmark. This is an application of the “learned intermediary” principle, where the clinician is central to the decision-making process. However, it

is often difficult to accurately differentiate between biased, incorrect, or inadequately explained AI guidance, and the quality of the clinician’s interpretation and use of the guidance. Collective knowledge from clinicians might be able to avoid these biases and design better XAI to help clinicians critically appraise AI guidance. AI in primary care should “augment (not subvert) the patient-physician relationship” [38].

*Discrimination conscious algorithms* to reduce bias and prejudices in databases and associated applications are also being developed [39]. Discrimination prevention ensures that data mining models do not lead to discriminatory decisions, even if the data set is inherently biased against vulnerable and disadvantaged groups [40].

Empirical studies of existing AI practices that demonstrate critical ethical issues, concepts, and solutions should be done. Trials comparing AI-based or AI-guided practice with usual practice are needed to provide the evidence to support policy recommendations and requirements for current poorly regulated AI practices [41]. We must minimise the opacity and complexity of AI methods, including deep machine learning and neural networks.

Modifications or alternatives to traditional medical research ethics principles are needed to guide the management and governance of AI platforms, methods, and outcomes; informed consent to participate in AI programs; recognising and protecting individual and group level harms and benefits; patient empowerment; patient-doctor relationship; research subject rights to access the inputs and outputs of AI and AI-supported projects; and data protection regulations on research and personal health/wellness services.

Research, Ethics, and Institutional Risk Management Committees must critically appraise and assess the ethical impact of AI applications to evaluate whether an intended action or actual outcome is morally right or wrong. An assessment at all stages in the data life cycle is important to understand “context-mechanisms-impacts”, anticipate and suggest early interventions to avoid or mitigate risks of unethical consequences, reinforce ethical processes and outcomes, and ensure ethical best practices



for AI tools are developed, implemented, maintained, and monitored [10]. Formal processes need to be developed and members of ethics and risk management committees trained to assess the ethical processing of data as it progresses through the data life cycle in AI-related research. Data governance committees should contribute to the oversight of this AI-related research and have processes to monitor data input, processing and outputs from AI implementations for fidelity, bias, safety, and quality.

Cloud technologies and platforms can enable the secure sharing of tools to access and use the data rather than sharing the actual data [42]. However, if poorly-designed or poorly-governed these platforms can compromise the FAIR principles, particularly accessibility and interoperability. Security can be enhanced by advanced encryption technologies. However, quantum computing may provide the computing power to break current encryption schemes [43], suggesting that quantum encrypting schemes are needed to anticipate this potential problem.

### Recommendation 6: Apply an ethical framework mapped to the data life cycle, including an assessment of data quality, to achieve effective data curation

We mapped the 14 ethical principles to the data life cycle to emphasise the importance of an integrated approach for dealing with professional and research ethics, institutional risk, and privacy protection. Table 1 indicates the ethical principles at various stages of the data life cycle with example mechanisms that are or may be used to implement the principles within a primary care setting. This applies particularly to projects using real world data, where privacy, risks, and ethics need to be considered to make sound decisions, including during ethical approval processes. The ethical framework (Figure 1) can guide study designs to ensure that data quality, privacy, and ethics are routinely considered in carrying out health research. Clarity of ownership of data within the context of reproducible research, intellectual property, clinical practice, and liability is important.

## Conclusions

Internationally, there are moves to scale and better integrate primary care into health systems, and the use of technology is part of this process. However, health services are not looking toward the primary care informatics community to ensure that data are integrated and used ethically and confidentially. This paper offers a framework to address this across the whole data lifecycle.

Within integrated care systems, data are often used remotely from recording and may go through all sorts of transformation on their journey. The ETL processes and AI need particularly careful scrutiny. Current systems and service users may not fully appreciate the risks, especially in making trade-offs between quality, safety, efficiency, and privacy.

The deliverable from this exercise is our integrated approach to using data, managing data quality, governance and ethics (Table 1, see next page), and six recommendations, reproduced below, to guide those involved in making greater use of data in primary and integrated care. The recommendations should help maintain trust in contemporary systems and their immediate planned developments.

1. Ensure consent and formal process to govern access and sharing throughout the data life cycle
2. Implement sustainable data creation/ collection to build and maintain trust as well as permission
3. Pay attention to Extract-Transform-Load (ETL) processes as they may have risks that go unrecognised
4. Integrate data governance and data quality management to support clinical practice in integrated care delivery and systems
5. Recognise that additional processes are needed to address the ethical issues arising from AI in primary care
6. Apply an ethical framework mapped to the data life cycle, including an assessment of data quality, to achieve effective data curation

## References

1. World Health Organisation Secretariat. Framework on integrated, people-centred health services. Geneva: World Health Organisation; 2016 15 Apr. Report No.: A69/39.
2. McGovern A, Hinton W, Corra A, Munro N, Whyte M, de Lusignan S. Real-world evidence studies into treatment adherence, thresholds for intervention and disparities in people with T2DM in the UK. *BMJ Open* 2016;6(11): e012801.
3. Coiera E. The Price of Artificial Intelligence. *Yearb Med Inform* 2019;28(01):14-5.
4. Eze B, Kuziemy C, Peyton L. A configurable identity matching algorithm for community care management. *J Ambient Intell Human Comput* 2019.
5. Abouelmehdi K, Beni-Hessane A, Khaloufi H. Big healthcare data: preserving security and privacy. *J Big Data* 2018;5(1):1.
6. Presser L, Hruskova M, Rowbottom H, Kancir J. Care.data and access to UK health records: patient privacy and public trust. *Technology Science* 2015;2015081103.
7. US Department of Health & Human Services. Security of certified EHR technology under meaningful use; 2018. [Internet][cited 2020 Feb 20] Available from: <https://oig.hhs.gov/reports-and-publications/workplan/summary/wp-summary-0000016.asp>
8. Goble C, Stevens R, Hull D, Wolstencroft K, Lopez R. Data curation + process curation = data integration + science. *Brief Bioinform* 2008;9(6):506-17.
9. Liyanage H, de Lusignan S, Liaw ST, Kuziemy C, Mold F, Krause P, et al. Big Data Usage Patterns in the Health Care Domain: A Use Case Driven Approach Applied to the Assessment of Vaccination Benefits and Risks. *Yearb Med Inform* 2014;27-35.
10. Liyanage H, Liaw ST, Di Iorio CT, Kuziemy C, Schreiber R, Terry AL, et al. Building a Privacy, Ethics, and Data Access Framework for Real World Computerised Medical Record System Data: A Delphi Study. Contribution of the Primary Health Care Informatics Working Group. *Yearb Med Inform* 2016(1):138-45.
11. Liyanage H, Liaw ST, Jonnagaddala J, Schreiber R, Kuziemy C, Terry AL, et al. Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges. *Yearb Med Inform* 2019;41-6.
12. de Lusignan S, Liaw ST, Krause P, Curcin V, Tristan Vicente M, Michalakidis G, et al. Key concepts to assess the readiness of data for International research: Data quality, lineage and provenance, extraction and processing errors, traceability, and curation. *Yearb Med Inform* 2011;112-21.
13. Liaw S, Tam M. Ethical research or research ethics? *Aus Fam Physician* 2015;44(7):522-3.
14. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016;3:160018.
15. Cate FH. The Failure of Fair Information Practice Principles. *Consumer Protection in the Age of the Information Economy*; 2006.
16. Di Iorio CT, Carinci F, Brillante M, Azzopardi J, Beck P, Bratina N, et al. Cross-border flow of

**Table 1** Mapping the ethical principles to the data life cycle

	Data collection & utility	Data extraction & storage	Data & metadata processing	Data extraction, linkage, processing & translation	Triangulate and validate data iteratively	Traceability	Curation & use	Mechanisms used to implement ethical principle
1. Autonomy	●	●	●	●	●	●	●	Informed consent
2. Respect rights and dignity of patients	●				●		●	Discrimination algorithms
3. Respect clinical judgment	●				●	●	●	Professionalism
4. Duty to provide care	●						●	Regulations, policies & professionalism
5. Protection of the public from harm	●				●	●	●	Regulations, policies & Professionalism
6. Beneficence	●						●	Professionalism
7. Justice	●						●	Regulations & policies
8. Non-maleficence (obligation to not inflict harm intentionally)	●						●	Professionalism
9. Reciprocity	●						●	Professionalism
10. Solidarity	●			●	●	●	●	Communitarian
11. Stewardship	●	●	●	●	●	●	●	Regulations, policies & professionalism
12. Trust	●	●	●	●	●	●	●	Reproducible research, Explainable AI
13. Lawfulness	●	●	●	●	●	●	●	Data protection legislation
14. Transparent project approval process	●			●			●	Explainable AI

- health information: is 'privacy by design' enough? Privacy performance assessment in EUBIROD. *Eur J Public Health* 2013;23(2):247-53.
17. Liaw ST, Powell-Davies G, Pearce C, Britt H, McGlynn L, Harris MF. Optimising use of EHR data: current issues, evolving opportunities, strategies and scope for collaboration. *Aust Fam Physician* 2015;45(3):153-6.
  18. Liaw ST, Taggart J, Yu H, de Lusignan S, Kuziemy C, Hayen A. Integrating electronic health record information to support integrated care: practical application of ontologies. *J Biomed Inform* 2014;52:364-72.
  19. Tham TY TT, Prueksaritanond S, Isidro JS, Setia S, Wellupillai V. Integrated health care systems in Asia: an urgent necessity. *Clin Interv Aging* 2018;13:2527-38.
  20. Azarm M, Backman C, Kuziemy C. System level patient-centered data sharing. *Proc 2019 IEEE/ACM 1st Int Workshop on Software Engineering for Healthcare, SEH 2019*; 2019:45-8.
  21. McLean R, DeLong D, Moyer D. The U.S. health care system is ill and needs a bold new prescription. *Ann Intern Med* 2020;172(2\_Supplement):S1-S2.
  22. Benchoufi M, Ravaud P. Blockchain technology for improving clinical research quality. *Trials* 2017;18(1):335.
  23. Sunyaev A, Dehling T, Taylor P, Mandl K. Availability and quality of mobile health app privacy policies. *J Am Med Inform Assoc* 2015;22(e1):e28-e33.
  24. Kim D, Ferrin D, Rao H. A trust-based consumer decision-making model in electronic commerce: the role of trust, perceived risks and their antecedents. *Decision Support System* 2008;44(2):544-64.
  25. de Lusignan S, Chan T, Theadom A, Dhoul N. The roles of policy and professionalism in the protection of processed clinical data: A literature review. *Int J Med Inform* 2007;76(4):261-8.
  26. Fiske A, Buyx A, Prainsack B. Health Information Counselors: A New Profession for the Age of Big Data. *Acad Med* 2019;94(1):37-41.
  27. Prasser F, Spengler H, Bild R, Eicher J, Kuhn KA. Privacy-enhancing ETL-processes for biomedical data. *Int J Med Inform* 2019;126:72-81.
  28. Culnane C, Rubinstein B, Teague V. Health data in an open world (arXiv:1712.05627). *Computers and Society*; 2017.
  29. Guo GN, Jonnagaddala J, Farhsid S, Huser V, Reich C, Liaw ST. Comparison of cohort selection performance of Aust Med Terminology to Anatomic Therapeutic Chemical mappings. *J Am Med Inform Assoc* 2019;26(11):37-1246.
  30. Ohno-Machado L. Understanding and mitigating the digital divide in health care. *J Am Med Inform Assoc* 2017;24(5):881.
  31. Lavigne M, Mussa F, Creatore MI, Hoffman SJ, Buckeridge DL. A population health perspective on artificial intelligence. *Healthc Manage Forum* 2019;32(4):173-7.
  32. Liaw S, Pearce C, Liyanage H, Cheah-Liaw G, de Lusignan S. An integrated organisation-wide data quality management and information governance framework: theoretical underpinnings. *Inform Prim Care* 2014;21(4):1-8.
  33. Carter RE, Attia ZI, Lopez-Jimenez F, Friedman PA. Pragmatic considerations for fostering reproducible research in artificial intelligence. *NPJ Digit Med* 2019;2:42.
  34. US Food & Drug Administration. [Internet] Artificial intelligence and machine learning in software as a medical device. [cited 2020 Feb 20] Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-software-medical-device>. 2020 Jan 28.
  35. Zhang B. Public opinion lessons for AI regulation. [cited 2020 Feb 20]. Brookings; 2019 Dec 10. Available from: <https://www.brookings.edu/research/public-opinion-lessons-for-ai-regulation/>
  36. European Commission. White Paper on artificial intelligence: A European approach to excellence and trust. Brussels; 2020 Feb 19.
  37. Birkeneder E. Explainable AI In Health Care: Gaining Context Behind A Diagnosis: *Forbes*; 2019. Available from: <https://www.forbes.com/sites/erikbirkender/2019/10/16/explainable-ai-in-health-care-gaining-context-behind-a-diagnosis/#46e2086c4bee>. [accessed 2020 Feb 20]
  38. Lin SY, Mahoney MR, Sinsky CA. Ten Ways Artificial Intelligence Will Transform Primary Care. *J Gen Intern Med* 2019;34(8):1626-30.
  39. Hajian S, Bonchi F, Castillo C. Algorithmic bias: From discrimination discovery to fairness-aware data mining. *Proceedings of the 22nd ACM SIGKDD international Conference on Knowledge Discovery and Data Mining*; 2016. p. 2125-6.
  40. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019;366(6464):447.
  41. Magrabi F, Ammenwerth E, McNair JB, De Keizer NF, Hyppönen H, Nykänen P, et al. Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications. *Yearb Med Inform* 2019:128-34.
  42. Eze B, Kuziemy C, Peyton L. Cloud-based performance management of community care services. *Journal of Software: Evolution and Process* 2018;30(7):e1897.
  43. Kobie N. The quantum clock is ticking on encryption – and your data is under threat. *Wired Security* [Internet] 2016 . [cited 2020 Feb 20]. Available from: <https://www.wired.co.uk/article/quantum-computers-quantum-security-encryption>

Correspondence to:  
Siaw-Teng Liaw  
E-mail: [siaw@unsw.edu.au](mailto:siaw@unsw.edu.au)