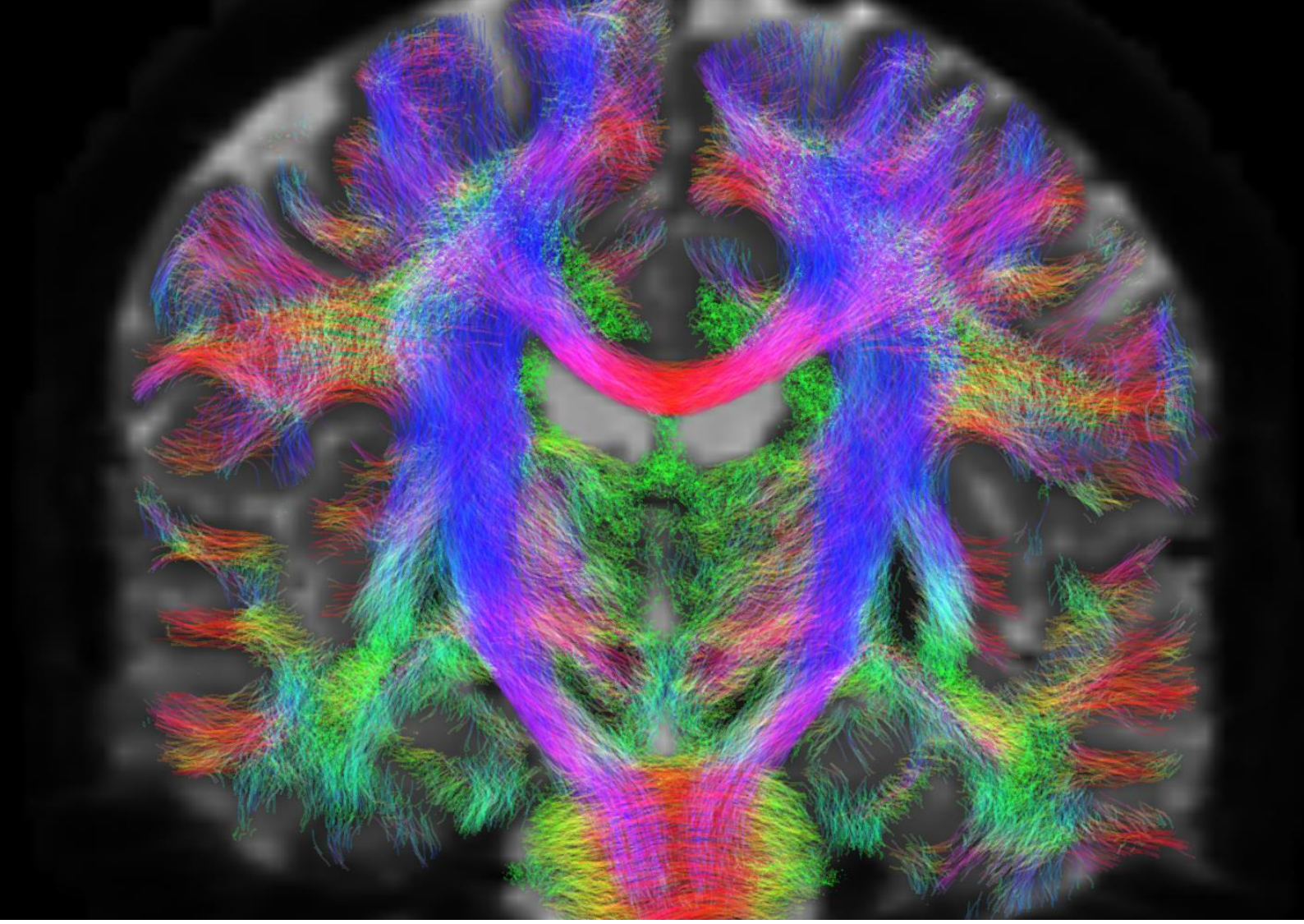


Exemplars of Artificial Intelligence and Machine Learning in Healthcare

Improving the safety, quality, efficiency and accessibility of Australia's healthcare system

July 2020

THE AUSTRALIAN E-HEALTH RESEARCH CENTRE

Queensland Government

CSIRO

Citation

Koopman, B., Bradford, D., Hansen, D. (Eds) (2020) Exemplars of Artificial Intelligence and Machine Learning in Healthcare: Improving the safety, quality, efficiency and accessibility of Australia's healthcare system. Report Ep203543. CSIRO, Australia.

Version 1.0 dated July 2020 is available at aehrc.com/ai

Case Study authors

Denis Bauer, Justin Boyle, Jason Dowling, Jurgen Fripp, Shaun Frost, Norm Good, John Grimes, David Ireland, Mohan Karunanithi, Sankalp Khanna, Michael Lawley, Kay Mann, Alejandro Metke Jimenez, Anthony Nguyen, Alex Pagnozzi, Parnesh Raniga, Christian Redd, David Silvera, Jim Steel, Susmita Saha, Natalie Twine, Janardhan Vignarajan, Laurence Wilson, Qing Zhang.

Cover image

Whole-brain tractogram obtained from diffusion MRI, with colours showing local orientation of white matter pathways. AEHRC uses AI to assess brain connectivity in brain development, ageing, and disease. Image: Kerstin Pannek.

Acknowledgements

At AEHRC, we acknowledge the Traditional Owners and custodians of the lands on which our offices are located and where we conduct our research. We pay our respects to Elders past, present and emerging.

Special thanks to our participants, collaborators and stakeholders.

Copyright

Important disclaimer

Foreword

This report provides a view of how Artificial Intelligence (AI) and Machine Learning (ML) are currently used in healthcare through a series of Case Studies from research undertaken at the Australian e-Health Research Centre at CSIRO. Our focus is to develop technologies to improve the safety, quality, efficiency and accessibility of healthcare in Australia and overcome the major challenges faced by Australia's healthcare system – ageing population, chronic disease, workforce pressures, access to services and inequities in health outcomes.

We wrote this report for clinicians, health service executives, academics and others with an interest in using data to improve today's health service, and in understanding how AI and ML supports the transition of the healthcare system to a digitally enabled healthcare system.

Our Case Studies demonstrate that AI and ML are part of how we undertake our research and development and are central to the digital health tools we develop – but the AI and ML do not drive our research. Rather, it is the aim to achieve an improved health outcome that is the driver, and AI and ML form part of the suite of tools we employ to achieve that aim. The Case Studies describe research and development in actual use in healthcare, demonstrating current practice in AI and ML. These Case Studies provide promise of future possibilities.

We included 34 Case Studies from across our medical imaging, genomics, data analytics, health services and health data interoperability research that demonstrate the myriad of forms in which AI features. They also demonstrate that integration into the health system is integral to success – understanding the actual problem and the implementation issues.

The Case Studies are not designed to be read in any particular order, rather the reader should move between Case Studies at will, guided by their interests and knowledge. The language is aimed at a general, albeit well educated, audience, although we recognise that specialist interest and knowledge are probably required on occasion (Case Study 16: Post-coordination and the SNOMED CT Expression Constraint Language, for example!).

We have also included a primer on the different types of AI and ML and how these are employed with health data to unravel the challenges posed in the healthcare system. Much of the discussion around AI at the moment tends to focus on the ML methods. This is due to the huge amounts of data available for the equally large amounts of computing power – with the result that ML algorithms can identify trends in data well beyond human computing ability. However, there are many other techniques that make up the family of AI and ML approaches, and it is important to understand the full scope of what is available. While the primer given in this report is by no means a complete compendium of all AI and ML methods, we hope that it provides readers with the breadth of current and potential opportunities and pointers for further reading.

We finish this report with eyes to the horizon, and the promise of AI and ML to improve health system efficiency – enabling targeted, tailored and timely interventions for all patients.

Dr David Hansen

CEO, Australian e-Health Research Centre

Contents

List of Tables and Figures

Acronyms and Abbreviations

ABI	Acquired Brain Injury
ADL	Activities of Daily Living
AEHRC	Australian e-Health Research Centre
AI	Artificial Intelligence
ALS	Amyotrophic Lateral Sclerosis
AMD	Age-related Macular Degeneration
AMT	Australian Medicines Terminology
ASD	Autism Spectrum Disorder
AWS	Amazon Web Services
CapAIBL	Computational Analysis of PET for the Australian Imaging, Biomarker & Lifestyle Study of Ageing
CNN	Convolutional Neural Network
COVID-19	Corona Virus Disease identified in 2019
CP	Cerebral Palsy
CPU	Central Processing Unit
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CT	Computed Tomography
CT-MR	Computed Tomography merged with Magnetic Resonance imaging
DICOM	Digital Imaging and Communications in Medicine
DL	Description Logic
ECL	Expression Constraint Language
ED	Emergency Department
EMR	Electronic Medical Record
FAF	Fundus AutoFluorescence
FHIR	Fast Healthcare Interoperability Resources
GPUs	Graphics Processing Unit
HL7	Health Level 7 - The International Standards Body for Pathology
ICD	International Classification for Disease
ICU	Intensive Care Unit
MICE	Medical Imaging Communication Exchange
ML	Machine Learning
MLlib	Machine Learning Library
MRI	Magnetic Resonance Imaging
NLP	Natural Language Processing
OCT	Optical Coherence Tomography
OWL	Web Ontology Language
PAPT	Patient Admission Prediction Tool
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PHEIC	Public Health Emergency of International Concern
QIMR	Queensland Institute for Medical Research
ResNet	Residual Neural Network
RNN	Recurrent Neural Network
SNOMED CT	Systematized Nomenclature of Medicine -- Clinical Terms
SPECT	Single-Photon Emission Computed Tomography
UWB	Ultrawide band
VEP	Variant Effect Predictor

1 AEHRC Research and Artificial Intelligence

The use of Artificial Intelligence (AI) in healthcare is not new – AI technologies have been used in healthcare for decades. However, the increasing capture of data electronically in Electronic Medical Records (EMRs), the increase in personal data captured through devices, sensors, imaging or genomics and the increase in computing power available – either through cloud-based computing platforms or on the phones in our pockets – is enabling a new generation of applications of AI through-out the healthcare system.

The Australian e-Health Research Centre (AEHRC) is the leading national digital health research facility applying information and communication technology to improve health service delivery for people living in Australia. Most of the technologies developed and research projects undertaken by the AEHRC's almost 100 data scientists employ some sort of AI and Machine Learning (ML) techniques in developing digital health solutions (Figure 1).

These AI techniques are powering digital health – from understanding whole genome sequences to enabling clinical decision support from EMRs, to determining risk prediction for patients, to understanding functional wellbeing of older people living independently.

This paper explores some of the key AI and ML technologies used by the AEHRC in developing digital health solutions. The paper providers a series of case studies from the AEHRC's five research groups to demonstrate the value that AI and ML brings to healthcare.

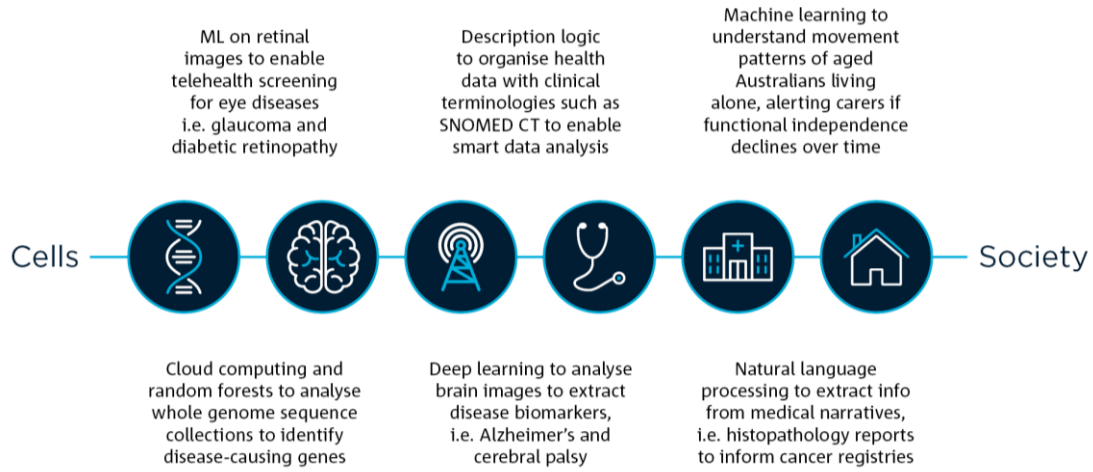


Figure 1: From genomic engineering to independent living, the Australian e-Health Research Centre (AEHRC) is using artificial intelligence techniques and machine learning approaches to overcome the challenges facing the healthcare system and improve health service delivery to people living in Australia.

2 A Primer for AI and ML in Health

AI is a somewhat amorphous term. There are many different types of techniques and algorithms that make up the wide family of AI techniques and the boundary can be somewhat subjective. In this Chapter, we provide an overview of the major trends and techniques that relate specifically to AI in Health. AI techniques are often described across a range of different components or scientific domains. For the purpose of this paper we focus on the four different domains highlighted in Figure 2.

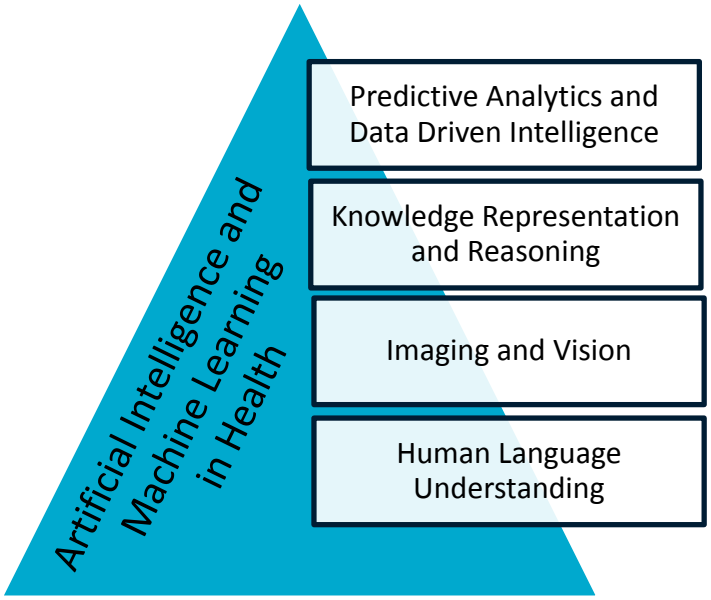


Figure 2: Domains of Artificial Intelligence techniques.

Predictive Analytics and Data-Driven Intelligence is concerned with extracting insights from existing data – often from large datasets where it is difficult for humans to derive such insights. In data-driven intelligence, the intent is for insights to be bottom-up; i.e., identify trends and insights from the, often, low-level data.

Knowledge Representation and Reasoning[1] is about how we represent or classify information about the world in a form that a computer system can utilize to solve complex tasks enabling us to infer (new) knowledge. In healthcare, this is typically about representing medical concepts (such as diseases) and their properties and relationships in a machine readable and understandable form. In many instances, solving the knowledge representation problem is the key challenge: once the data is represented in the right form, the problems become "tractable" – able to be processed using compute power in an appropriate timescale.

[1] Hommersom A., Lucas P.J.F. (2015) An Introduction to Knowledge Representation and Reasoning in Healthcare. In: Hommersom A., Lucas P. (eds) Foundations of Biomedical Knowledge Representation. Lecture Notes in Computer Science: 9521. Springer, Cham.

Imaging and Vision[2] involves analysing images or videos to derive insight into the cause and impact of medical conditions. Computer vision and image processing are two areas that have been transformed by new AI methods, particularly deep learning-based methods.

Human Language Understanding[3] uses AI methods aimed at understanding natural language. While striving to standardise data and make it machine readable, humans communicate in natural language and as such there will always be data in this form. AI methods, therefore, aim to handle human language by extracting meaning, searching, summarising and classifying such language.

2.1 Symbolic or Statistical Artificial Intelligence?

In AI, there have traditionally been two schools with contrasting approaches – symbolic and statistical. Symbolic AI concerns itself with attempting to explicitly represent human knowledge in a declarative form (i.e. facts and rules). Statistical (or non-symbolic) AI, instead, involves presenting the data in a more-or-less raw form to a machine and having it learn patterns and create its own representation of the knowledge.

There are a number of different AI technologies that we describe below. These different technologies use one of two key AI paradigms to solve problems or perform tasks.

Symbolic AI

Symbolic AI refers to the representation or encoding of human knowledge into a form of known facts and/or rules. These facts or rules can then be combined with mathematical logic to undertake verifiable and explainable problem solving.

This paradigm of AI often uses an ontology – a collection of concepts with properties including the relationships between the concepts – to describe a part of the world. Mathematical logic or rule-based systems can then use the concepts, relationships and properties of the ontology to 'reason' across a data set.

The 'semantic web' is a prime example of this symbolic approach to AI, with the Web Ontology Language (OWL) as the basis for describing data across the semantic web.

While symbolic AI and explicit encoding of knowledge has become less popular in the general domain, it has a specific place in AI in Health. This is mainly because in the health domain considerable effort has been made to capture and explicitly represent health data in various standards such as the SNOMED CT ontology. Thus, symbolic AI methods can make use of this already curated medical domain knowledge.

[2] Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing*. Pearson.

[3] Deng, L., & In Liu, Y. (2018). *Deep learning in natural language processing*. Springer.

Statistical AI

Statistical AI takes the opposite approach – rather than predefining the knowledge and rules, it 'learns' these from the data itself. This approach uses existing data and evidence along with computational techniques to extract patterns and insights and thus reason about the world. This process involves training a model using available data. In healthcare, this enables us to identify relationships in large sets of data – such as estimating the complex biological relationships that underpin the world's most complicated diseases or identifying the factors that increase the risk of certain outcomes for patients.

ML and its family of algorithms are the key techniques used in Statistical AI. These new approaches are required due to the large increases in the availability of data captured electronically and the computing power of high-performance supercomputer clusters. ML algorithms have been able to attain a level of precision that was not possible in the past with smaller data sets and less computing power.

While ML is one of the better-known sub-disciplines of AI, its success depends on being combined with many other sub-disciplines of AI to deliver solutions for real world problems.

Both paradigms play an equally important role, especially for AI solutions developed for domains such as healthcare.

2.2 Data, Data Models, Data Quality and Data Standards

AI depends on data, and the quality of the data used to either train AI models or for AI based analysis will have a direct impact on the quality of outputs and downstream tasks. Health data comes in a myriad of different forms. As many AI approaches are intrinsically linked to the data type, we outline some of the more common types of health data (Table 1).

2.3 Machine Learning

ML is a field of study that gives computers the ability to learn without being explicitly programmed. There are three main techniques for ML: statistical ML aims to find some type of predictive function from the training data; reinforcement learning approaches provide AI algorithms with 'rewards' or 'penalties' based on their problem-solving performance; and deep learning approaches make use of artificial neural networks.

Traditional ML includes techniques such as regression, decision trees, support vector machine, Bayes network, K-nearest neighbour, principal component analysis. These techniques analyse various types of data – numerical, categorical, binary, time series, text, image, audio and video data – to identify relationships in the data.

Classification vs regression

There are two main ML tasks: classification and regression.

Classification involves using a ML model to 'classify' some data according to a finite set of categories; for example, classifying the type of cancer found in a pathology report: breast, lung, etc. The simplest case being a binary classification – yes/no, true/false, cancer/not cancer, etc.

Regression, in contrast, involves using a ML model to predict a continuous value rather than a category. For example, predicting length of stay for a patient given their condition.

Most ML models perform either regression or classification; however, there are models that can be implemented to cover both.

Table 1: Categories of health data

Data Type	Description	Format and Standards	Avg. size/patient
Clinical Data	Includes a wide range of data used by health organisations. This includes patient records and is what large electronic health record systems would store. It also includes laboratory data (e.g. pathology and radiology reports).	Includes data which is unstructured (free-text narratives) or highly structured (e.g. Intensive Care Unit (ICU) observations). There is a strong drive to develop standards for clinical data. • HL7 for messaging/transfer of clinical data; • FHIR to replace many older HL7 for better patient data interchange; • SNOMED CT - an ontology for medical knowledge representation; • International Classification of Diseases (ICD) - a classification for diagnostic coding. • Human Phenotype Ontology for describing features associated with genetic disease	100s MB
Genomics Data	Growing source of data as precision medicine becomes more commonplace. Includes single gene tests, panel tests, whole exome sequencing, whole genome sequencing. With precision medicine, the boundary between clinical and genomic data will narrow.	Genomic data can be provided in multiple forms of processing stages each with its own set of standards or uniquely structured data, ranging from raw sequence data such as FASTA (consensus genome sequence) and FASTQ (sequencing reads), processed data such as SAM/BAM (aligned reads), to tertiary derivatives such as VCF (genomic variant information) or unique structures for annotated VCFs, methylation, transcription etc. This is more dominant in the non-human domain such as metagenomics or pathogenomics.	138 GB (whole genome, FASTQ); 10s-100s GB/cohort (VCF)
Imaging	Medical images typically arising from radiology. Imaging results from MRI, CT, x-ray, PET etc.	Standards such as DICOM exist to capture medical images; however, these are still pictures (or 3D images) represented with pixels.	Modality dependent, 100s MB–GB
Administrative data	Data not associated directly with a patient's medical condition. Includes billing, insurance, financial data, and efficiency metrics and patient flow numbers.	A mix of structured and unstructured. Use of some terminology systems such as ICD-10.	100s MB
Sensor and wearables Data	Data provided by sensors. Sensors vary widely from ICU observations through to wearables in the home. Rapidly growing areas with the Internet of Things (IoT).	Often numeric, structured and time-series based. Can be image, sound, number or text.	Varies greatly

Supervised vs unsupervised

ML models learn from data, in either a supervised or unsupervised manner.

Supervised learning means that the machine is provided with the correct answers or labels along with the data. For example, the machine is provided with lung x-rays labelled with either "cancerous" or "not cancerous" and then learns characteristics of the image that indicate these classifications. The key feature of supervised learning is that the answers/labels are needed along with the data. For some problems these labels are easily available; for other problems obtaining labels can be a costly exercise.

Unsupervised learning means the learning is not directed by any manually provided answers or labels. An example would be clustering algorithms that might group similar medical images together.

Deep learning

Deep learning is an approach that uses artificial neural networks for either classification or regression, both supervised and unsupervised. The name deep learning comes from the fact that their architecture has many "deep" layers of neural networks. These are able to capture richer and more high-level features in the data they are provided; for example, deep learning for facial recognition will capture both coarse grained location of the eye within the face as well as more fine-grained parts of the eye.

Deep learning has shown impressive results in variety of domains, but in particular in image processing and speech processing.

2.4 Predictive Analytics and Data Driven Intelligence

Predictive analytics and data driven intelligence refer to the analysis of data and the discovery of patterns that provide novel insights to inform workflow and decision making in the chosen business domain. These are often used in the context of big data and employ AI and ML techniques to extract meaning and knowledge from this data to discover relationships and trends, forecast more accurately and reliably, guide informed decision-making, and optimise business operations.

Predictive analytics and data-driven intelligence might use health data to help improve capacity management and patient flow through the health system, develop patient-centred evidence-based models of care, address the burden of chronic disease, health monitoring and management of home-based care. Equally this area of AI and ML might identify the complex biological relationships that underpin the worlds most complicated diseases.

While in knowledge representation, information about the world is explicitly represented using symbols, here the representations are learned from the actual data. For example, from accelerometer sensor data, learning what a fall looks like from previous falls rather than explicitly defining the characteristics of a fall.

2.5 Knowledge Representation and Reasoning

Knowledge representation and reasoning is a core tenant of AI. An intelligent system needs some means to represent information and knowledge in a form computers can understand to solve complex problems such as diagnosing a medical condition or understanding natural language. Once knowledge is represented in a suitable form, reasoning systems can then apply this knowledge in new situations, to acquire or infer new knowledge. In symbolic AI, knowledge is represented as symbols (often pre-defined human concepts) and reasoning is done by manipulating symbols in an automated manner.

Formal logic is a form of symbolic reasoning that represents knowledge in declarative form as a set of facts and rules. Description logic (DL) is a family of knowledge representation languages with (usually) decidable semantics and efficient reasoning algorithms. They are often used to describe and reason about relevant concepts in an application domain and are widely employed to represent clinical concepts. The foundation of the reasoning process is known as classification.

From a knowledge representation perspective, health is quite unique. This is because considerable effort has been made to capture and represent medical knowledge in a standardised and machine-reasonable manner. Today, some of the largest domain-specific knowledge representation systems are health based.

While there are a number of terminologies for representing medical knowledge, one that adheres to the formal logic mentioned above is the SNOMED CT ontology. In SNOMED CT, knowledge representation is achieved through three core components:

- Concepts: a numerical code that identifies a specific medical concept; e.g. 74400008 is the code for appendicitis.
- Descriptions: textual descriptions for a concept, for which there may be multiple; e.g. the descriptions 'heart attack' and 'myocardial infarction' both pertain to concept 22298006.
- Relationships: which connect concepts together; e.g. 'appendectomy' can actually be represented as a relationship between the 'appendix' and 'excision' concepts.

A sample portion of the SNOMED hierarchy as visualised by the CSIRO Shrimp browser for cardiomegaly is provided in Figure 3.

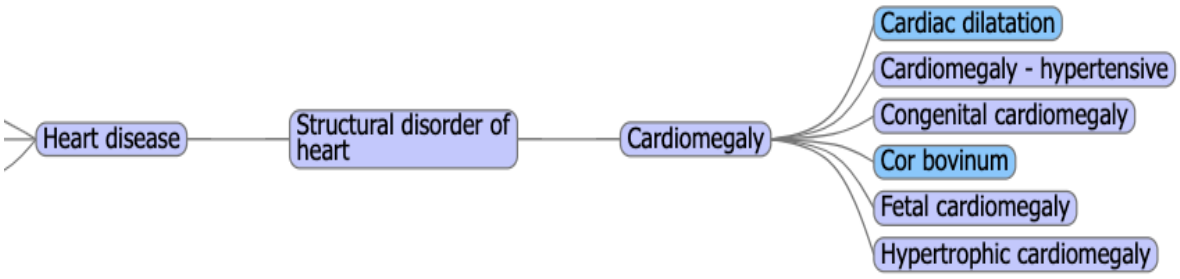


Figure 3: A portion of the SNOMED CT hierarchy as visualised by the CSIRO Shrimp browser. The relationships are vital in allowing specific clinical information about a patient to be captured, e.g. "Cardiomegaly – hypertensive" while allowing the clinical information system to still be able to find all patients with "Cardiomegaly".

The AEHRC has a long history of active development in SNOMED CT and various tools that support the use of SNOMED CT. This includes a reasoning engine that is able to use SNOMED CT to infer new knowledge, a terminology service that allows third parties to easily look up, retrieve and leverage SNOMED CT concepts, and support for analytics and genomics that utilise SNOMED CT. A number of Case Studies in this vein are outlined in Chapter 5.

In addition to the representation of knowledge, there needs to be an *information model* in which to store data, with the data coming from the terminology. The easiest way to think of an information model is of a simple form for collecting data. In this case the fields in the form provide the information model while the data which is collected in each field is from the terminology. For instance, there might be a field in the form to capture diagnosis and the data captured in the form should come from the "clinical finding" hierarchy of the SNOMED CT terminology. In healthcare, HL7 (Health Level 7) is the international standards body for standards supporting the transfer of clinical and administrative health data. Their standards, including HL7 v2, CDA (Clinical Document Architecture) and FHIR (Fast Healthcare Interoperable Resources) are the bedrock of the sharing of information in healthcare. The information model – whether it is a HL7 standard or one of many others – provides the "meta-data" for any data used in AI or ML applications.

2.6 Human Language Understanding

Humans communicate in natural language, which is ambiguous, does not follow strict syntactic or semantic rules and is thus difficult for machines to understand. In response, there has been a strong emphasis to avoid using natural language and adopt standard terminologies (e.g. SNOMED CT and ICD). However, for certain tasks – such as communication between humans – natural language remains the most effective format. In response, computational methods for natural language processing (NLP) and, more generally, natural language understanding have emerged.

With the advent of new deep learning models for NLP, considerable improvements have been made in the field. A key technique in the health space is in information extraction – that is analysing and then extracting structured information from unstructured text. Information extraction can be used to identify specific disorders, procedures or test results. This can then be used for a whole host of downstream tasks from clinical decision support through to billing and administration. In many cases, the task involves mapping the text to a standard medical terminology; e.g. mapping a text to SNOMED CT codes. In this way, NLP is used to bootstrap the knowledge representation and reasoning resources outlined in Section 2.5. Alternatively, new deep learning methods are able to learn representations of word meaning independent of standard terminologies, akin to how facial recognition systems learn to recognise images of people without being told how to explicitly.

Beyond information extraction, AI has been used to identify the relationships between medical concepts extracted from free text. Relationship extraction plays an important part in extracting meaning from free text. For example, it can identify the relationship between drugs and adverse events or conditions and risk factors. New deep learning models, again, have shown promising empirical results in this domain.

NLP can be used for document categorisation; for example, from a patient record, recording risk or triage categories. New NLP methods also involve generating succinct summaries of long documents (e.g. generating patient summaries). Natural language understanding also encompasses new techniques in search engine technology, where AI models now enable 'semantic search' – search based on the underlying meaning of query keywords. New search methods are ML based: they take a set of existing queries and associated relevant documents to 'learn' how to rank results according to an unseen query.

The advent of conversational agents (e.g. Siri, Alexa, etc.) and custom chatbots have meant that natural language interfaces and data are on the rise. This in turn requires models that successfully interpret, reason and respond in natural language. This natural language understanding is a fast-growing area – new techniques, availability of vast amounts of training data and the exponential growth in computing power (including specialised processors called GPUs) have also all contributed to significant growth. It is likely this area of AI will play an increasing role in healthcare.

2.7 Imaging and Vision

Imaging and Vision are key applications of AI – used to extract information from images and in order to make decisions. Imaging applications can range from using images taken by regular cameras through to images acquired by advanced imaging machines typically used in healthcare. In the case of vision applications, images are processed in real time and used for applications in Robotics, which are increasingly deployed in medical applications from surgical robots to social robotics.

Medical imaging (along with genetics) has become one of the key elements in precision medicine in advanced healthcare systems. Since the first use of X-rays in 1896, the fields of Radiology and Nuclear Medicine have helped revolutionise the diagnosis and treatment of a whole range of health conditions. There is currently a long list of available modalities including X-ray radiography, magnetic resonance imaging (MRI), ultrasound, endoscopy, elastography, tactile imaging, thermography, medical photography, and nuclear medicine functional imaging techniques such as positron emission tomography (PET) and single-photon emission computed tomography (SPECT).

In healthcare, outcomes are dependent on the acquisition technology, interpretation and communication of the medical images. Traditionally, medical imaging is usually interpreted qualitatively by trained experts. One of the foci of our centre is to develop technologies that extract and analyse quantitative imaging biomarkers for use in screening, risk stratification, diagnosis and treatment for various clinical and research applications. The developed technology turns images into information that is used for earlier detection of diseases and improved diagnostic accuracy.

A large component of this work has involved the use of AI and in particular ML and deep learning approaches that provide a powerful approach to perform or improve image-based tasks such as image acquisition, reconstruction, quantification (segmentation) and analysis. This enables interpretation roles including clinical scoring (of pathology), diagnosis and prognosis.

Not all imaging requires expensive medical imaging machines. Other applications of AI use images taken with regular cameras – for example when examining skin lesions or burns – or with cameras which take photos of parts of the eye, for retinal image analysis. Techniques include automated methods for registration of retinal images that are collected over time (longitudinally), or obtained using different retinal imaging devices/modalities, or are captured from different angles.

3 AI, ML and AEHRC Research Groups

AEHRC's 100 scientists and engineers and over 30 research students work across five research groups. The groups have each developed a number of platform technologies and digital health solutions that enable them to work together and with our stakeholders and collaborators to tackle the key challenges of twenty-first century healthcare.

3.1 The Health System Analytics Group

The Health Systems Analytics group develops AI and ML based tools to help increase health system productivity, improve patient safety, and deliver higher quality care. The tools are aimed at optimising patient, clinician and resource level flows, and providing intelligent decision support.

The tools developed employ a mix of AI and ML technologies integrated with allied science areas such as simulation modelling to predict and optimise hospital workflows and support clinical and administrative decision making. Working closely with clinicians and health system administrators, the group has delivered significant impact in the areas of patient flow analytics and hospital avoidance.

Case Studies from the Health System Analytics group are provided in Chapter 5 AEHRC Case Studies in Predictive Analytics and Data Driven Intelligence.

3.2 The Health Informatics Group

The Health Informatics research group aims to improve patient outcomes and health system performance and productivity through the use of data collected in EMRs and other clinical information systems. The group uses international data standards to enable data analytics and text processing to support clinical decision making in electronic health systems. The group applies ML, NLP, formal logic, and statistical and simulation approaches to problems involving decision support, systems modelling and reporting.

The group works closely with two key international standards, SNOMED CT – the international standard clinical terminology – and FHIR. As well as using these standards in many applications across analytics and text processing, the group is recognised for contributing to the development and adoption of both SNOMED CT and FHIR.

Our solutions have been developed in partnership with practitioners from across healthcare, including cancer registries, hospital radiology and emergency medicine departments and international standards bodies. They leverage the wealth of clinical data from health industry stakeholders to aid in decision support and reporting.

Case Studies from the Health Informatics group are provided in Chapter 6 AEHRC Case Studies in Knowledge Representation and Reasoning and Chapter 7 AEHRC Case Studies in Human Language Understanding.

3.3 The Transformational Bioinformatics Group

The Transformational Bioinformatics group are world leaders in cloud-native genomics research, using the latest in ML and cloud computing technology to drive innovation in the use of genomics in the health system.

The group applies AI and ML across two main genomics disciplines – Genome Sequence Analysis and Digital Genome Engineering.

The Genome Sequence Analysis research is aimed at understanding the functionality of the genome and how it relates to disease. Our team develops ML based tools to evaluate large genomic datasets for population genomics health applications, such as novel biomarker discovery.

The Digital Genome Engineering research aims to develop technology to support applications which edit a genome for biosecurity and medical applications. Our team works closely with our collaboration partners to develop bespoke ML algorithms to enable Gene Therapies or design Gene Drive applications.

Case Studies from the Transformational Bioinformatics group are provided in Chapter 5 AEHRC Case Studies in Predictive Analytics and Data Driven Intelligence.

3.4 The Biomedical Informatics Group

The Biomedical Informatics group develops innovative medical technologies for the discovery and communication of meaningful patterns in biomedical data. Especially valuable in areas rich with recorded information such as medical images or genomics, these technologies rely on the simultaneous application of statistics, computer programming, and applied mathematics. The group also develops techniques to report and visualize complex biomedical information for clinical diagnosis and screening and to communicate insights to clinicians and clinical researchers.

The biomedical informatics group works with clinicians, industry, and patients to develop, validate and translate imaging-based AI and ML tools for improved disease diagnosis, treatment planning and treatment delivery, blood-based diagnostics and innovative new therapies. This includes areas such as neuroimaging and neuroplasticity, soft tissue imaging (e.g. prostate cancer) and imaging for musculoskeletal conditions, including following knee surgery.

AEHRC medical image analysis research is leading a paradigm shift in radiology from qualitative to (semi-) quantitative imaging and the development of a new generation of imaging biomarkers. The developed technology turns images into information that is used for earlier detection of diseases and improved diagnostic accuracy. This provides new insights and will reduce healthcare costs.

In addition to imaging data, the group works with multiple types of other biomedical data, including physiological and blood derived, and applying benchmark statistical methodologies to elucidate the complex relationships between biomarkers and disease pathology. Working with many international collaborators, group members utilise AI and ML technologies to discover new biomarkers from many thousands of candidates. Software is designed to automate workflows and deliver reports to customers around the world.

Areas include:

- Outcome prediction
  - to improve detection and diagnosis leading to a better understanding of prognosis for neurodevelopmental disorders
  - AI combined with MRI to predict joint pain, disorders, or abnormal movement patterns.
- Clinical decision support tools
  - using ML to understand neuropathology and relation to an age-matched typically developing cohort for children with neurodevelopmental disorders
  - to provide clinician friendly reports for aiding the quantification of PET and MR images used for the early diagnosis of Alzheimer's disease
  - generation of synthetic CT and organ delineation from MRI for image-guided radiotherapy
- Advanced neuroimaging
  - to measure localisation and extent of neuroplasticity in response to evidence-based interventions for developmental disorders and brain trauma
  - to develop novel analysis techniques that leverage deep learning technologies for faster and more accurate quantification of MR brain images.

Case Studies from the Biomedical Informatics group are provided in Chapter 8 AEHRC Case Studies in Imaging and Vision.

3.5 The Health Services Group

The Health Services group develops diagnostic and decision support systems for delivery of virtual, remote, community, and aged healthcare. The multi-disciplinary group brings together expertise in clinical research, tele-medicine systems and AI for medical image and data analysis. We work with key stakeholders and collaborators to develop and trial these solutions to demonstrate improved health outcomes and health service delivery.

The group's research uses three key paradigms in the delivery of health services: sensors, mobile health, and store and forward tele-health. In all these paradigms, the team develops image and data analytics which use AI and ML. The goal is to improve access to services, health outcomes and quality of life while realising cost savings for the healthcare system.

Chronic disease rates are increasing in Australia and around the world, stretching existing health resources and increasing the cost and complexity of care. One strategy for meeting this additional demand has been to shift care to the home, using mobile health technologies to more holistically manage complex chronic conditions.

AI can be used along with mobile health technologies to enable timely and precise interventions, shifting the model of care from reactive to proactive. Mobile health data is integrated into deep learning models to predict when a patient may be at risk of experiencing a complication, and then can provide an alert to the care provider to intervene or increase support.

Predictive AI models can incorporate mobile health data from all aspects of the patient's life, allowing precise management not available in the traditional clinic.

Case Studies from the Health Services group are provided in Chapter 5 AEHRC Case Studies in Predictive Analytics and Data Driven Intelligence and Chapter 8 AEHRC Case Studies in Imaging and Vision.

4 Real World Applications

The National Primary Health Care Strategy[4] identified four main challenges faced by Australia's healthcare system – the growing burden of chronic disease, an ageing population, workforce pressures, and unacceptable inequities in health outcomes and access to services. In this Chapter, we take an outward looking approach and cover four key areas within the health system where existing or emerging AI is being used to meet these challenges.

4.1 Data, Big Data and Health Records

The modern health system is increasingly data rich with electronic mediums now replacing paper-based systems. This transition has happened somewhat independently and at times organically – with some areas of the health system transitioning earlier (such as laboratory results), with other areas being more recent converts (such as full electronic health records). A consequence is that while the system may be data rich, areas within it can be siloed, disconnected and lacking shared data standards.

Recognising the need for improved management and sharing of data, considerable effort has been directed at standards and infrastructure to better represent health data. In addition, with so much data available, there is strong drive to extract valuable insights through big data analytics. While this vision may be shared by many, the reality is that breaking down data silos and providing a universal, well-structured and interoperable representation of medical knowledge has been a long and difficult journey. However, progress has come from several different efforts, with AI working behind the scenes to help make this happen.

How to represent health data is a key objective. The primer in Chapter 2 outlined the area of knowledge representation and reasoning, specifically work on data models such as FHIR and medical ontologies such as SNOMED CT. While not immediately coming to mind when thinking about AI in health, good knowledge representation provides the foundation for many technical solutions, both AI and non-AI. Having structured, standardised and sharable medical knowledge has a swathe of benefits: it enables different systems to talk to each other; it greatly increases data quality, reduces errors (which can have major consequence in health data); it unlocks the opportunities for secondary use (big data analytics); and decouples the data from the system used to capture and store it.

Once the foundation has been laid with good knowledge representation, a host of analytics options become available. This is often an area where AI methods can be applied – and where they rely on well-organised underlying data. Standards such the SNOMED CT and FHIR facilitate this by providing AI models with such organised data.

Getting the data right and developing the AI on top of it still does not constitute success for AI in health. The final – and critical – step is integration within the healthcare system. The challenge of embedding an AI system within complex clinical workflows is often overlooked. However, new AI

[4] National Primary Health Care Strategic Framework, Standing Council on Health, 2013.

infrastructure platforms are emerging that help with integration. This technology both eases the effort of integration while adding rigour to the process to ensure safety and quality.

Much of the effort around data is not seen at the patient bedside. Nevertheless, data is pervasive in the background of all parts of the health system and effective management is a core criterion for success in digital health.

4.2 Precision Health

Precision health is about adapting healthcare to the individual. While it originated as adapting care based on genetic profile, it now encompasses the broader view of developing systems, technologies and therapies tailored to the individual.

Early detection of health conditions allows timely invention which is highly correlated with good health outcomes. Childhood neurological conditions in particular benefit from early intervention, because therapy can be implemented while the brain is at its most malleable, ameliorating atypical development in response to neural injury. Detection is often non-trivial, relying on subtle cues that may not be obvious for some time. In these cases, AI can help to pick these early indicators. ML of MRI scans can be used to detect early signs of Cerebral Palsy or Autism Spectrum Disorder. While MRIs provide hugely valuable insights, getting a scan is expensive and can be invasive for young children. Alternatives include using non-invasive sensor technology and ML of sensor data to provide another source of screening for early detection of disease. As more sources of data become available, there will be increasing opportunity for both early detection and targeted invention.

Analysis of the human genome continues to give important insights into human health. Increasing availability of genetic data is met with increasing efforts to analyse this data for better treatment options. There are a range of avenues from an AI perspective in this space. There is a large body of work on the practical side of how to produce ML methods that can cope with vast amounts of genetic data. This work has been given a significant boost with the advent of cloud computing, making it much easier to setup and run genetic analyses on shared computing infrastructure. There is also the innovation into new ML methods that can reveal genetic risk factors for certain diseases. Critical to much of this work is explainable AI: ensuring that the AI method is able to provide an explanation for why a certain prediction or output was given. In genomic medicine this might involve specifying which gene mutation was responsible for the risk of a certain disease. As AI systems become more integrated into the clinical decision-making process, the need for explainability will be paramount.

AI is helping to improve targeted treatments for major interventions which can have large side effects. Certain medical inventions, while certainly effective, can have quite adverse side effects. For example, radiation therapy for cancer treatment is effective against cancer but can harm healthy tissue. In such cases, it is important to deliver the radiation to a very precise area of the body. AI can be used to model and map out this area; following this the AI can, under the guidance of a clinician, deliver the radiation to just this specific area. Because each patient's anatomy and disease are different, the method needs to be calibrated to the individual. AI based modelling of anatomy and targeted treatment enables this precision. Going forward, we expect to see

increased patient specific analysis and targeted treatment, with explainability being a key criterion for integration into clinical practise.

Social interactions may not be thought of as an area where technology and AI play a role in health, but AI methods have emerged to help people with autism, speech issues and to assist in emotion detection in tele-medicine. AI systems have been used to communicate in a predictable and tailored way to people with autism, greatly increasing these people's ability to communicate comfortably. Coupled with advances in robotics, socially aware robots will increasingly be able to assist patients in a range of different scenarios. This area is being driven by rapid advances in conversational agent technology (e.g. chatbots). In the general domain, such technology is now standard in agents such as Siri and Alexa. Being able to understand and respond in natural language is the key underlying advancement; this is likely to have a range of applications in the health domain.

4.3 Virtual care

Australia's healthcare system has long been dogged by the tyranny of distance, resulting in inequitable access to healthcare. Virtual care aims to provide healthcare without being tied to a physical location (See Box 1). This reduces many of the overheads of having patients travel to central locations or allows patients to remain independent at home for longer.

There are many benefits to supporting elderly patients to continue to live at home. This both reduces the cost on healthcare system as well as maintaining patient wellbeing. However, ensuring that people are safe and healthy at home involves a range of healthcare activities; AI can assist in these. Regular health assessments are one important activity in providing care. These are done manually, at period intervals and often as qualitative assessments; this makes them costly, reflective of just a snapshot in time and not always an objective assessment. With the advent of low cost, non-invasive sensor technology, it is possible to get real-time, objective assessments of people’s health conditions. This can both inform early inventions if someone's health declines, as well as emergency monitoring for acute events (e.g. a fall at home). While the sensor technology provides the basis for this type of virtual care, the key addition is AI and ML methods that use the sensor data to provide insights in someone’s health. With new AI methods such as deep learning, greater insights from sensor data will be possible.

People in regional and rural areas have far less access to healthcare services; tele-health aims to provide access without the need for people to travel to specific health facilities. The need for social distancing as a result of the COVID-19 pandemic has also driven the demand for tele-health. Simply, tele-health might just involve virtual consultation with a specialist, but a host of other opportunities exist in which AI can assist. Basic screening and testing can be done outside of specialist facilities; for example, routine retinal scanning. AI can then be used for routine checking of results for specific risk factors and, if flagged, images can be sent to specialists in metropolitan centres for review.

Box 1: Virtual Care: From Mobile Health to Wearable Devices

AEHRC has been at the forefront of mobile health research since 2014, when we conducted the world's first validation of a mobile health enabled cardiac rehabilitation program. Under-utilisation of cardiac rehabilitation results in adverse patient outcomes and increased hospital admissions. We aimed to improve rehabilitation participation by providing an alternative to standard in-clinic care through mobile app delivered cardiac rehabilitation – including physical exercise, counselling, health monitoring and education support. The trial, with 120 participants, clearly demonstrated the acceptability of this option, with greater uptake, adherence, and a 30% increase in completion of cardiac rehabilitation compared to an in-clinic delivered program.

Since then the AEHRC has undertaken a number of trials in mobile health using a similar approach, while also expanding our research to include other sources of data beyond mobile phone apps. This includes sensors in homes and integration with medical devices to monitor and support the health and wellbeing of aged people living alone. We have developed wearable sensors for babies to monitor movement for early detection and diagnosis of conditions such as Cerebral Palsy. We have also investigated chatbots delivered via mobile phones to provide monitoring and support for a wide range of physical and mental health conditions.

These innovations and integration in mobile health, connected devices, and wearables shift the paradigm of tele-health and tele-medicine using video conferencing to 'virtual care', in which access to healthcare and data is pervasive, and can be delivered to the patient at home or in the community. This provides capacity for new and larger datasets, allowing novel analytics to be developed using AI, from monitoring, diagnosis to prediction of likely events at individual level.

The model of virtual care where data is captured remotely and sent to a central portal for review provides many opportunities for automated data analytics using AI and ML that benefit patients and clinicians. For virtual care to deliver the promised efficiencies of scale, the systems need to do more than aggregate the data for review – they must generate actionable information to allow prioritisation of patients, trigger explainable alerts and alarms to ensure safety, and they should support decision making by patients or clinicians. The potential to deliver smart services is only one component of the power of AI in remote monitoring. As the volume of longitudinal data collected via virtual care platforms grows, new opportunities for insights into human behaviour, disease trajectories, treatment outcomes, care pathways and more will emerge. This approach to virtual care also provides an opportunity for continual improvement on the AI tools developed for delivering care. The data being collected can be used to continually improve the AI and ML algorithms and provide more accurate personalised intervention, therapy and care.

References: Box 1: Virtual Care: From Mobile Health to Wearable Devices, page 75.

4.4 Hospital and Health Services

In many countries, demand for hospital beds has increased through an increased population, increased elderly numbers, increased community expectations, increased provision of hospital services, and a higher availability of therapeutic interventions[5]. These demand increases have taken place in finite capacity health systems, where bed stocks and staff resources have remained fairly constant on a per-capita basis.

As more and more data are available about hospital workflows, AI can help at optimising operations and predicting events to improve healthcare delivery. Risk prediction is a key area where AI can help. Using historic data on many patients, it is now possible to predict the risk of a specific individual patient. This might be the risk of deterioration (and need for specialist care), the risk of readmission to hospital, the risk of falls, the risk of self-harm, as well as many others. In all these cases, identifying the risk can drive early invention that may prevent an adverse event. New ML methods will continue to improve the accuracy of these predictions and the likelihood of adoption in routine clinical practise.

There is a strong drive to ensure efficient operation of hospitals. This involves both optimising the use of hospital resources as well as ensuring there are sufficient resources for surges in demand. Optimisation and scheduling algorithms have been used to derive the best use for resources (e.g. scheduling of surgery and operating theatres). Prediction of future demand (e.g. emergency department presentations) can also be done with a high degree of accuracy using AI. All these methods can be used by hospital administrators and clinicians to improve hospital operation and respond to changes in demand. Going forward, hospital operations will be increasingly data-driven with AI providing the support.

AI can provide an effective real-time monitoring of outbreaks or anomalies and errors. AI-based syndromic surveillance can be used to predict possible disease outbreaks (such as influenza). This will increasingly involve corroborating multiple data sources (including non-traditional sources such as social media) to make predictions. AI has been used to predict errors, anomalies and adverse events in clinical data; for example, flagging adverse drug events or antibiotic resistant bacteria in laboratory tests. This can reduce the incidence of medical errors, which have a significant negative impact of health. With more and more data available, and better AI methods, we will see an increasing number of these types of AI assistance in clinical practice.

[5] Epstein, et al. (2009) The national report card on the state of emergency medicine: Evaluating the emergency care environment state by state, Ann. Emerg. Med., vol. 53(1):4–148.

5 AEHRC Case Studies in Predictive Analytics and Data Driven Intelligence

The health system is a complex and distributed environment. In this Chapter, we present a wide range of different Case Studies on how predictive analytics and gaining insights from data can have a real impact on the health system.

Data-driven insights for hospitals and health systems

Hospitals now have large amounts of data: electronic patient records contain a wealth of information about individual patients, while other data are available on hospital services, resources and their utilisation. AEHRC has been working closing with various health departments and hospitals to help use this data for improved service delivery. Large amounts of patient data allow for predictions to be made about the outcomes of new patients.

At an individual patient level, we can use electronic records to identify at risk patients. Case Study 1 presents a project where ML was used to predict the risk of hospitalisation / readmission.

Case Study 2 looks at optimising elective surgery by modelling all the inter-connected departments requiring access to share resources. Analytics applications are really made possible when health data is standardised and in a form that supports analytics. Case Study 3 shows how real time analytics is made possible through interoperable data efforts such the FHIR standard. Case Study 4 looks at how these analytics can be used to predict future demand and patient flow and Case Study 5 shows how we can identify deteriorating patients and thus intervene earlier to prevent their condition worsening.

Insights from the human genome

Computationally processing genomic data presents a real challenge: the scale of genetic data can be huge, and many ML methods require significant compute power. Yet the rewards from insights into the human genome have the potential to revolutionise human health. In the past, the lack of compute power has hampered what is possible. These days, with the advent of faster computers and cloud-based deployments, much more is possible.

AEHRC's genetic computing projects aim to explore this, combining ML with cloud deployment. Case Study 6 aims to use random forest models to identify the underlying genetic causes of neurodegenerative diseases, thereby opening up new treatment avenues. Case Study 7 uses ML to help with the laborious curation task that pathologists must perform with genetic data. In addition, the whole process is ported to a serverless, cloud-based environment – greatly increasing performance and ease of access. Case Study 8 uses ML to guide effective gene editing. Case Study 9 presents a cloud architecture with ML to visualise and track the genomic fingerprint of the COVID-19 virus.

Insights from sensors for better health

Sensors have become ubiquitous in the home environment. Using sensor data, a number of valuable health insights can be made. Sensors in the home can aid elderly people to live independently in their homes for longer, which has many health and economic benefits. Case Study 10 used passive (non-wearable and non-intrusive) sensors to accurately measure how someone is coping at home and identify when they might need assistance. Where multiple people live together, Case Study 11 used ML to identify the different individual people. From the elderly to infants, Case Study 12 used miniature wearable sensors for early identification of infants at risk of Cerebral Palsy.

Case Study 1: Risk Stratification for Identifying Patients at Risk of Hospitalisation

Partners: Queensland Health, Metro South Hospital and Health Service, Australian Government Department of Health

Providing hospital services is much more expensive than primary care. Strategies to reduce unplanned hospitalisations include integrated care in the community based around General Practice (GP), while hospitals use discharge planning to reduce unplanned re-hospitalisations. However, identifying the patients who need these interventions can be a challenge.

AEHRC have developed algorithms that use data captured in clinical and administrative systems to identify patients that would benefit from appropriate hospital avoidance interventions. This involves linking various datasets containing information about patients and their interaction with the health system and extracting feature predictors from the linked dataset. The data is divided into training and test datasets. A model is then developed on the training data using a range of ML algorithms and validated on the test dataset. Logistic regression, naïve Bayes, artificial neural networks, random forests, generalised boosting were some of the candidate algorithms employed for model development. While each model offered its own set of advantages, the final models were chosen based on predictive performance, implementation constraints, and explainability.

For the Australian Government Health Care Homes integrated care trial, we used a linked data set of GP, hospital and death registry data, to identify patients who were at high risk of being hospitalised over the next 12 months. The algorithms have also been employed across more than 150 general practice clinics across Australia to identify chronic disease patients for the trial.

For a major metropolitan Queensland hospital, we used data collected during patient stays and linked to death registry data to build an explainable AI model. The model was designed to inform in-hospital care and the discharge processes, to identify patients who needed additional support to reduce the likelihood of readmission. For the last two years, this decision support tool has provided patient risk information to chronic disease nursing staff members involved in their care.

Ensuring that the results are explainable to the doctors and nurses is important. Figure 4 shows the web-based decision support tool using an approach to explainable AI in showing which risk factors contributed most significantly to the predicted risk level to assist clinicians in care and discharge planning.

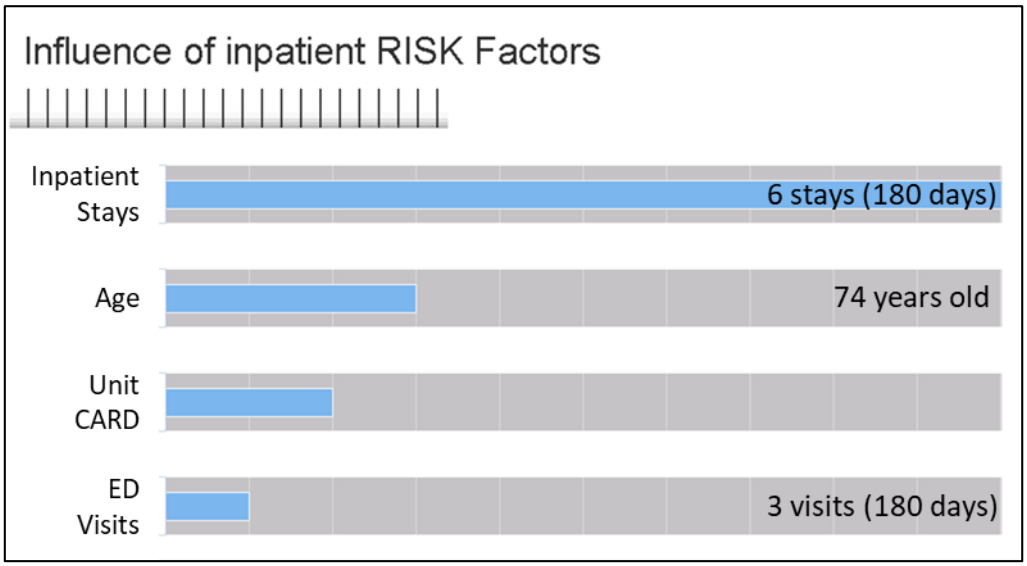


Figure 4: Employing explainability on machine learning output to support clinical decision making. This deidentified report shows the patient has had six inpatient stays in the past 180 days, is aged 74 years, has had previous admission to Cardiology Unit and three emergency department visits in the past 180 days. These represent the top four factors contributing to the patient's risk of readmission within 30 days post discharge. The blue bars represent relative influence of these risk factors.

References: Case Study 1: Risk Stratification for Identifying Patients at Risk of Hospitalisation, page 75.

Case Study 2: Improving Scheduling of Elective Surgery

Partners: Queensland Health, Griffith University

Elective surgery is a planned, non-emergency surgical procedure, which can be scheduled at the patient's and surgeon's convenience. However, efficient scheduling of elective surgery is difficult for a variety of reasons – including the availability of surgeons, nurses, theatres and equipment. In AI this is referred to as a distributed scheduling problem.

This problem is typically addressed by the hospital bookings department assigning patients to the surgery schedule based on their position in the queue and in consultation with relevant surgical teams. Other departments are responsible for scheduling of other clinical staff (nurses and anaesthetists) and equipment to support the schedule. In addition, 10% to 40% of scheduled elective procedures can be cancelled often due to lack of theatre time on the day caused by over-run of other surgeries. Accurate surgery duration estimation is therefore essential for optimal planning and hence efficient use of operating rooms.

AEHRC have developed novel algorithms to improve the scheduling of elective surgery. We used a Multi-agent System and the Distributed Constraint Optimization Problem (DCOP) formalism for modelling and solving naturally distributed optimization problems such as efficient elective surgery scheduling. Here surgical staff and other resources were represented as agents. Novel measures were then developed to emphasize the interconnectedness between each agent's local and inter-agent sub-problems, and used to guide dynamic agent ordering during a distributed constraint reasoning process.

An agent-based distributed constraint reasoning processes works by allowing intelligent software agents to generate optimal schedules on behalf of their respective departments (Figure 5). Each agent, trained with the appropriate constraints, preferences and priorities, optimizes schedules for their respective department and then negotiates to resolve inter-agent constraints. For example, an agent optimising the bookings schedule may require the agent managing theatre resources to have a certain machine available for a particular procedure. The resulting Dynamic Complex Distributed Constraint Optimization Problem (DCDCOP) algorithm significantly outperformed existing DCOP approaches. A proof-of-concept Multi-agent System implementation using this algorithm modelled the scheduling activity of four agents: Bookings, Nursing, Anaesthesiology, and Theatre Resources, and demonstrated the efficacy of the approach.

Figure 5: Agent interaction for distributed elective surgery scheduling.

Another project employed administrative and perioperative data from a large metropolitan Queensland hospital to investigate the performance of different ML approaches for improving procedure duration estimation. Initial ML models employed linear regression, multivariate adaptive regression splines, and random forests. The random forest model outperformed other methods, reducing mean absolute percentage error by 28% when compared to current hospital estimation approaches. Additional investigation focussed on operations with only a single procedure being performed. Ensemble boosting and bagging based ML algorithms delivered a 55% reduction in mean absolute percentage error when compared to the current hospital estimation approaches, a reduction of 18% when compared to the performance of random forests.

Hospital information systems and scheduling processes at each department remain disparate systems without the interconnectedness needed to support the use of these sorts of models. As systems mature, these intelligent algorithms will support optimal scheduling of staff and resources in our hospitals.

References: Case Study 2: Improving Scheduling of Elective Surgery, page 75.

Case Study 3: Real Time Analytics using Fast Healthcare Interoperable Resources (FHIR)

Partners: Alcidion and eHealth NSW

The deployment of electronic medical records (EMRs) at Australian hospitals is increasing the capture of high-quality patient data that can be used for patient care as well as safety, quality and research purposes. The increased adoption of standards – including terminologies, such as SNOMED CT, and common data models, such as FHIR (Fast Healthcare Interoperability Resources, Box 4), greatly enhance the usefulness of information in EMR systems and improve interoperability of health information shared among healthcare providers and organisations.

The FHIR resources, which contain standard terminology, demonstrate the importance of agreed and standardised data models and terminologies – the cornerstone of knowledge representation. Leveraging these standards facilitates the deployment and scalability of real time clinical analytics and decision support applications and is a key research area for the Health System Analytics group.

A recent partnership project demonstrated the efficacy of this approach. A predictive risk stratification algorithm developed by CSIRO was added into Alcidion's Miya Precision Real-Time FHIR Platform. CSIRO also supplied middleware that enabled data from the NSW Health EMR (from American company, Cerner) to be sent as FHIR resources from the Alcidion Miya platform whenever certain trigger conditions were met, e.g. a new pathology report was received (Figure 6).

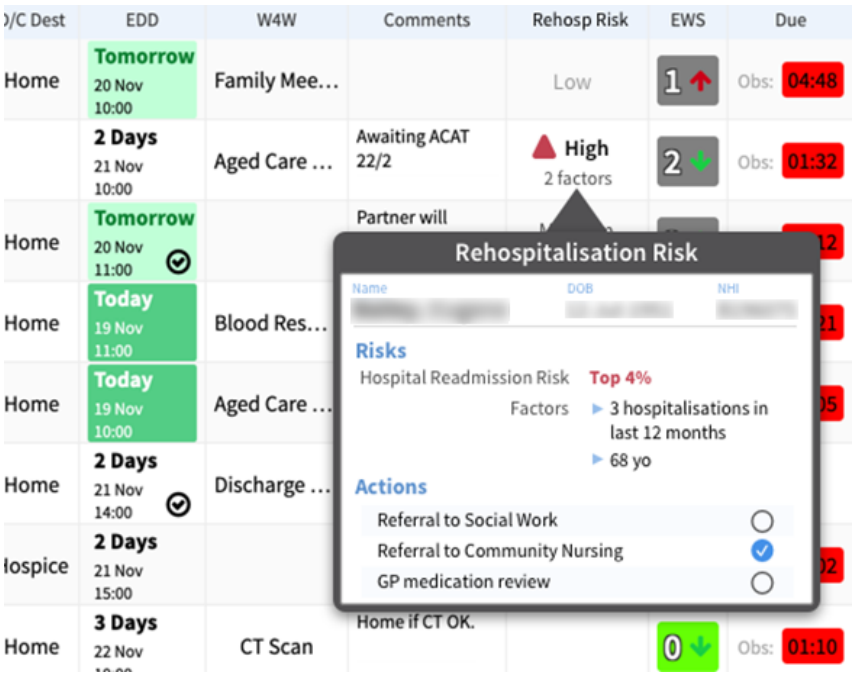


Figure 6: CSIRO rehospitalisation risk algorithm embedded within Alcidion's Miya flow dashboard.

The proof-of-concept CSIRO algorithm then calculated a risk score based on the data in the FHIR observation resources and returned a FHIR resource into the Miya platform for display on dashboards to support real-time decision making. This work demonstrates the potential for improved detection and management of patients at risk of readmission.

Case Study 4: Predicting Demand for Hospital Services

Partners: Queensland Health, and several other health services across Australia

Bed capacity management is a component of managing hospital performance. Using predicted bed demand information enables improvement in bed utilisation and effectively reduces access issues, allowing for better forward planning. In addition, predicting emergency and elective operating theatre case load on a daily basis is vital for improving patient access to surgery and improving theatre management.

AEHRC's Patient Admission Prediction Tool (PAPT) (See Box 2) provides estimates of future emergency department patient presentations, inpatient admissions and separations, and operating theatre arrivals (Figure 7).

The development of PAPT explored a number of modelling approaches to ensure accurate predictions including exponential smoothing, Poisson regression, autoregressive integrated moving averages, regression trees, generalised additive models using basic splines, exponentially smoothed state space models, and various ensemble approaches.

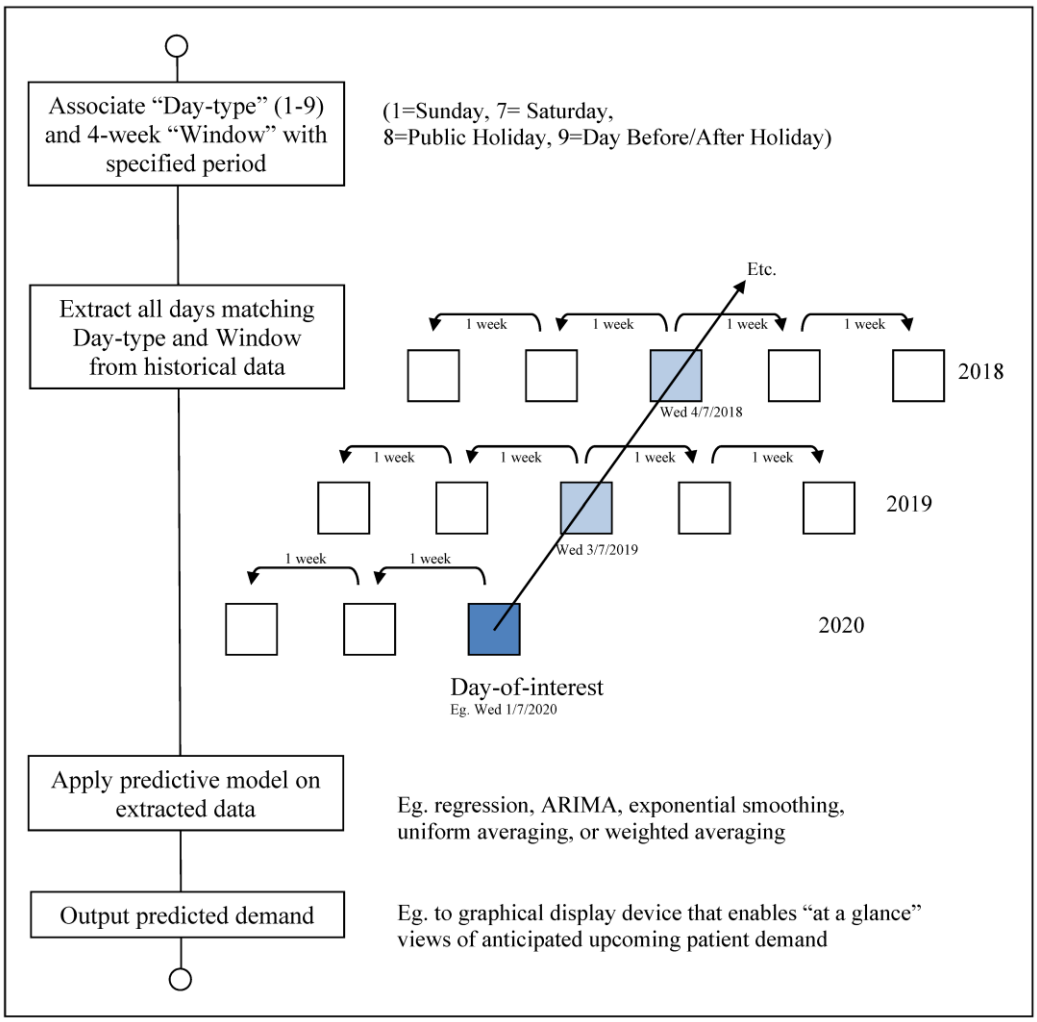


Figure 7: The Patient Admission Prediction Tool (PAPT) prediction algorithm.

Validation of predictive accuracy found that hospital demand can be predicted with a daily accuracy of around 90%. The approach which had the best predictive accuracy (a rolling 4-week window centred on the day of interest and providing weights to observations of matching 'similar' days) was subsequently implemented as PAPT.

Initially developed for Queensland Health, it remains a component of daily bed management across many Queensland public hospitals. It has also been implemented across health services in Victoria and Western Australia where it is used daily for short-term planning (1-3 days) in advance and for understanding overnight demand. Users have commented on the predictions provided by PAPT as very accurate and they use the information generated at Emergency Department huddles daily. They also use PAPT predictions for 2-week forecast and bed stock planning in the hospital and exploring options to implement advance bed management planning within the hospital.

References: Case Study 4: Predicting Demand for Hospital Services, page 75.

Case Study 5: Predicting Patient Deterioration

Partners: Queensland Health

Identification of patient deterioration can reduce the need for higher acuity care, hospital lengths of stay and admission costs and can even impact survival.

Patterns in illness trajectories and clinical intuition, such as peaks and falls, slow or sudden change, greater or lower frequency of vital sign measurement, can be used to enable potential life-saving clinical interventions. An electronic medical record with device integration can collect real-time physiological data including vital signs (e.g. heart rate, blood pressure, temperature, oxygen levels, etc.) as well as clinical data and reports.

Working with a major metropolitan Queensland hospital, we have collected 100 million data points from physiological data as well as pathology results from over 50,000 annual patient encounters. This data is being utilised to identify time-based features capturing temporal patterns from hospital encounters associated with patient deterioration.

Initial investigations used neural network-based semantic representation models to quantify the wealth of longitudinal and demographic patient information embedded in medical text and inform appropriate inputs to other learning models. Various statistical and ML algorithms including logistic regression, hidden Markov models and deep learning are currently being trained on these data and derived features used to identify and predict patterns that indicate patient deterioration (Figure 8).

The model will provide a live, continuously updated, real-time risk prediction that allows for early alerts when likely patient deterioration is identified.

References: Case Study 5: Predicting Patient Deterioration, page 76.

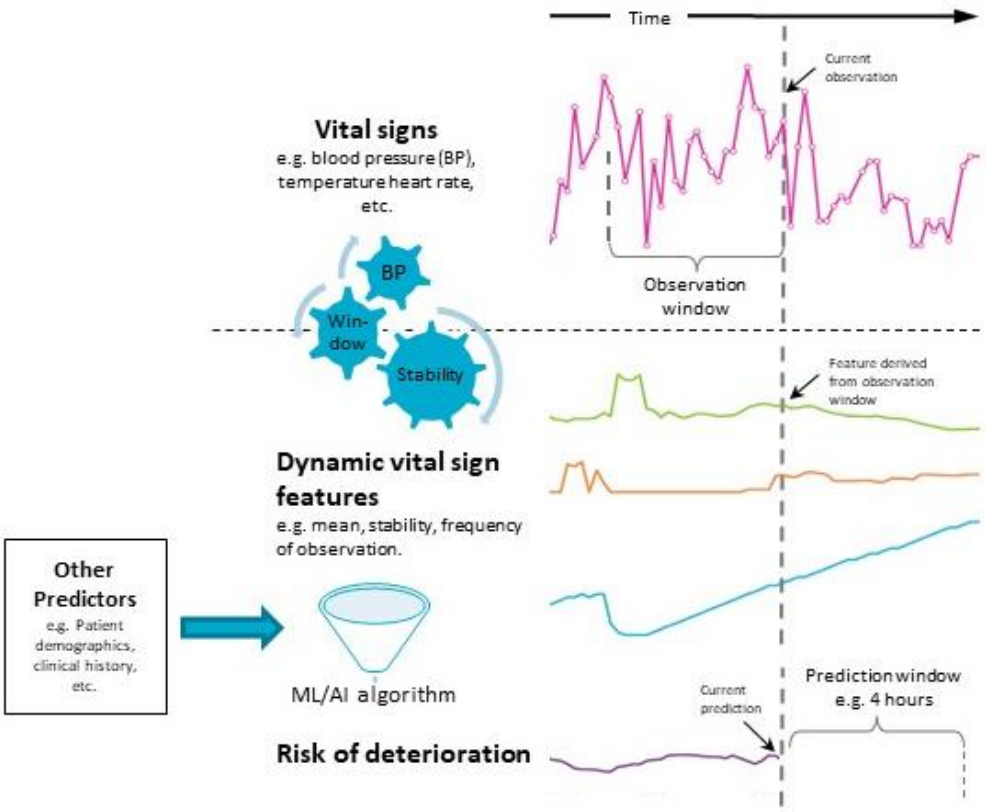


Figure 8: Risk of patient deterioration algorithm utilises real-time integrated electronic medical record data.

Box 2: Helping Health Systems through Times of Crisis

The Australian health system is often faced by crisis situations that create varying levels of additional and often unforeseen demand, affecting the ability of an already encumbered health system to deliver care. In instances such as these, AI solutions from CSIRO have proven especially useful to address the crisis at hand. Examples have included:

- The Patient Admission Prediction Tool (See Case Study 4) was used to predict demand on Emergency Department services for the annual Schoolies activity on the Gold Coast informing planning and deployment of services.
- The Patient Admission Prediction Tool was used to predict demand over winter months under various scenarios. This was used to predict capacity shortfall and quantify the impact of winter on state-wide hospital bed planning. The results were incorporated into the Queensland Government's winter bed planning as part of a 5-point plan to improve access to emergency department care when demand surges during the winter months.
- Syndromic surveillance algorithms, originally developed in response to the 2009 (H1N1 – 'swine flu') influenza outbreak and to support influenza-like-illness monitoring and other analytics, have been extended to the help manage the COVID-19 crisis.
- The algorithm developed for the Australian Government Health Care Homes integrated trial has been made freely available by Precedence Healthcare during the COVID-19 pandemic to member General Practitioner practices; it helps identify high-risk chronic disease patients for preventative care planning.

Case Study 6: Discovering Genetic Drivers for Amyotrophic Lateral Sclerosis

Partners: Project MinE consortium and Macquarie University

Amyotrophic Lateral Sclerosis (ALS) is a devastating neurodegenerative disease that currently leads to death within two years of diagnosis for approximately one third of patients. The underlying genomic drivers of the disease are not known and currently there is no treatment. We aim to use ML to identify potential genetic causes of ALS. Identifying the underlying genetic causes can open up new treatment avenues.

In our collaboration with Macquarie University we have the whole genome sequence of 800 individuals with ALS. Each sequence is three billion base pairs (or letters in the sequence) and there are on average five million differences (or mutations) in each sequence. Identifying the mutations that might cause diseases is similar to looking for a needle in a haystack.

With our VariantSpark technology we used a random forest classification algorithm. Random forest classifies datasets by generating a multitude of decision trees (potential solutions) for different probabilities of events, and then identifies the most likely solution by inferring across all decision trees. (See Box 3 for more on VariantSpark.) In this work we create decision trees based on the mutations in the genes of patients. Gene mutations that are deemed critical in multiple trees are likely causing the disease.

Due to the large size and sparseness of whole genome sequencing data and the computing power required to calculate all of the decision trees, our ML method is only possible with new distributed and cloud computing architectures that enabled us to develop VariantSpark (Figure 9).

Figure 9: Illustration of information flow in VariantSpark: getting insights from big and wide dataset through the use of distributed computing (Apache Spark).

Using VariantSpark in a study of Australian ALS patient genomes, we have identified a potential novel disease gene for ALS active in the Australian population. We are now engaging with Project MinE, an international research project on ALS with information on 22,000 individuals.

References: Case Study 6: Discovery Genetic Drivers for Amyotrophic Lateral Sclerosis, page 76.

Case Study 7: Machine Learning and Cloud Computing for Annotation of Human Genome Sequences for Clinical Reporting

Partners: Queensland Genomics and the Queensland Institute for Medical Research (QIMR)

When a patient's genome has been sequenced a genetic pathologist must review any changes in the sequence (called variants or mutations) to identify any variants that are the likely cause of a disease or would help inform treatment decisions.

As part of this clinical reporting, the pathologist categorises variants in the sequence as either 'pathogenic' or 'non-pathogenic' or levels of uncertainty in between. An automated pre-annotation stage reduces the huge number of variants in each sequence to a set of potential variants for the pathologist to review. This pre-annotation stage is slow and does not take into account a range of information – such as previous decisions made by the pathologist, data in local or international knowledge bases or data which is only reported in academic publications.

We have developed a new version of the variant effect predictor (VEP) software from the European Bioinformatics Institute which is in widespread use. Addressing both the flexibility and time constraints of VEP, our version uses cloud-native architectures to provide a highly modular and resource-efficient implementation of the variant annotation pipeline, serverless VEP (sVEP). This includes using ML methods to incorporate previous decisions from the curator (Figure 10).

Serverless VEP is estimated to be 99% faster than traditional VEP which will increase the turnaround time for pathology labs considerably and therefore indirectly improve the diagnostic process.

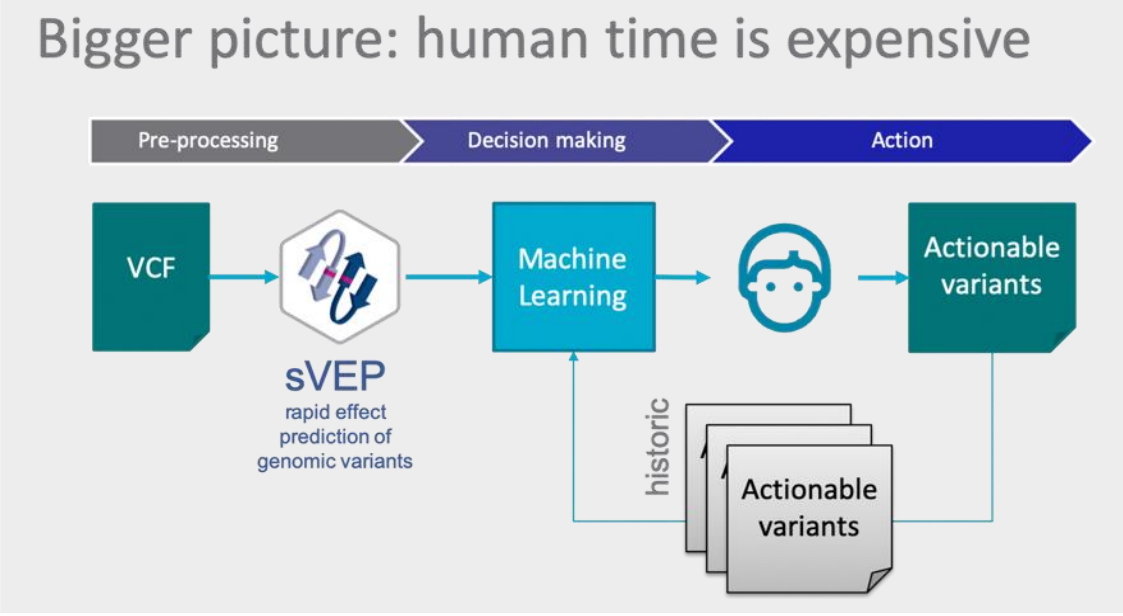


Figure 10: Illustration of the bottlenecks for variant curation (human time) and the use of the serverless variant effect predictor to speed up pre-processing of gene sequence variants, stored in Variant Cell Format (VCF).

We will partner with pathology labs to develop a ML-based methodology that fits into their workflow by using their historic annotation data to train a variant prioritization tool that mimics the decision-making process of a human.

Box 3: Cloud Computing for Machine Learning

Machine learning (ML) is data intensive – it is an iterative process that requires information to be kept in memory between each step. However, as the amount of data has grown, the workload has moved to be shared across multiple Central Processing Unit (CPUs). This combination of persistence and parallelisation is a difficult combination to achieve natively on traditional high-performance-compute cluster.

The ML community has re-implemented several ML algorithms on Apache Spark, a distributed computing framework, in a library called MLlib. While MLlib can be run on a local Spark-cluster, it is designed for elastic scalability, which is best offered by cloud providers such as AWS, Azure, Google Cloud. The genomics discipline has taken advantage of this scalability to a point where MLlib was not able to handle the scale of data now available.

To address this issue, a new random forest algorithm, VariantSpark, had to be developed. VariantSpark is 90% faster than traditional compute frameworks and requires 80% fewer samples to detect statistically significant signals. The benefits of speed and higher sensitivity enables VariantSpark to open up the usage of advanced, efficient ML algorithms on high dimensional data.

At the other end of the spectrum, where speed and efficiency rather than data-intensive workloads are required, cloud computing offers a different kind of distributed solution: serverless computing. Here, individual elements like compute, storage and communications are decomposed into individual services and only incur cost when consumed. This allows solutions that are agile and light weight, such as web services that serve the predictions of a trained ML method to a client.

References: Box 3: Cloud Computing for Machine Learning, page 75.

Case Study 8: Digitally Guided Genome Editing for Gene Therapy

Partners: Children's Medical Research Institute, Westmead, Australian National University

Genome engineering using methods such as CRISPR has revolutionized a wide range of research and industry sectors by allowing the precise editing of the genome of a living cell, thereby altering its function. However, delivering this precision in practice is difficult due to a wide range of influencing factors ranging from differences in the genomic makeup to environmental influences. It is critical to precisely understand the governing factors to ensure that the gene editing results in the desired change to the gene – shown below in the diagram (Figure 11).

We are improving our understanding of gene editing through the use of ML. Algorithms such as random forest and deep learning are able to analyse and learn from past experimental observations in order to discern rules governing gene editing activity. These rules can then be used predict and optimize the effectiveness of future applications. This has enabled us to develop our GT-Scan suite of software tools which aims to find the right editing spot for specific genes for specific applications.

Importantly, we found that the biggest difference to overall effectiveness is the way the training data is prepared, rather than the ML method per se. By preparing the training dataset that is appropriate for the specific editing application, we can substantially improve accuracy and effectiveness.

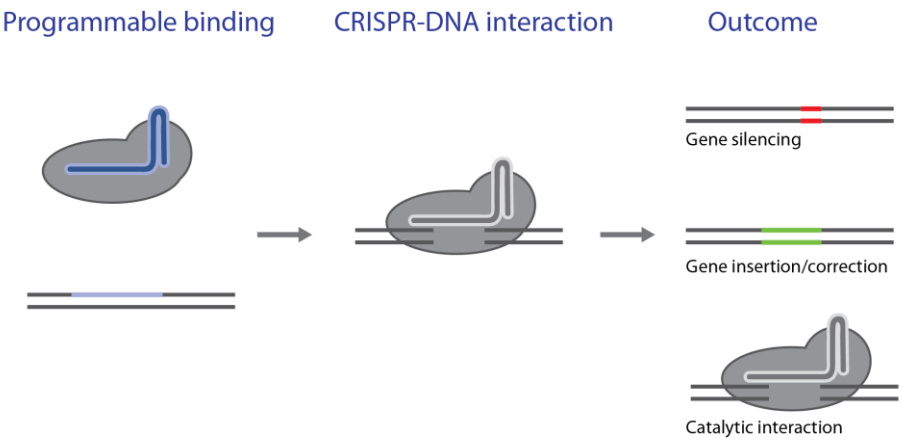


Figure 11: Illustration of function and use cases for CRISPR/CAS9 genome editing.

The trained ML models are used by our Australian National University collaborators to improve functional genomics experiments by using historic data from the lab to create a bespoke model that fits into the workflow. For example, the use of our CUNE software can help researchers design and optimize their approaches for inserting new sequences into DNA. Alternatively, VARSCOT can be used to identify the best gene editing approach for a specific individual, taking into account their natural genetic variation.

Our models will also contribute to Gene Therapy applications with our collaborators at Children's Medical Research Institute. The aim of this treatment is to correct causative mutations within a patient's DNA, reducing the need for invasive surgery. Our software helps by identifying the best gene-editing approach to take, as well as using a patient's genomic profile to personalize the treatment.

References: Case Study 8: Digitally Guided Genome Editing for Gene Therapy, page 76.

Case Study 9: Informing Public Health Responses for Infectious Disease Outbreaks such as COVID-19

Partners: CSIRO's Australian Centre for Disease Preparedness, India's Council of Scientific & Industrial Research Institute of Genomics and Integrative Biology

Due to the ever-increasing mobility of the world's population, localised infections can rapidly escalate into global threats – with an infectious disease declared a 'Public Health Emergency of International Concern' (PHEIC) on six occasions since 2009.

Novel analysis technologies are required to rapidly gain insights into large volumes of information to inform public health measures more efficiently.

We are aiming to make sense of the evolutionary drift a pathogen experiences while the outbreak progresses. We use dimensionality reduction methods (principal component analysis, T-distributed Stochastic Neighbour Embedding (t-SNE) and manifold learning) to visualise the genomic fingerprint unique to each virus isolate sequenced around the world (Figure 12). Dimensionality reduction methods enable the visualisation of objects by a two- or three-dimensional graph so similar objects are represented by nearby points and dissimilar objects by distant points.

In 2020, this technology was used to gain insight from the publicly available COVID-19 sequences and inform the choice of animal model generation at CSIRO's Australian Animal Health Laboratory (now the Australian Centre for Disease Preparedness), as seen in the below image. This is a crucial step in the development of diagnostics, treatments and vaccines.

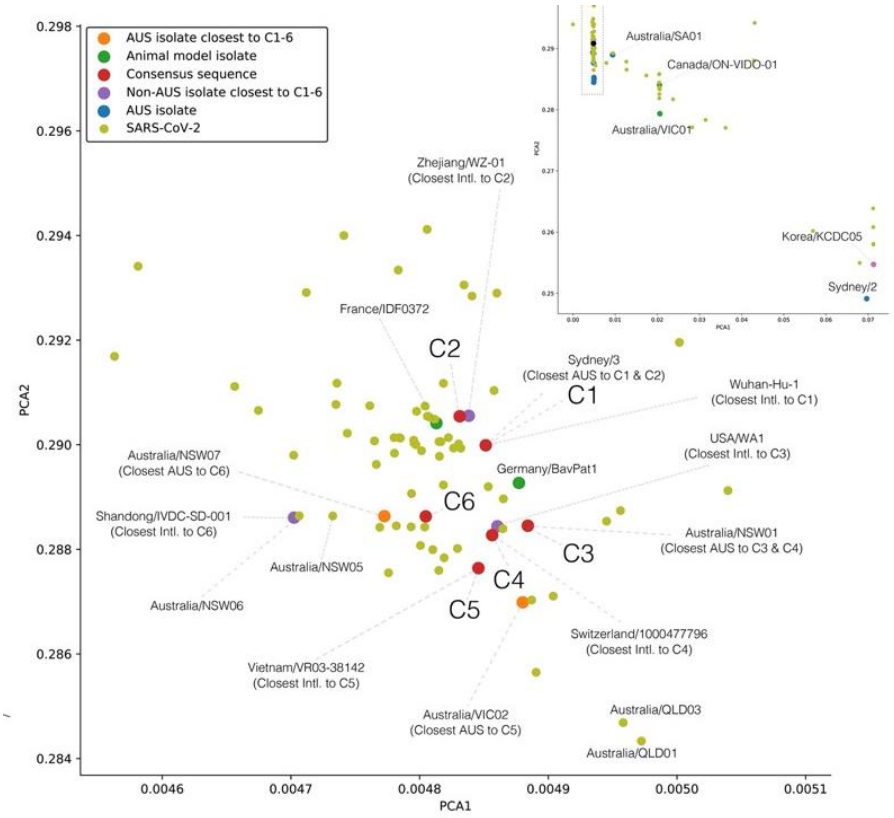


Figure 12: Principal component analysis of the SARS-CoV-2 genomes was performed in order to identify the best potential candidates for vaccine development.

For public health responses the virus data should be shared in a way that enables continuous real-time insights. Sharing using a cloud-architecture allows such distributed analytics – as demonstrated in Figure 13.

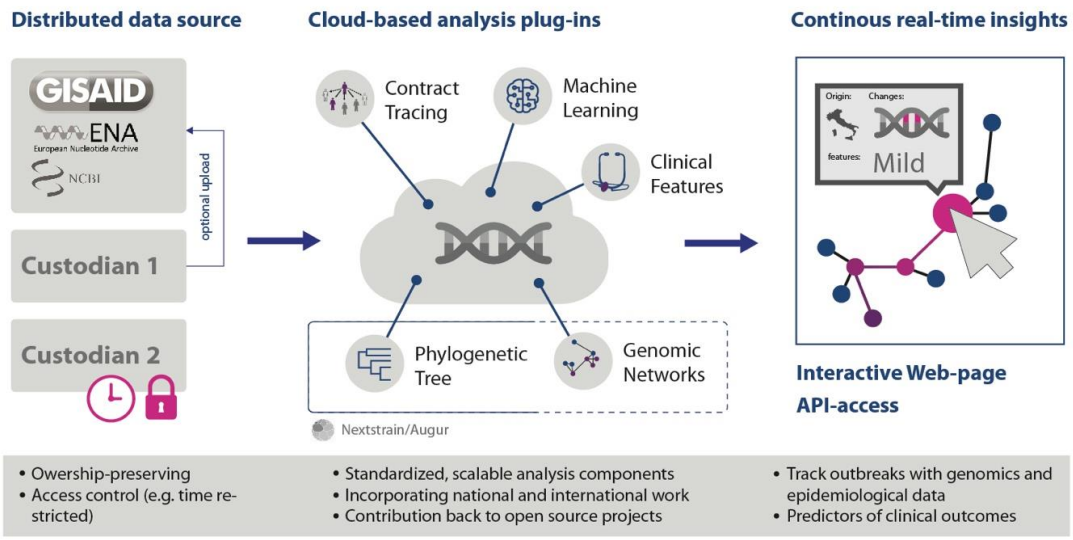


Figure 13: Illustration of how a distributed data system and cloud-architecture can help in providing robust standardised responses to public health emergencies like COVID-19.

References: Case Study 9: Informing Public Health Responses for Infectious Disease Outbreaks, page 76.

Case Study 10: Aging in Place and Objective Activities of Daily Living

Partners: Aged care service providers

Technology such as in-home sensors are increasingly being used to support elderly to continue to live at home. Sensor data can provide a rich stream of valuable information beyond simple alerts.

CSIRO's Smarter Safer Homes Platform (SSH) uses the sensor data to calculate a measure of functional independence called the 'objective activities of daily living' (OADL). The platform uses wireless sensor technology in the home including motion sensors, temperature sensors, and power sensors, to capture data that can be used to provide a view on the aged person's functional independence without the need for any wearable sensors.

The real-time time data streams are processed to provide the OADL measure. SSH uses unsupervised learning to extract features from the time series data. Multiscale clustering then detects patterns of interest – such as peaks and trends in the time series data. A decision tree fuses these patterns to compute the OADL measure**.** This measure provides the capacity to determine a person' ongoing functional level of independence referenced to their baseline profile (Figure 14).

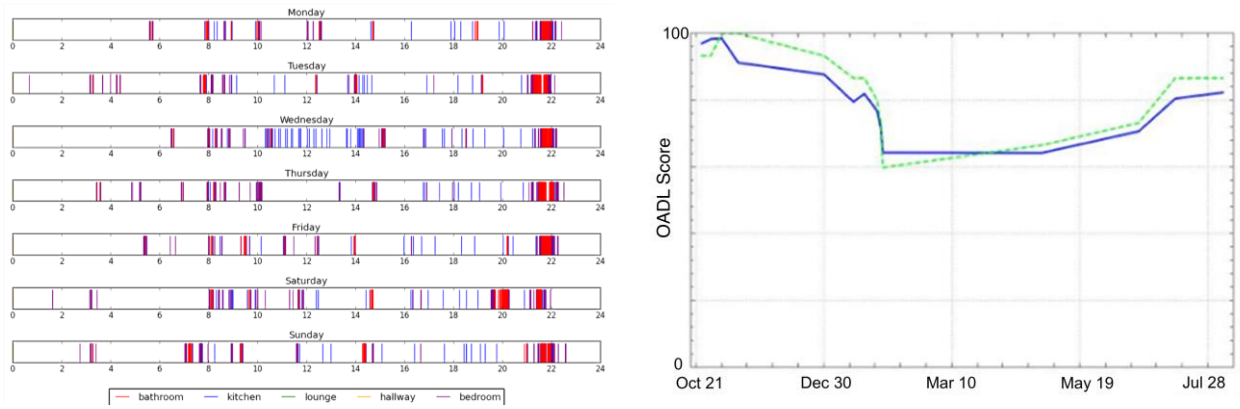


Figure 14: (Left) One week of sensor data from parts of the house which are then processed to provide the Objective Activities of Daily Living (colours denote rooms). (Right) Objective ADL (blue) computed from sensor data vs. self-reported level of functional independence (green).

A 200-home trial of the SSH platform is currently underway. Changes in the OADL are being used by aged care service providers to inform a number of interventions for older people living alone. Results of the trial will be available in late 2020.

References: Case Study 10: Aging in Place and Objective Activities of Daily Living, page 76.

Case Study 11: Using Deep Learning to Identify Individuals in a Multi-resident House

Partners: Aged care service providers

The Smarter Safer Homes Platform (See Case Study 10) analyses data from sensors deployed in a home setting to measure a person's 'activities of daily living'. This works well when there is only one person living alone in the house. However, in a home with multiple residents, activity recognition models designed for single living environments fail to yield satisfactory results because it is difficult to know whose data the sensors are capturing. Individual identification is one of the most crucial problems faced by multi-residential smart homes to fully realise their functionality and potential.

To deal with multi-resident houses the platform needs to identify sensor data from each person. An automated approach to distinguishing the identity of multiple residents is the use of an ultrawide band (UWB) radar sensor. This is a credit-card sized sensor unit deployed on the ceilings of the home. The data collected from the UWB sensor is a high frequency time series data stream. This time series data is processed in real time to extract unique patterns of individual residents.

The processing involves firstly using band-pass filters to encode the sensor data generated by residents as spectrograms. Figure 15 shows these spectrograms for two different residents with different walking patterns. Unsupervised feature learning is then used, in this case we use a support vector machine, to extract patterns of individual resident from those spectrograms. A 16-layer convolutional neural network was used to classify individuals. The location of each individual at a certain time is then known and can be correlated against the data captured by other sensors to inform the calculation of the objective activities of daily living.

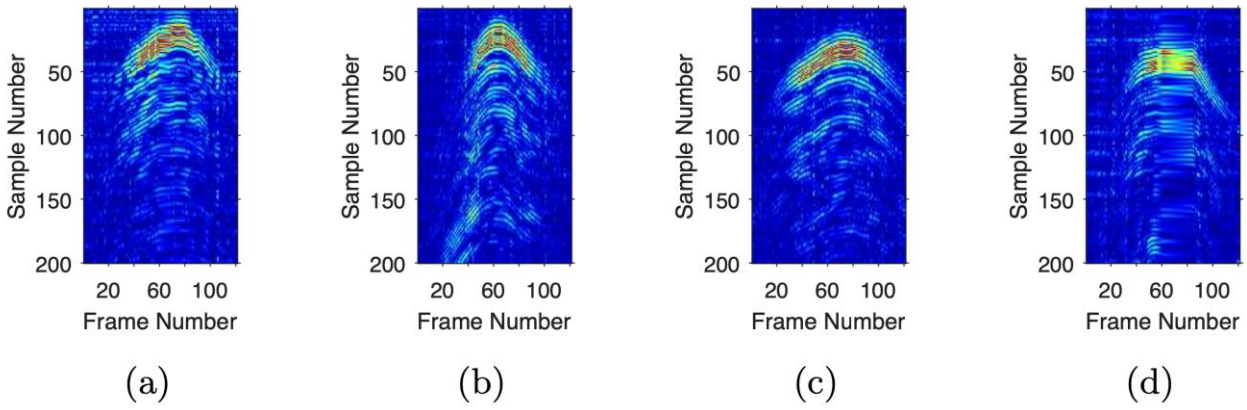


Figure 15: Ultrawide band sensor data of two individuals walking within the vicinity. (a) and (b) indicate Person 1 walking straight and diagonal; (c) and (d) indicate Person 2 walking straight and diagonal.

This method is considerably advanced compared to previous approaches – such as video surveillance which has privacy concerns, or using wearable devices which can be inconvenient. Currently the technology is being evaluated in an office environment with 14 persons. It can identify an individual person with 92% accuracy using only the UWB sensor data. This will greatly increase the accuracy of an individual's activity recognition in a multiple-resident home.

References: Case Study 11: Using Deep Learning to Identify Individuals in a Multi-resident House, page 77.

Case Study 12: Movement Classification for Early Diagnosis of Cerebral Palsy and Related Neurodevelopmental Disorders

Partners: Queensland Cerebral Palsy Rehabilitation and Research Centre (QCPRRC) and Queensland Early Detection and Intervention Network (QEDIN-CP)

Cerebral Palsy (CP) is the most common movement disability in childhood, affecting 1 in 700 children in Australia. The average age of diagnosis of CP is 18 months, despite the underlying causal neural injury being present at or around the time of birth.

We have developed a suite of miniature wearable movement sensors that can be used to measure motor development in infants from 3-5 months of age, towards early identification of infants at risk of CP.

The sensors use an inertial measurement unit to measure absolute orientation of the hands, feet, head, and trunk of the infant without the need for clinical supervision (Figure 16). Time series sensor orientation data have been used to develop a recurrent neural network (RNN) to classify infant movement along a developmental motor trajectory to determine whether the infant is at high risk for CP. Time and frequency domain movement characteristics are extracted and used as input features for the RNN, while clinical observations of the same movements from neonatal physiotherapists are used to create a labelled training dataset. This approach allows us to use a validated clinical metric to train the classification algorithm, thereby expanding neonatal motor assessments beyond the traditional medical environment while maintaining their diagnostic utility.

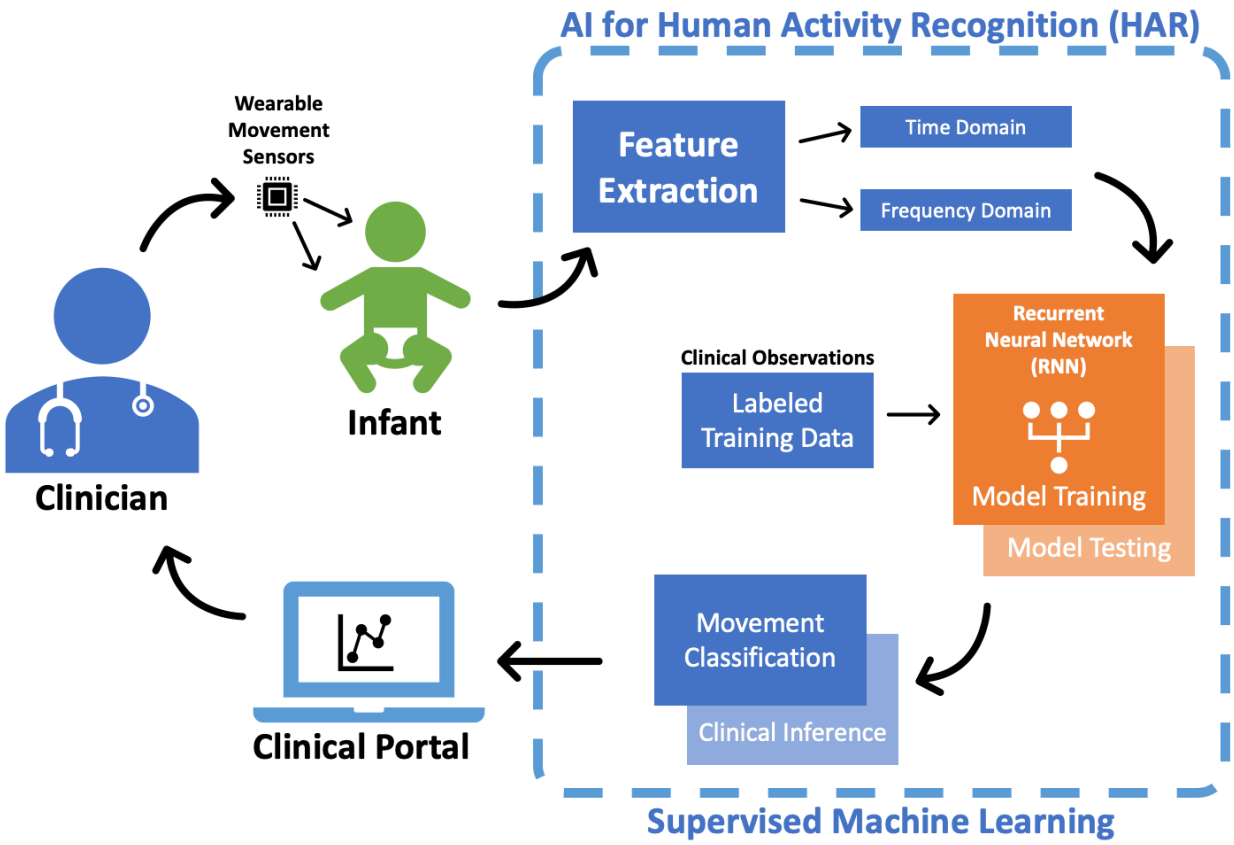


Figure 16: Data workflow for automated classification of infant movements indicating high-risk of Cerebral Palsy.

In partnership with the Queensland Early Detection and Intervention Network (QEDIN-CP), we are expanding screening of infants at high-risk of Cerebral Palsy in regional and remote areas of Queensland to improve early detection and facilitate targeted early intervention.

References: Case Study 12: Movement Classification for Early Diagnosis of Cerebral Palsy, page 77.

6 AEHRC Case Studies in Knowledge Representation and Reasoning

A number of groups across the AEHRC develop solutions which use AI for knowledge representation and reasoning.

Knowledge representation using ontologies

The primer in Chapter 2 outlined that ontologies are how knowledge of certain parts of the world is described – they use formal logic to define shared knowledge in a way that a computer can process and automatically infer new knowledge. From this a software program (called a reasoner) can use those rules to draw inferences and create new knowledge based on each ontology and the type of description logic which is used to build the ontology. Case Study 13 covers the "Snorocket" reasoner, developed by AEHRC, to reason using the SNOMED CT medical ontology. Snorocket is an extremely fast reasoner able to leverage the vast medical knowledge in SNOMED CT.

Extending medical ontologies

One key advantage of the formal logic of ontologies and reasoners is that it can be extended to support new domains. In Case Study 14, we outline how ontologies and reasoners can be extended to provide support for medications. Medication definitions require the use of reasoning with numeric values such as dosages – something that was not possible in existing formal logics in ontologies. AEHRC solved this by extending the Snorocket reasoner, thus ensuring the consistency and quality of the modelling of pharmaceutical information. Medications change rapidly with new products on the market all the time, so it is difficult to keep medications ontologies up to date. Case Study 15 solves this problem by analysing medication lists and automatically generating the appropriate medications knowledge in the medical ontology. In Case Study 16, we show how new medical knowledge can be added through 'post-coordination', whereby new concepts can easily be defined using the existing formal logic.

How knowledge representation supports analytics

Knowledge of how to use the ontology and its rules and properties supports the use of the ontology in many applications – including data analytics, search engines and NLP. The representation of knowledge in this way is a core part of AI. Case Study 17 presents Pathling, an advanced analytics service that exploits this standardised medical data to provide APIs that enable data visualisation, dashboard analytics, patient cohort selection and data preparation services.

Integrating AI into clinical workflow

A lot of attention is focused on new AI methods and how they might help health. Less attention though is focused on how to integrate such methods into the complex workflow of clinical practice. Good AI methods often fall at this last hurdle. Case Study 18 presents FORTE, a FHIR-based Workflow Platform for integrating AI into the Radiology Clinic. This provides a means of integrating automated methods into an existing clinical workflow, including ensuring clinician input when needed and or desired.

Box 4: Fast Healthcare Interoperable Resource (FHIR)

Fast Healthcare Interoperable Resources (FHIR) is the new standard from HL7 International describing data formats and elements (known as "resources") and an application programming interface (API) for exchanging electronic health records (EHR). FHIR provides the *Information Models* as described in Section 2.5 that enable health data to be exchanged in a way that different systems know what each element (or field) in the information model contains.

FHIR is designed so that there will be a manageable number of Resources. For example, there are Resources for the Patient, Condition and Procedure, but not for fine-grained or overly specific concepts such as address or blood pressure (which are elements in other Resources), or for larger concepts such as a patient record (which might be made up of many Resources). This has led to a dramatically faster uptake of the standard in the health sector, particularly in the USA, where the standard is now regulated by law as the mechanism for patients to access their health data.

SMART on FHIR builds on the FHIR standard by providing a standard – and standards-based – mechanism for third-party apps to access FHIR data stored in an electronic medical record. It does this through the definition of security protocols (based on the OAuth2 standard) for controlling which data can be accessed, and a context system that allows the sharing of clinical context between the app and the EMR. The main benefit of this standard is that it enables the same application to run in multiple systems, effectively enabling the creation of "app stores for health".

We have been instrumental in building the clinical terminology sub system within the FHIR standard and our Clinical Terminology Server, Ontoserver, is the world leading FHIR terminology server (Box 5). We have also built a number of technologies using the FHIR technology – such as the Pathling data analytics engine (See Case Study 16). We are also developing a number of applications using SMART on FHIR – including applications for genomics, medical education and medical image analysis.

The FHIR standard is driven by a large community of users and developers around the world who often meet in large and small *connectathons* to develop FHIR based applications. FHIR has also introduced what is known as a Community Process – where the technical and clinical community get together to develop a *FHIR Implementation Guide* – which describes the resources and terminology value sets for a particular use case. We have recently led the development of the FHIR Implementation Guide for General Practice in Australia – and an international one in collaboration with GSAID to capture COVID-19 related clinical data.

Case Study 13: "Reasoning" over Medical Knowledge – the Snorocket Classifier and SNOMED CT

Partners: SNOMED International

SNOMED CT is the largest medical ontology currently available and is now in widespread use for collecting clinical data to support clinical and secondary use. SNOMED CT is an ontology which uses the EL+ form of description logic (DL) to describe the rules and properties, or axioms, of the

ontology. The ontology is thus a knowledge base describing the interrelationships between each of its concepts. The axioms underpin these relationships and provide formal and computable definitions of each of the concepts. Figure 17 shows a part of SNOMED CT representing pain in the lower limb.

DL classifiers are special programs that use ontologies such as SNOMED CT to 'reason' about medical knowledge. For example, in Figure 17, the classifier is able to understand the calf is a lower limb body structure and pain is a clinical finding. Doing such reasoning in very large ontologies such as SNOMED CT is computationally expensive. When SNOMED CT was originally created, it used a very restrictive type of DL – a practical decision due to the computational cost, at the time, of the best-known algorithms for more expressive DL. However, this did mean that there were restrictions on what was able to be defined in each clinical concept.

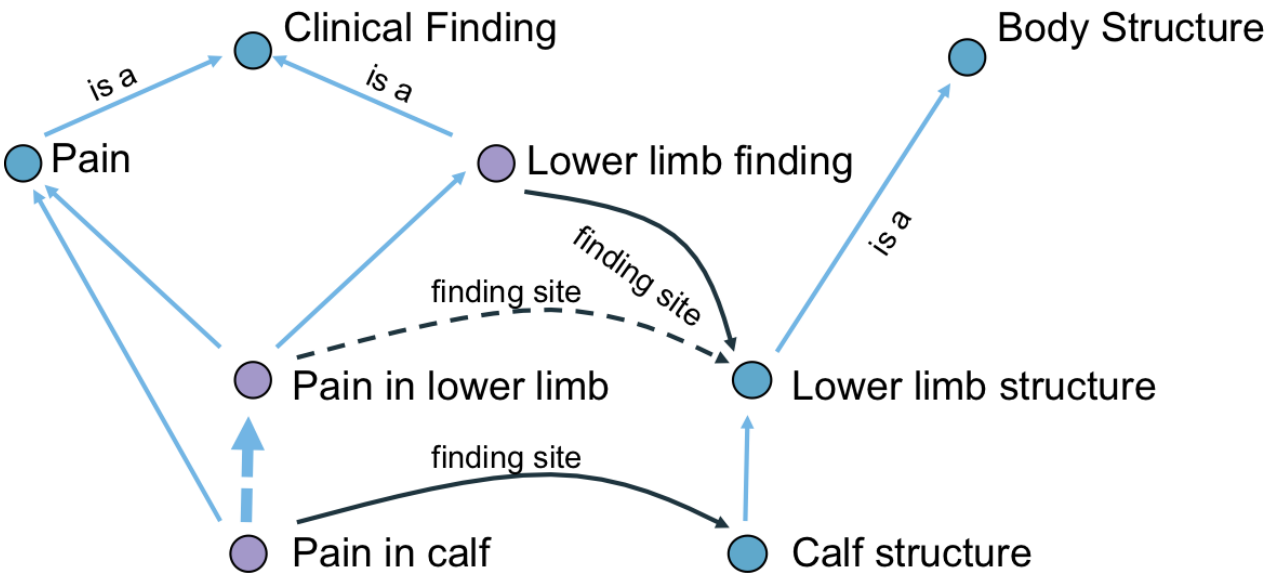


Figure 17: A sample part of SNOMED CT showing different clinical concepts and the relationships between them. The dashed lines represent relationships between concepts that can be inferred by a description logic automated reasoner.

To overcome this, we developed the first version of the Snorocket reasoner, which supports a more expressive form of DL. Snorocket was extremely fast — able to "classify" all of SNOMED CT in under 60 seconds as opposed to over an hour for other reasoners.

Snorocket was licensed to SNOMED International for use in the tooling used to author new clinical concepts and support the biannual release of SNOMED CT. The speed of the Snorocket reasoner in classifying SNOMED CT contributed to improving the quality of SNOMED CT. Subsequent reasoners such as ELK, and the latest version of Snorocket, use concurrent algorithms and have further reduced reasoning time to a few seconds. Such a speed-up results in a qualitative improvement to the ontology authoring process as it enables near real-time feedback on changes.

CSIRO still contributes to the international standard to develop SNOMED CT's use of this increased expressive capability.

References: Case Study 13: "Reasoning" over Medical Knowledge – the Snorocket Classifier and SNOMED CT, page 77.

Box 5: Ontoserver, Shrimp and the National Clinical Terminology Service

Having great standardised medical knowledge has many advantages but does not guarantee that others can make use of such resources in a straightforward manner. We have developed a number of software tools that help with the implementation and use of SNOMED CT in a wide variety of use case. Many of these tools are mentioned in the knowledge representation and reasoning Case Studies and they all make use of the relationships encoded in the SNOMED CT ontology.

These tools include:

- *Ontoserver* – our world-leading terminology server that supports fast, incremental search, and context-specific result ordering. Ontoserver supports the use of multiple terminologies including SNOMED CT, Australian Medications Terminology, LOINC, and other FHIR-based CodeSystems.
- *Shrimp* – our free SNOMED CT search, visualisation, and exploration tool that runs in your browser, and talks to any Ontoserver instance using its RESTful FHIR API (accessed via https://ontoserver.csiro.au/shrimp/. Many of the diagrams of the SNOMED CT concepts in this Chapter are taken from the Shrimp browser.
- *Snapper:Author* – enabling the creation and maintenance of the FHIR terminology resources: CodeSystem, ValueSet, and ConceptMap.
- *Snapper:Map* – streamlining simple ConceptMap authoring and maintenance
- *Pathling* – a FHIR server with functionality designed to ease the delivery of analytics-enabled apps, and to augment tasks related to health data analytics. Information about Pathling is given in Case Study 16.

Our tools also support national digital health infrastructure. The Ontoserver terminology server is part of the National Clinical Terminology Service from the Australian Digital Health Agency. As part of our collaboration with the Australian Digital Health Agency, Ontoserver can be licensed fee-free for end-use in Australia. The aim of this licensing arrangement is to support both the private and public health sectors in Australia in the implementation and use of SNOMED CT.

Case Study 14: Representing Medication Information – The Australian Medicines Terminology

Partners: Australian Digital Health Agency

A machine readable, standardised representation of medications allows different health systems to interoperate easily and allows AI systems to ‘reason’ about this data for a host of applications from decision support to advanced analytics.

To support consistency across medical concepts and medicines, the current version of the Australian Medicines Terminology (AMT) was developed with the same variation of description logic, called EL+, as SNOMED CT. This presented a problem as most EL+ reasoners do not support reasoning with numerical values, that is, they do not support numerical operations such as “equals”, “less than” or “greater than”. When dealing with medicines, these operations are vital to enable the correct modelling of pharmaceutical concepts such as defining strength and quantity.

This capability is also required to support the use of description logic (DL) reasoning for decision support algorithms.

AEHRC solved this problem by extending the Snorocket reasoner to support the DL concept of concrete domains. Concrete domains extend the DL reasoning to concrete objects such as numbers and time intervals – requiring considerable changes to the Snorocket reasoner.

Adding support for concrete domains to Snorocket enabled AMT to define knowledge such as "this tablet contains 500mg of Paracetamol as its active ingredient" (Figure 18).

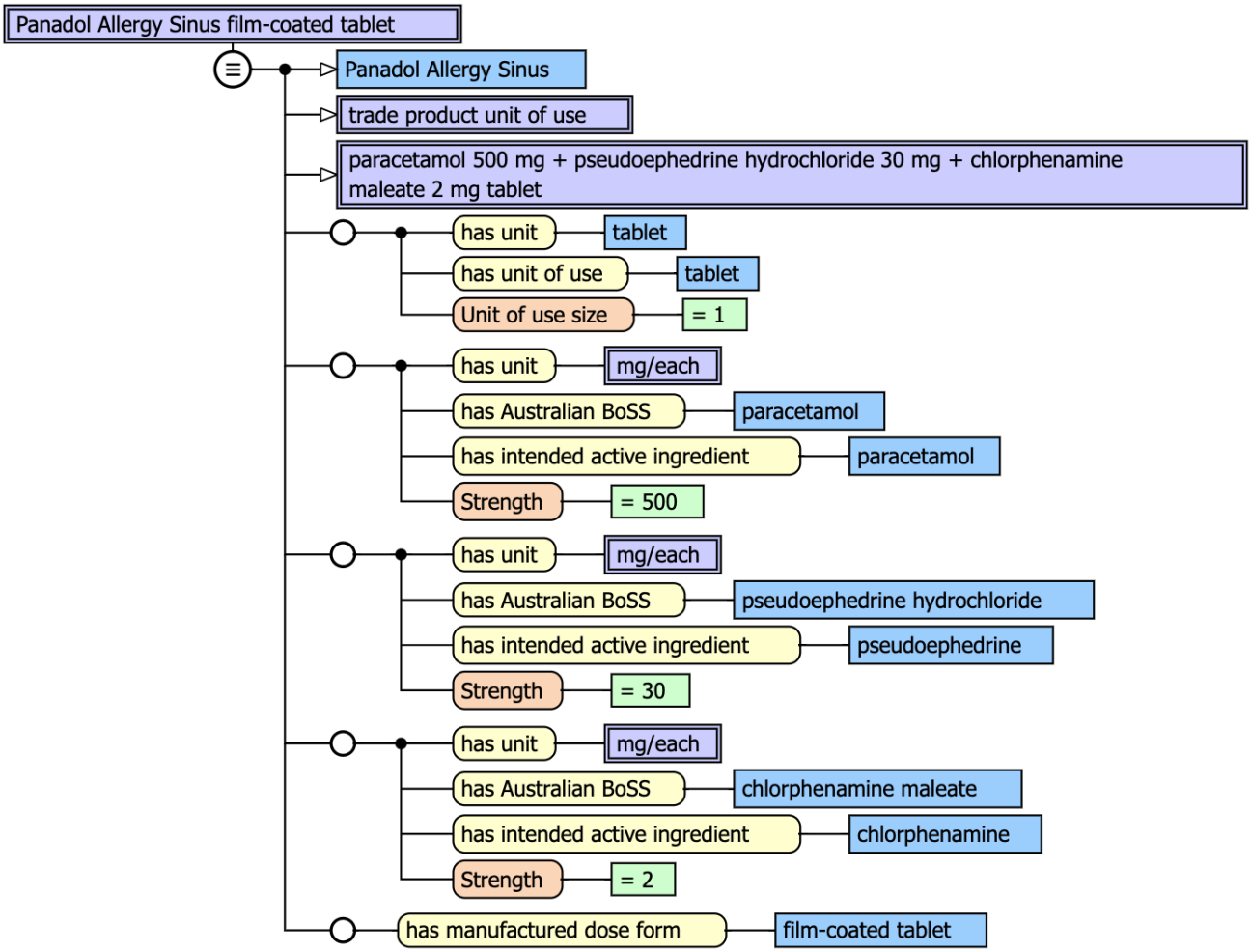


Figure 18: A diagram showing the definition of a medication concept including its properties. In this case we can see the active ingredients as well as the units of each active each ingredient.

The use of numeric values in the new model significantly increased the consistency and quality of the modelling of pharmaceutical information in the AMT. SNOMED International has since adopted the more expressive DL foundation that our work and tooling pioneered. This has paved the way for significant improvements to the modelling quality of pharmaceutical concepts (following the success of AMT) as well as anatomy concepts, that latter of which provide a grounding for most of the disease/disorder and clinical procedure concepts.

AMT is now included in all releases of SNOMED CT-AU and is increasingly used in clinical systems. This provides the basis for advanced clinical decision support, such as adverse drug reaction monitoring and medicines reconciliation.

Case Study 15: Generating Knowledge Representations: Developing Medication Ontologies from Semi-structured Text

Partners: Australian Digital Health Agency

The Australian Medicines Terminology (AMT) provides unique and unambiguous codes and descriptions for all commonly used medicines in Australia. AMT is part of the Australian release of SNOMED CT and, just like the other clinical content, medicines are modelled using description logic, enabling clinical decision support, prescribing, and dispensing.

As part of a feasibility study done in partnership with the Australian Digital Health Agency, the AEHRC developed a rules-based system that takes the inputs used to maintain and add new medications and attempts to create the corresponding AMT axioms automatically. The system uses a segmentation algorithm, implemented as chain of functions, to assign labels to different sequences within the text. These labels are used to create a tree structure that represents the text, which is then used as input to create the axioms. The approach was evaluated using existing data in AMT (Case Study 14) and showed promising results.

The data used to maintain AMT is largely drawn from the Therapeutic Goods Administration and the Pharmaceutical Benefits Scheme, with new medications being sourced from semi-structured information from the Australian Register for Therapeutic Goods. To ensure up to date medicines information is available, AMT is released monthly, meaning that any new medications, or changes to existing ones, are incorporated into each new release. However, this is mostly a manual task, making it both time-consuming and costly. The promising results of this study suggest that the approach developed by the AEHRC can be integrated into the AMT release workflow and will have a significant impact in reducing the effort and cost of maintaining AMT up to date.

Case Study 16: Post-coordination and the SNOMED CT Expression Constraint Language

Partners: SNOMED International, HL7

Coded data selected from a terminology such as SNOMED CT is stored within an information model such as a form within an Electronic Medical Record (EMR). Ensuring that the information model can use the terminology properly is key to extracting the most value out of the collected data – and providing value to the clinician who entered the data and downstream use of the data.

The AEHRC has worked closely with the core HL7 FHIR team and SNOMED International to ensure that the advanced knowledge representation and reasoning features employed by SNOMED CT are accessible through the FHIR API. This work is embodied in two key features: ensuring the SNOMED Expression Constraint Language (ECL) is easily usable from FHIR and including post-coordination support across the broader FHIR API.

SNOMED's ECL represents a bridge between the open-world semantics of the underlying description logic, and the more familiar and more easily manageable closed-world semantics of modern database systems as characterised by relational algebra and SQL. This was done by formulating ECL's semantics in terms of a relational model over a pre-computed "substrate" (both

concept hierarchy and concept to concept relationships). For simple SNOMED use-cases, a description logic reasoner can be used for a once-off pre-computation of the substrate. An example use of ECL might be to find all tumour staging concepts relating to skin cancer (Figure 19). This ECL statement will return all of the relevant concepts which can then be used to within a query to the EMR to identify patients with those stages of skin cancer.

```
descendantOf 385356007|Tumour stage finding|:
  363698007|Finding site| = descendantOrSelfOf 39937001|Skin structure|
```

Figure 19: The expression constraint language syntax defining all the tumour staging concepts relating to skin cancer.

Post-coordination is the act of creating a new SNOMED CT concept dynamically by adding additional properties to an existing concept within an existing information system. The new concept is not part of a formal release of SNOMED CT and exists only within that system. To efficiently support post-coordinated concepts, incremental reasoning is required, as pioneered by Snorocket, so that the substrate can be extended before evaluating the expression for use such as data analytics or decision support (Figure 20).

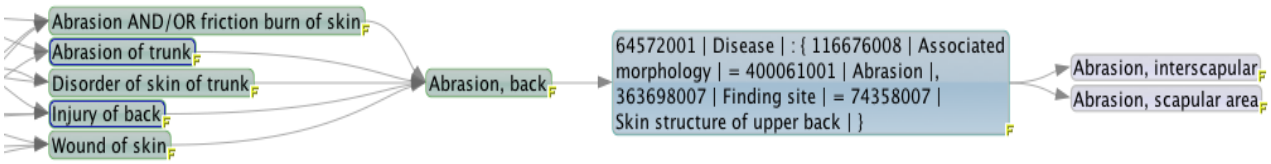


Figure 20: Illustration of the SNOMED CT hierarchy including a post-coordinated expression for an abrasion to the skin of the back showing its immediately more general concept (abrasion of back) and more specific concepts (interscapular and scapular area abrasions).

SNOMED's post-coordination grammar provides a mechanism to define an unbounded number of post-coordination expressions that combine the atomic concepts of SNOMED to express additional clinical detail. In this context, SNOMED becomes an infinite code system, which poses particular challenges for use. For example, in a request to return all descendants of Appendicitis, there are a theoretically infinite number of such descendants when including all potential post-coordinated expressions.

FHIR allows additional information to be associated with codes in an existing code system using a mechanism called a *supplement*. We can construct such supplements of SNOMED CT with post-coordination expressions harvested, for example, from data in EMRs. These supplements then form a library of post-coordination expressions. Using FHIR code system supplements in this way establishes a clear boundary between the open world semantics of the terminology and the closed-world reasoning required when working with actual health data.

Both of these capabilities are provided through the AEHRC's world leading FHIR Terminology Service, Ontoserver.

References: Case Study 16: Post-coordination and the SNOMED CT Expression Constraint Language, page 77.

Case Study 17: Data Analytics with FHIR and SNOMED CT

Partners: Australian Genomics, Royal Brisbane and Women's Hospital, Kidgen project

There is an increasing imperative for generalized tools for data analytics and business intelligence to insulate their users from the complexity inherent in health data. Users often deal with this complexity by writing large volumes of code to cover the domain-specific functionality missing from their tools. This code is sometimes shared in the form of libraries and packages but is more often written and maintained in isolation from other users.

More data is now being collected using SNOMED CT and shared using FHIR – which provides an opportunity to use these two standards to build advanced analytics tools on top of this data. We have developed Pathling, an advanced analytics service that exploits this standardised medical data to provide APIs that enable data visualisation, analytics dashboards, patient cohort selection and data preparation services.

Pathling encapsulates a set of functionality useful for health data analytics application development and workflow into a server implementation. Pathling understands the FHIR data model it ingests and serves, and it can integrate with a FHIR terminology server to enable the use of the description logic underpinning SNOMED CT.

We recently used the Pathling tool to perform an advanced analysis of genomic phenotype data which was collected using FHIR and SNOMED CT. In this set of data, differential diagnoses were collected at stages through the patient journey using SNOMED CT. As more testing was undertaken (including whole genome sequencing) Pathling was able to use the SNOMED CT semantics to understand the change in diagnosis – from a general diagnosis to a more specific diagnosis, or potentially to a completely unrelated diagnosis. The below Sankey diagram was automatically generated from the query response data returned from the server.

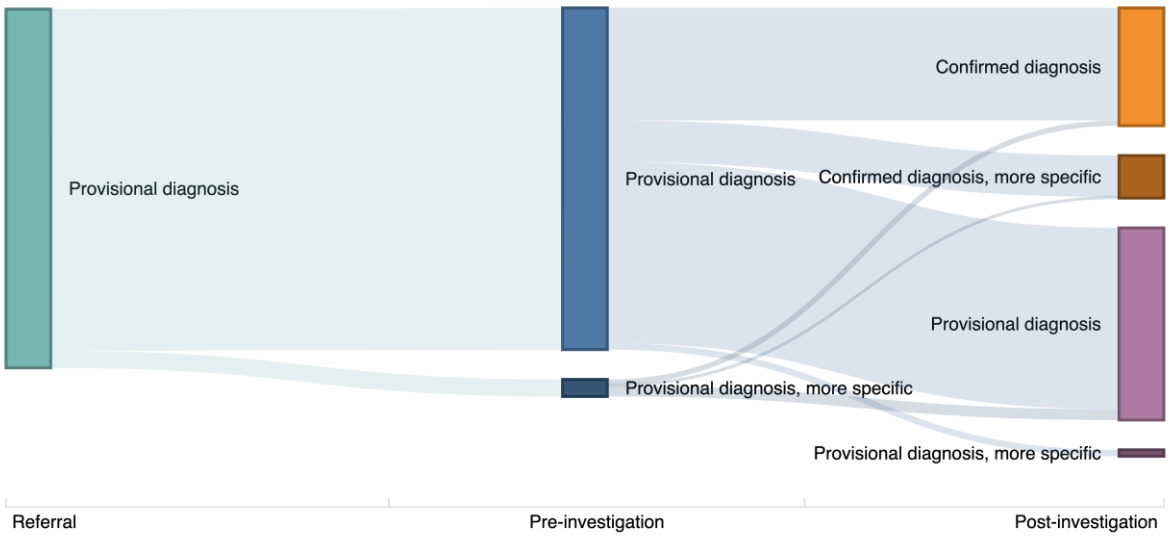


Figure 21: Pathling analysis of genomic testing data. Number of genomic test patients by encounter type, diagnosis verification status and diagnosis specificity relative to previous diagnoses (counts redacted).

We have recently released an open source version of Pathling (https://pathling.csiro.au) and are working with the HL7 community to support its use for analysing HL7 FHIR data.

Case Study 18: A FHIR-based Workflow Platform for Integrating AI into the Radiology Clinic

Barriers to wide adoption of AI based algorithms in radiology include an inability to easily integrate algorithms into the workflow, security and privacy, and the lack of explainability and provenance.

We are developing a web-based workflow platform called FORTE, for mapping standard operating procedures and clinical workflows into the Fast Healthcare Interoperability Resources (FHIR) framework. Each workflow in FORTE consists of steps which can be automated or manually performed. Furthermore, FORTE workflows are dynamic and rule based, which allows for the embedding of expert knowledge. For example, typical medical procedures can be coded such that subsequent steps (e.g. medical tests) are only allowed to be executed depending on output of previous steps unless vetoed by clinicians.

In FORTE, we leverage the capabilities of the FHIR framework which is a new standard for describing data elements and interfaces for exchanging electronic medical records. By utilising this model, we can standardise communications and integrate with hospital systems using FHIR. FORTE workflows can leverage off other workflows by exchanging information in FHIR.

An example of ordering a workflow for MRI based prostate radiotherapy is presented in Figure 22. The workflows consist of four steps, three of which are manual (white boxes) and one is automated (grey box). The automated step is described in detail in Case Study 28 and Box 6.

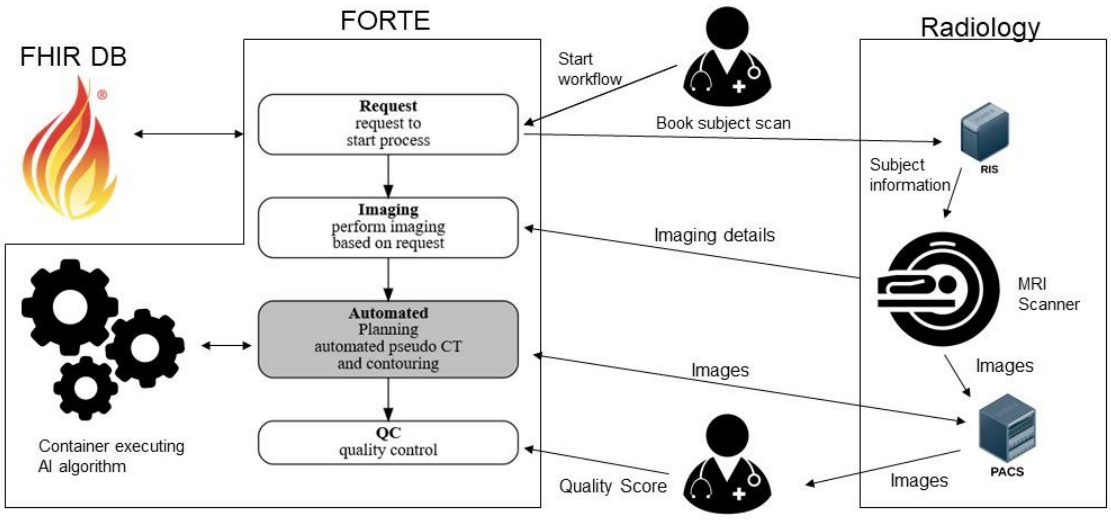


Figure 22: A workflow for performing MRI based prostate radiotherapy planning. The diagram shows the interactions between clinicians, FORTE and typical Radiology services.

FORTE overcomes many of the issues in bringing AI to the radiology clinic. Workflows are computer readable and described in FHIR, thus providing a high level of provenance throughout the workflow lifecycle. Since workflows integrate automatic and manual systems in a simple web-based platform, AI algorithms can easily be integrated into the clinical workflow. FORTE executes algorithms locally, thereby eliminating security and privacy issues associated with cloud-based implementations. However, FORTE can also support a cloud-based implementation. Currently, FORTE is in prototype stage as a proof-of-concept and we are continuing development for its use in radiology. We are also actively exploring its applicability in other areas of healthcare.

7 AEHRC Case Studies in Human Language Understanding

Several groups across the AEHRC develop solutions which use AI for human language understanding of health data. The data is typically in the form of unstructured human language, often electronic health records. While there has been a strong emphasis on standardised and structured records, many still exist as free-text and hence specific methods are needed to work with this data. Extracting and processing this data falls within the area of NLP, while searching such data is in the area of information retrieval. In addition, AI systems that aim to interact with users in a more natural, conversational manner fall within the area of conversational agents or 'chatbots'. We briefly review these three areas below before providing several Case Studies related to each.

Natural language processing

There are two main automated approaches to NLP: rule-based and ML based. Rule-based techniques use of pattern matching of words/phrases and dictionary lookup to extract specific concepts. Rules are typically developed manually, working with a domain expert and via manual review of the textual data being classified. ML approaches, on the other hand, have proven effective in a multitude of different NLP tasks. Key ML approaches for NLP include supervised ML techniques such as conventional classifiers (e.g. support vector machine, random forest, and naïve Bayes). Newer deep learning algorithms, which are based on artificial neural networks are now the state-of-the-art. Deep learning approaches also require less manual preparation of the data and exhibit more flexibility in addressing various automation tasks. Many NLP methods that use ML require large amounts of labelled data in order to train and validate a learning model.

Case Study 19 presents a good example of how rule and deep learning approaches can be combined to extract valuable information on cancer from a range of free text medical documents. Case Studies 20 and 21 show how machine-learning based NLP can be integrated into hospital workflow to detect missed limb fractures and to identify patients with antibiotic resistant infections. Case Study 23 shows how NLP can be used to automatically quantify the semantic similarity between sentences in medical literature for evidence-based medicine.

Information retrieval

Information retrieval is the scientific discipline behind search engine technology. In the ever increasing and diverse range of health data – from patient records to medical literature – effective methods to quickly and effectively search this data are needed.

Information retrieval applications typically use many other AI techniques – including probabilistic and statistical models and machine-learning. Key techniques often rely on other methods from NLP and knowledge representation. For example, using NLP to construct knowledge graphs that can be represented using SNOMED CT concepts. Modern search ranking methods are often machine-learning based. Case Study 22 demonstrates how a range of machine-learning based information retrieval methods can be used to help produce better systematic reviews.

Conversational agents

With the rise of social and communication technologies, conversational agents, or chatbots, provide a means for users to become engaged in conversation, continuing and progressing the dialogue in the same way human-to-human interaction occurs. They have more recently been used within the health paradigm for their ability, amongst other benefits, to autonomously gather key data about the user for analysis by clinicians. For example, monitoring speech degeneration in patients with Parkinson's Disease by asking the user to undertake a short speech assessment and forwarding a recording to the speech pathologist. Ultimately, this data can be used to tailor and improve the care clinicians provide.

Some examples where chatbots have been implemented previously include disease self-management, encouraging behaviour change, and provision of health education[6]. Case Study 24 presents a project to develop a chatbot to assist patients in decision making around the provision of genomic information.

Case Study 19: Automating Cancer Registry Tasks to Enhance Clinical Data Quality

Partners: Cancer Alliance Queensland

Information about cancers are gathered from a variety of different modalities – including imaging and from biopsy and resections – and then typically written into a narrative report for sending to the treating clinician. CSIRO has worked with Cancer Alliance Queensland to extract information from pathology and radiology reports and death certificates, using AI technologies, for a variety of reporting purposes – including cancer notifications, cancer staging and synoptic reporting.

The AEHRC Medtex technology uses a mix of symbolic and statistical AI methods to process the clinical reports. A natural language processing (NLP) engine is used to break the discourse of the text into statements and then features are extracted from each statement. The meaning of these features is then inferred through using ML models, which are trained from ground truth (human judgements) data using deep neural networks. For some features we use a formal logic rule-based approach using the relationships encoded in SNOMED CT (See Section 2.5).

Cancer Alliance Queensland uses Medtex to extract cancer information from new histopathology, radiology reports and death certificates and stores it in the Queensland Oncology Repository. This supports the clinical coding workflow to improve data collection capture within the Queensland Cancer Register (Figure 23). It also improved the multi-year delay in the reporting of cancers by providing more up-to-date population-level statistics on Queensland cancer incidence and mortality.

[6] Cited in Ireland, D. et al. (2016) Hello Harlie: Enabling Speech Monitoring Through Chat-Bot Conversations. Stud Health Technol Inform., 227:55-60.

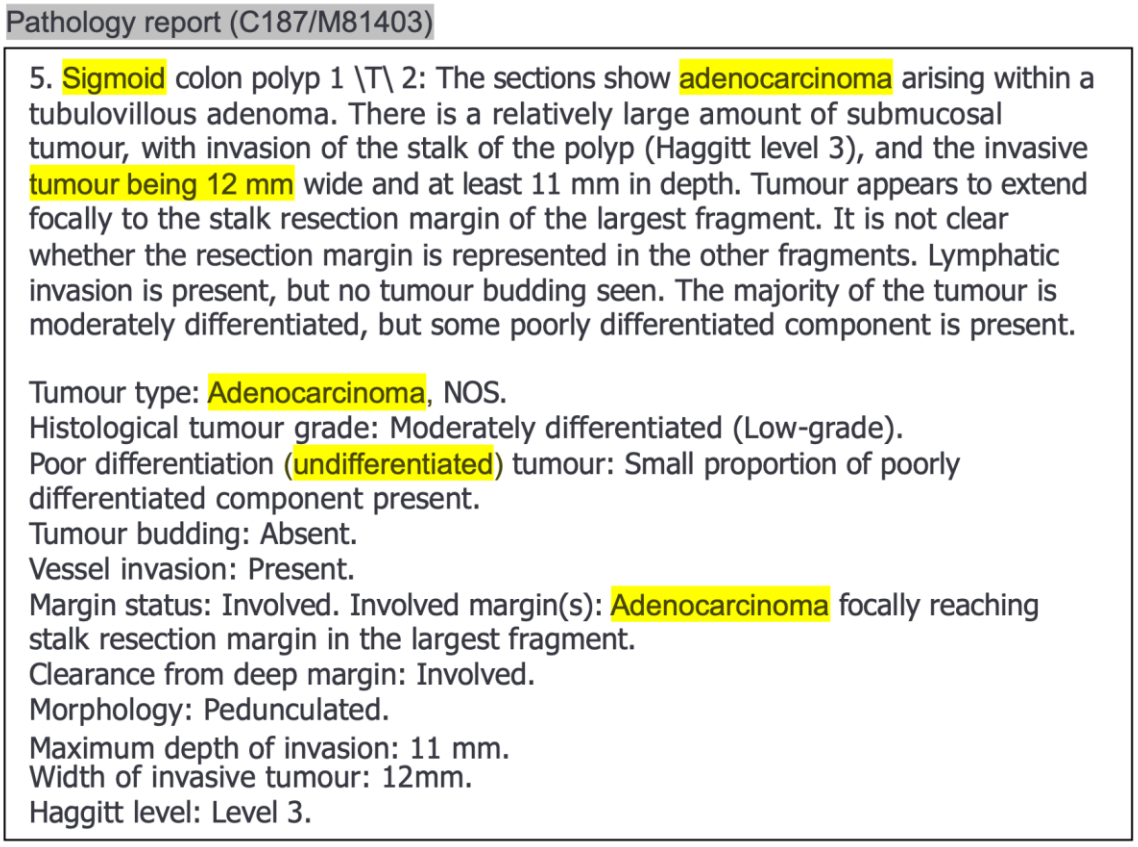


Figure 23: Medtex software processes narrative pathology reports and generates structured data with attentional mechanisms for explainability to aid clinical coders in cancer abstraction tasks (Source: Cancer Alliance Queensland).

The software now supports the extraction of over 20 different clinical features from the text of the histopathology reports covering a range of cancers. Studies have shown that the accuracy of the AI algorithms is very high. The algorithms have a 96% recall and precision for classifying notifiable cancers. Detailed extraction and coding of specific cancer notification items include basis of diagnosis, histological type and grade, primary site and laterality. Visual explanations and feedback from AI decisions are supporting clinical coders in their cancer abstraction task.

References: Case Study 19: Automating Cancer Registry Tasks to Enhance Clinical Data Quality, page 77.

Case Study 20: Checking Radiology Reports to Prevent Missed Fractures

Partners: Royal Brisbane and Women's Hospital and Gold Coast Hospital Department of Emergency Medicine

Patients admitted to a hospital emergency department (ED) with a suspected fracture are X-rayed, treated and then discharged. However, when the X-ray report is later finalised by a radiologist, ED specialists have to manually match the report from the radiologist with the patient's discharge diagnosis to ensure that subtle fractures were not missed. The manual checking process is an essential but laborious task that can occur up to 7 days after the patient is discharged.

Automation of this process can help ensure abnormal findings are not missed and that patients receive appropriate follow-up after discharge.

We used our Medtex system (See Case Study 19) to develop a solution that performs this check automatically and then flag any potential inconsistencies. The solution uses NLP to extract features from the reports. ML models including support vector machines and deep neural networks are then used to find associations between features in the radiology report. Standard clinical terminology (SNOMED CT) concepts are used as features to reliably identify limb fractures and other abnormalities documented in radiology reports (Figure 24).

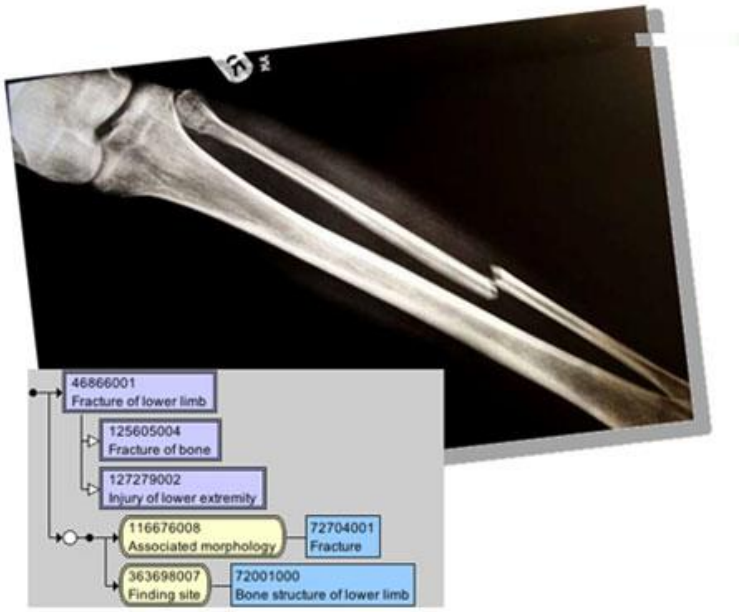


Figure 24: Medtex uses SNOMED CT, the internationally defined set of clinical terms, and other syntactic and semantic textual features from radiology reports to inform its ML models.

Our system automatically matches fractures identified in the radiology reports with patients' ED discharge diagnosis to provide decision support to the current manual checking process. Studies have shown that this checking can be done with high precision and recall across three different hospital ED settings. By fast-tracking diagnoses and streamlining test result reviews, emergency departments can save time and deliver improved patient outcomes.

References: Case Study 20: Checking Radiology Reports to Prevent Missed Fractures, page 78.

Case Study 21: Tackling Antimicrobial Resistance with Test Result Reviews

Partners: The Prince Charles Hospital Emergency Department and Pathology Queensland

Antibiotic overuse contributes to antimicrobial resistance, which could cost the global economy US$100 trillion by 2050 and cause up to 10 million deaths per year. Patients with suspected infections are tested for the presence of bacterial organisms with antibiotic resistance. These test results are then manually reviewed to ensure patient's infections are not resistant to the antibiotics they are taking. This project aims to automate this process in two parts: 1) streamline Emergency Department microbiology test result review to identify bacterial organisms and their antibiotic sensitivities; and 2) match these with antibiotic prescriptions extracted from Emergency Department discharge letters.

Our NLP methods extract antibiotic prescriptions detailed in discharge letters. Then we parse microbiology reports for bacterial organisms and antibiotic sensitivities. Given these two sources, we exploit the semantics in clinical terminology such as SNOMED CT to match antibiotic prescriptions (e.g. generic and trade names) with the bacteria's sensitivities for a given antibiotic class. This provides clinical decision support to identify patients that have been prescribed an antibiotic for which the bacterial organisms are resistant. The patient can then be contacted for follow-up treatment, such as a change of antibiotic treatment (Figure 25).

An example scenario is where the discharge letter noted that a patient was prescribed with "ampicillin". When the microbiology test result returns, it notes the bacteria present was "Escherichia-coli" (E. coli): a bacterium known to be resistant to ampicillin. Our system would pick this up immediately and alert the clinician, enabling the patient to be contacted and provided with a more appropriate antibiotic.

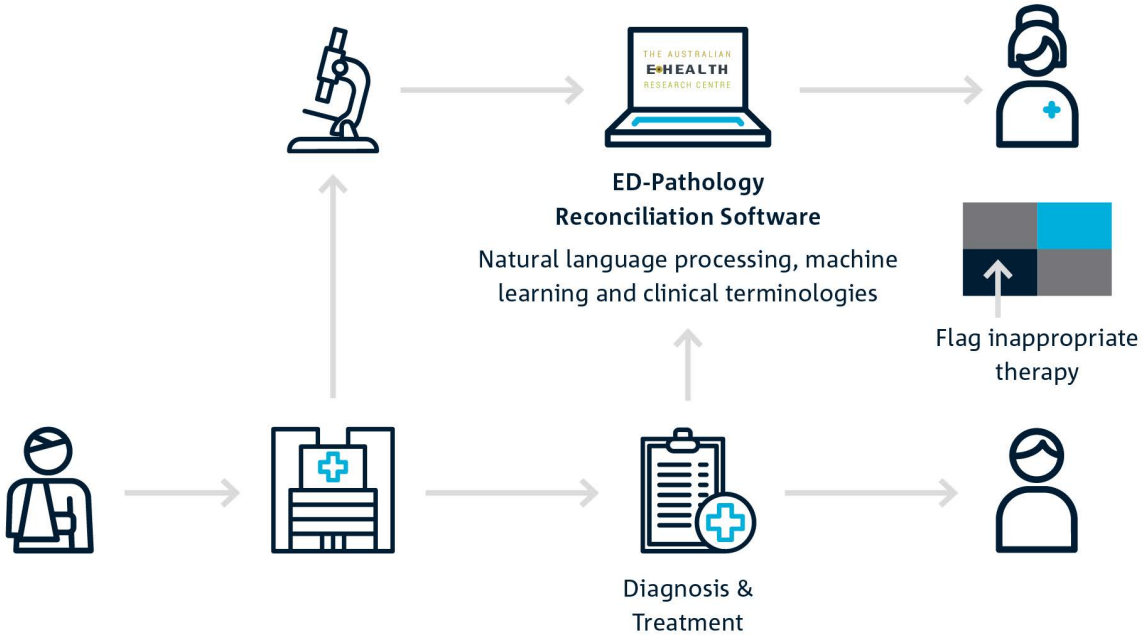


Figure 25: Clinical decision support workflow for streamlining the Emergency Department microbiology test result review to ensure important diagnoses are recognised and correct antibiotics are prescribed.

The system supports clinicians with faster and more accurate test result reviews, improving patient outcomes and helping reduce the resourcing strain on hospital staffing. This also contributes to the hospital Antimicrobial Stewardship program to ensure that antimicrobials do not become ineffective and are available for future generations.

References: Case Study 21: Tackling Antimicrobial Resistance with Test Result Reviews, page 78.

Case Study 22: Search Engines for Compiling Systematic Reviews

Partners: NHMRC Centre for Research Excellence in Digital Health, Bond University, Cochrane

Systematic reviews synthesise the work of primary research studies to answer a specific research questions; they are the key evidence for clinical guidelines and health policy. However, systematic reviews are time-consuming (>2 years), labour-intensive (>1,100 hours), and expensive (>$350K). Most of the costs arise from the lengthy screening phase, whereby large numbers of results are retrieved according to a search query and manually reviewed (See Figure 26). How people formulate their queries determines how many results they need to screen; but good queries are difficult to formulate, even for specialist librarians.

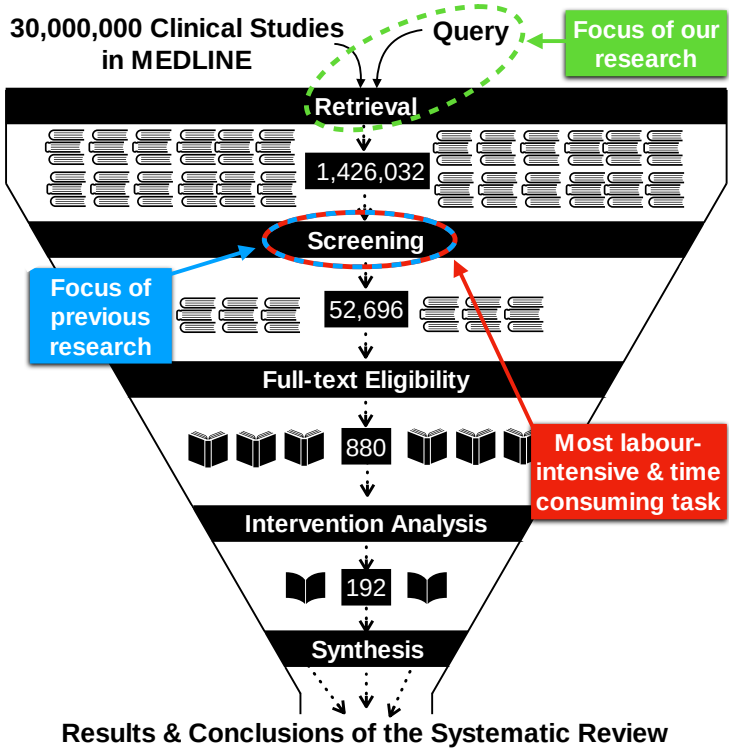


Figure 26: Phases of systematic review production. Query formulation determines how many results are retrieved and consequently require manual review. This project aims to help people formulate effective queries.

This project aims to devise AI-based query formulation methods, helping people write and understand better search queries. A number of AI-based methods were developed:

- Given the research question for a systematic review, a Query Formulation model uses NLP to build an effective Boolean search query.
- ML based Query Performance Prediction model that provides a quantitative assessment for the quality of a query. This measure of quality can either be provided directly to the user or guide the automated method below.
- A Query Chain Transformation model to help people formulate better queries. This model takes a searcher's initial queries and applies a series of semantic transformations such as adding new search terms, changing the Boolean clauses and broadening or narrowing the search scope. Queries can be presented to the user for explainability or automatically applied.

Empirical evaluation of our methods for improving Boolean query formulation shows that we can significantly reduce the number of non-relevant search results that need to be screened, and that this is not at the expense of missing any relevant studies. Furthermore, we provide more objective measures that reduce bias in how the search is done and hence in the conclusion of the review.

This software is now being used by our collaborators in the NHMRC Centre for Research Excellence in Digital Health and Bond University as part of an increasingly AI driven process to reduce the time it takes to develop a systematic review.

References: Case Study 22: Searching Engines for Compiling Systematic Reviews, page 78.

Case Study 23: Linking Clinical Evidence in Medical Literature for Evidence Based Medicine

Partners: University of Melbourne, Macquarie University, NHMRC Centre of Research Excellence in Digital Health

Published clinical trials and high-quality peer reviewed medical publications are considered as the main sources of evidence used for synthesizing systematic reviews or practicing evidence-based medicine. However, finding all relevant published evidence for a particular medical case is a time- and labour-intensive task. In particular, finding the relationships and linking different pieces of evidence is a real challenge.

This project aims to aid researchers and clinicians in linking and quantifying the strength of evidence from medical literature. A generalisable AI approach was developed to quantify semantic similarity between biomedical sentences. This was done for different classes of evidence, including clinical interventions, population information, and clinical findings. An example of sentence similarity between two articles' abstract is shown in Figure 27.

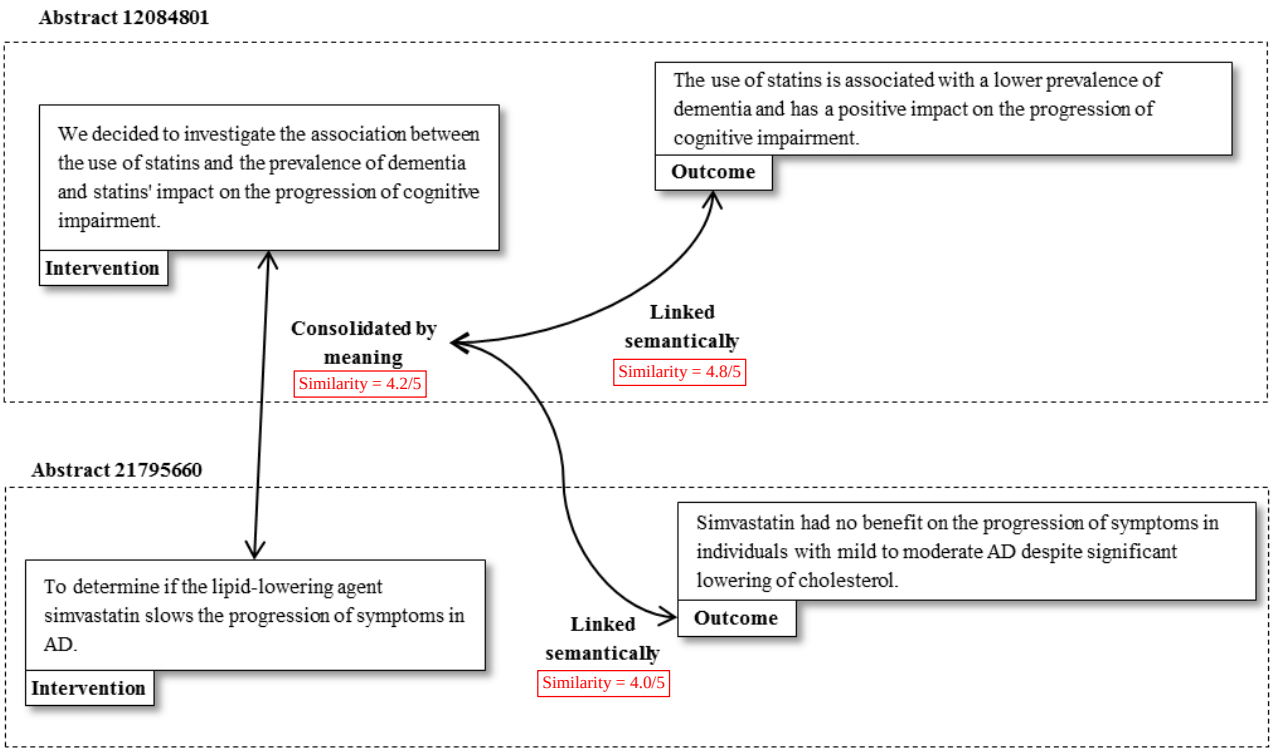


Figure 27: Clinical evidence relationships from two sample abstracts. Both abstracts have conceptually similar Interventions (statins and simvastatin refer to the same medical concept) but have almost contradictory Outcomes. Predicted similarity scores range from 1 (low) to 5 (high).

A series of generic and domain-specific semantic similarity measures, including ontology-based and vector-space models of similarity measures, were used to quantify the similarity of sentences containing clinical evidence. They make use of:

- lexical similarity derived from terms, syntactic part of speech, and other syntactic and lexical elements,
- conceptual similarity based on medical terms or concepts and their corresponding nodes in the ontology graph, and
- contextual information based on concept embeddings derived from unsupervised neural network-based vector representation of medical concepts.

All the above methods are combined into the single workflow (Figure 28) to generate an overall similarity score.

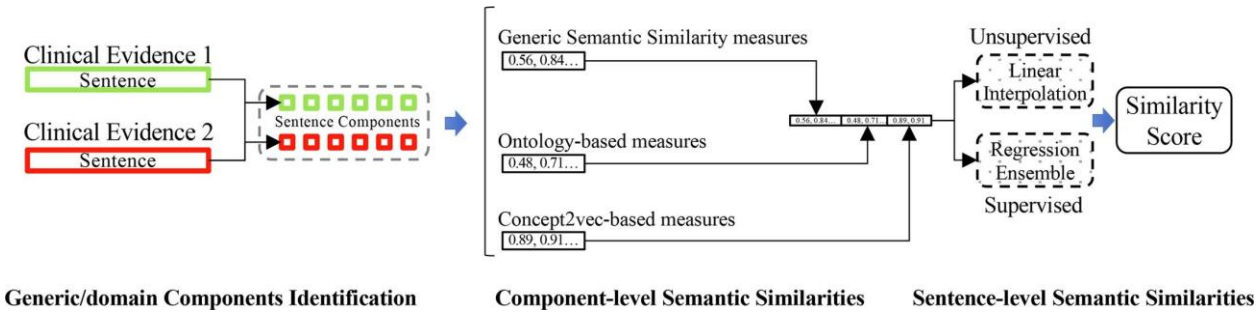


Figure 28: Process of quantifying the semantic similarity between two sentences. Generic and domain-specific methods are combined to build an overall measure of similarity.

Human assessors reviewed 1000 pairs of sentences from biomedical publications to provide a gold standard reference set. Comparing human judgements with our automated approach showed a 0.80 (80%) Pearson correlation on five different clinical evidence types. The approach also showed high effectiveness and generalisability when evaluated on an external dataset.

Our automated linking method is aimed to expediate the synthesis of medical literature as part of evidence-based medicine and the production of systematic reviews. It allows for the automatic synthesis of studies to understand the similarity between study cohorts, how interventions for addressing a particular medical condition differs from one study to another, and which group of reported findings support the effectiveness of a particular intervention. It underpins the knowledge used by health professionals in clinical decision-making and researchers in producing systematic reviews.

References: Case Study 23: Quantifying Semantic Similarity of Clinical Evidence in the Biomedical Literature, page 78.

Case Study 24: Edna, the Chatbot Designed to Assist with Decision Making for Provision of Additional Findings Information.

Partners: Melbourne Genomics Health Alliance

Genomic healthcare is a rapidly advancing field - with patients undergoing testing to determine if their condition has a genetic cause. However, genomic sequencing creates data on thousands of genes, not only those known to cause the patient's presenting condition. Disease-causing gene variants unrelated to the medical indication for a genomic sequencing test are referred to as Additional Findings. A subset of these genes cause conditions that are preventable or treatable, and it has been argued that patients should be able to access this information if they choose.

Many patients are interested in having their genomic data reanalysed so they can receive additional findings, but they must make an informed decision on whether they want the information. Conversational agents, or 'chatbots', which interact with the user via speech or text, provide a good means to assist patients with this decision-making process, as these digital health options minimise the impact on health services. Chatbots can be accessed at any time, at home, leading patients (or families) through topics pertinent for informed decision making and answering general questions.

To support informed decision making about reanalysis of genomic data to ascertain future risk of preventable or treatable conditions, we developed a genomics chatbot, Edna (eDNA), to

complement genetic counselling. Edna uses AI to both provide and collect information (Figure 29).

Edna uses two main AI approaches: case-based and symbolic logic reasoning. The case-based reasoner provides most of the responses using two main algorithms, a syntactic matching algorithm and a sentiment analysis algorithm. The syntactic matching algorithm analyses the structure of the written or spoken data (technically termed an 'utterance') to recognise words (including their arrangement) and compute responses. Once a user provides their name, for example, Edna will respond with "*Nice to speak with you [name]*". The sentiment analysis algorithm identifies the probability of negative sentiment. Multiple responses are encoded, with each specific response chosen based on the sentiment of the last utterance and overall conversation. This function is important to alert Edna that interaction with a human may be required.

The logic reasoner acts on the computed semantics using a large lexicon database consisting of word types, synonyms and antonyms. The logic reasoner converts natural language to the language of predicate logic (simply put, converting words to mathematical equations that can be understood and manipulated by the computer). This allows representation of natural language data, detection of logical contradictions, and response to queries such as "*What genetic conditions, available through additional findings analysis, have high prevalence*?"

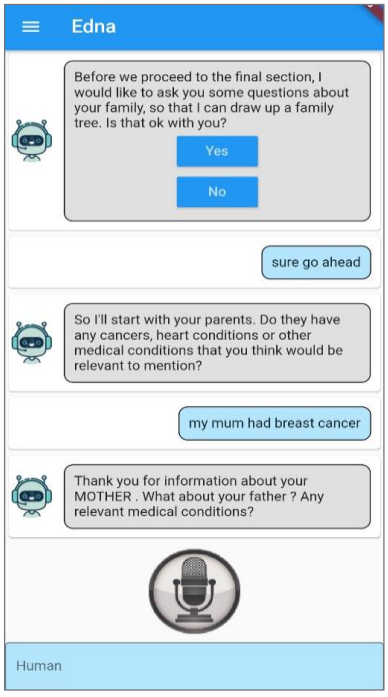


Figure 29: An interaction with Edna around provision of family history information.

Edna is programmed to converse on the many facets of additional finding such as the advantages and disadvantages of testing, mental health, insurance and family implications. Genetic conditions, associated medical terms and implications of testing can be explained to family and friends, and family history can be collected. At any point the user can opt to be contacted by a genetic counsellor. We are currently trialling Edna with patients who have had genomic sequencing.

8 AEHRC Case Studies in Imaging and Vision

A number of groups across the AEHRC develop solutions which use AI in medical imaging and vision.

Medical image analysis

Medical image analysis employs a range of supervised and unsupervised AI and ML techniques to extract clinically relevant information or knowledge from medical images. This may involve tasks such as image enhancement, feature detection, tissue segmentation, registration and normalisation, shape appearance and relationship modelling, classification and clustering, comparisons of groups, and investigation of longitudinal changes. These tasks are generally fully automatic and involve a combination of statistical, machine and deep learning methods.

Using medical resonance imaging for early detection of abnormal development

One of the key uses of medical imaging is the early detection of disease. Early detection results in early intervention, which can have lifechanging outcomes. AI-based image processing is one area that can provide significant capabilities for early detection, especially if it is well integrated into often complex clinical workflows. Case Study 25 presents a cloud-based 'Developing Brains' toolbox using ML to analyse MRI scans of very preterm-born infants to identify biomarkers that predict later motor, neurological and neurobehavioral problems. Case Study 26 describes a clinical analysis tool, AssessCP, the first fully automated web based clinical support tool for paediatric brain injury.

Image guided treatment and disease monitoring

AI-based imaging technology can also help guide treatments. Case Study 27 shows how software that integrates with MRI machines can be used to obtain fast, accurate and reproducible quantification of the changes in cartilage indicating osteoarthritis – this guides surgery such as joint replacements. Case Study 28 uses MRI images to help guide the delivery of radiotherapy for prostate cancer; the software that integrates with existing radiation oncology workflows. This is a complex and multi-faceted AI approach with further details on the techniques outlined in Box 6.

Case Study 29 uses PET imaging (often easier to obtain than MRI images) to generate quantified measures of risk of Alzheimer's Disease. In Case Study 30, deep learning methods are used on ocular images for automated detection of macular degeneration that can cause blindness. Similarly, Case Study 31 uses image processing for segmentation of flecks in the eyes to track Stargardt disease progression. These latter cases can also rapidly identify at risk patients for early intervention.

AI-based tele-health

An important part of tele-health is to increase early diagnosis and timely treatment of a range of diseases. This can be particularly true for regional and rural areas where specialists are limited to metropolitan hospitals. Case Study 32 presents a tele-oral care system that provides AI-driven oral mucosal disease classification and specialist-based clinical decision support.

A drawback of tele-health is the lack to human-to-human cues of emotion and reaction that are lost over limited bandwidth communication channels. Case Study 33 aims to overcome some of this through face detection and automated classification of emotion from video for tele-health. This provides a means for clinicians to obtain feedback on a patient's emotions and reactions even when they are not in the room.

Finally, tele-health relies on appropriate platforms to support communication and information sharing. Box 7 presents a platform to support the collection, screening and forwarding of medical images from remote locations to specialist to support tele-health.

Robotics for health

The use of robots is set to have a big impact on the way medicine is delivered. Robotics in health are often thought of as specialised robots performing complex operations such as surgeries. But there are many other ways robots can help. Case Study 34 shows how socially-assistive robots have been proposed to supplement traditional therapy and education for children with autism. In this Case Study, AI is used to train the robot how to behave during a conversation, by 'observing' how people interact with each other.

Case Study 25: Early Prediction of Neurodevelopmental Outcomes of Preterm Infants

Partners: Queensland Cerebral Palsy Rehabilitation and Research Centre

Events occurring during prenatal development and preterm birth can impact on vital brain development processes. This can lead to reduced cognitive and motor abilities in infants – and a diagnosis such as Cerebral Palsy (CP) or Autism Spectrum Disorder (ASD).

We are developing the cloud-based Developing Brains toolbox, which uses ML methods to analyse MRI scans of very preterm-born infants and provides imaging biomarkers of brain structure and microstructure. ML algorithms use these biomarkers to predict later motor, neurological and neurobehavioral function, helping to guide clinical decisions around early interventions.

The neonatal MRI analysis includes i) a ML classifier to remove motion artefacts in the MRIs, ii) atlas registration to perform brain tissue segmentation and labelling, iii) calculation of diffusion tensor metrics for assessing local white matter microstructure, and iv) constrained spherical deconvolution to perform white matter tractography. Deep learning and logistic regression are used to elucidate which biomarkers are the most predictive of patient outcomes. Using a large clinical cohort, deep learning is used to learn the relationship between hundreds of biomarkers and 2-year clinical outcomes. This can then predict each individual's 2 year developmental "fingerprint", their estimated brain age, and the likelihood of them developing CP or ASD (Figure 30).

Using deep learning, we can identify with high sensitivity, specificity and accuracy, infants at a risk of developing motor impairment from their early brain scans. Motor cortex and somatosensory regions were identified to have the highest association with later adverse motor outcome, which is consistent with the known functional roles of these areas.

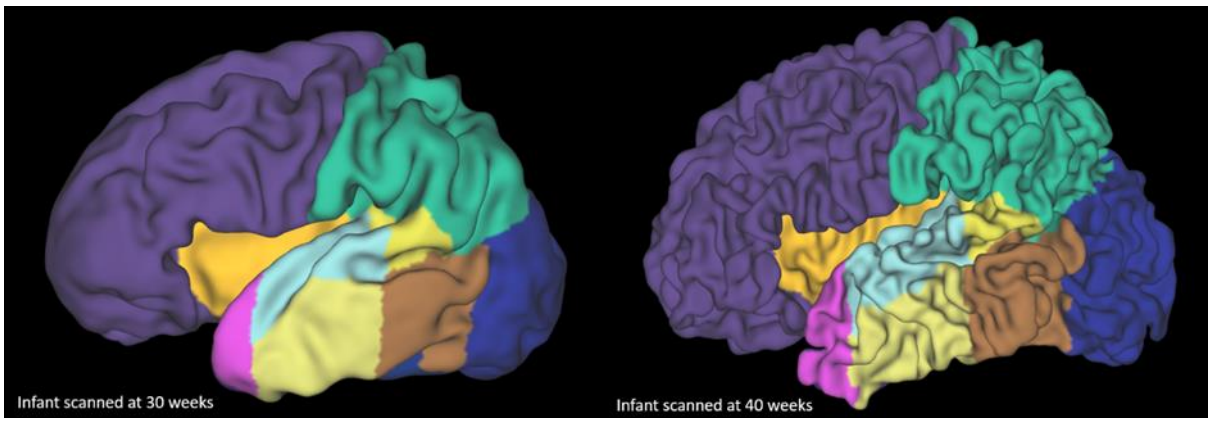


Figure 30: A 3D rendering of the cortical surface, illustrating the rapid development occurring between MRIs acquired at 30 weeks post menstrual age (left) and at term equivalent age (right).

The tool is currently used in research studies by our partners and has future application in supporting clinical decision making.

References: Case Study 25: Early Prediction of Neurodevelopmental Outcomes of Preterm Infants, page 78.

Case Study 26: Clinical Decision Support Tool for Children with Neurodevelopmental Disorders

Partners: Queensland Cerebral Palsy Rehabilitation and Research Centre, Monash University and the Phoenix Children's Hospital.

Cerebral Palsy (CP), Acquired Brain Injury (ABI) and Autism Spectrum Disorder (ASD) affect 1% of Australian children, however, the quantitative link between injury and health outcomes for these conditions remains poorly understood. While advanced structural imaging, typically MRI, is performed to assess the extent of brain injury, and guide physical therapy, current methods for quantifying injury have been limited by the inability to model severe pathology.

We are developing a clinical analysis tool, AssessCP, to be the first fully automated web based clinical support tool for paediatric brain injury.

The neuropathology of at-risk infants is often too severe to use traditional brain atlases when processing the infants' MRI scans, as current atlases are developed from images of healthy brains. To obtain accurate measures of brain injury in infants and older children, AssessCP uses ML approaches such as Expectation Maximisation (EM) to cluster image intensities. This allows segmentations of brain tissue and lesions, and measurements of cortical morphology, to be accurate and robust to brain injury. The tool generates an automated report quantitatively detailing brain structure (such as is shown in Figure 31), which can then be compared to an age-matched typically developing cohort. Clinicians can use the report to identify and visualise brain injury and estimate patient risk of long-term impairments; while researchers can use the report to understand how the brain adapts in response to both injury and treatment.

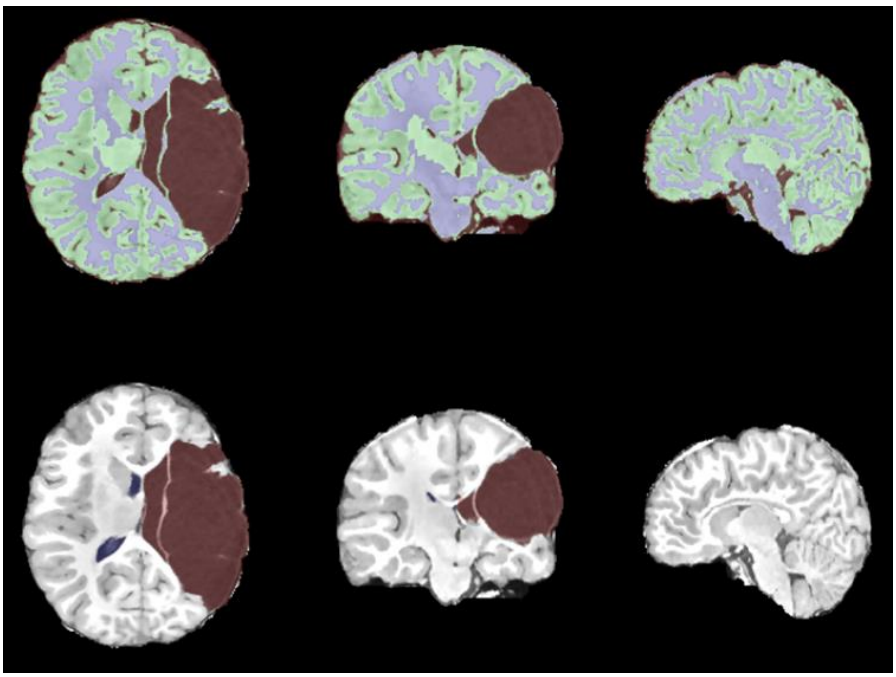


Figure 31: Structural brain segmentation of a child with Cerebral Palsy, presented in the AssessCP clinical report.

AssessCP has been validated on cohorts of children with CP and ABI and is currently being evaluated on children with ASD. AssessCP has clinical potential in quantifying neuroplasticity during evidence-based interventions and explore the association between structural changes and genetic markers.

References: Case Study 26: Clinical Decision Support Tool for Children, page 78.

Case Study 27: AI for the Diagnosis of Musculoskeletal Disorders

Partners: University of Queensland, Siemens Healthineers

Osteoarthritis (OA) is a major cause of chronic pain and disability and involves progressive degenerative changes in the anatomical structure of articular cartilage. Magnetic resonance imaging (MRI) provides excellent visualization of articular cartilage (Figure 32) along with other joint structures. This has generated extensive clinical interest in using new technologies to provide quantitative analyses of joint structures. This information is important to guide early diagnostic and management options for OA and other significant joint pathologies.

For MRI-based joint analyses to be clinically viable, it is essential to obtain fast, accurate and reproducible quantification of the morphometric and biochemical characteristics of cartilage, bone and surrounding structures.

To achieve this, AEHRC and the University of Queensland have jointly developed software (ChondralHealth) that integrates with existing Siemens MRI scanners. This involved the development of state-of-the art algorithms that model the shape and appearance of joints in the body. ML based algorithms are then used to objectively measure the health and condition of cartilage tissue as seen from MR images of the knee, shoulder, spine and hip. The algorithms used include atlas-based segmentation and deformable models – where a dictionary of example joints is used to match data or models to a new MR image of a patient. The deformable models utilize

learned anatomical and appearance information in a training process, which provides an understanding of the shape and appearance variation in the population.

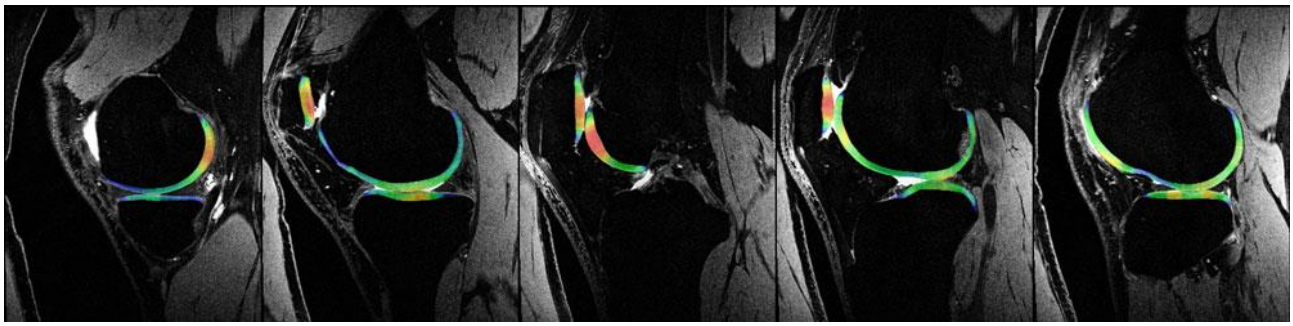


Figure 32: Sagittal slices of MRI Knee scan with cartilage segmentation colour coded using cartilage thickness.

AEHRC is working with Siemens to incorporate ChondralHealth into their next generation of MRI scanners. To further improve ChondralHealth, current research has involved developing ML methods to improve quantification of lesions and injuries, while deep learning methods have been used to improve the resolution and image quality of MRI scans.

References: Case Study 27: AI for the Diagnosis of Musculoskeletal Disorders, page 79.

Case Study 28: AI Driven Prostate Radiation Therapy

Partner: Calvary Mater Newcastle Hospital

Modern image-guided external beam radiation therapy relies on computed tomography (CT) scans to predict radiation absorption and scattering during treatment delivery. However, MRI is considerably better than CT at identifying soft tissue boundaries - reducing side effects including damage to normal tissues such as the bladder and rectum which can lead to rectal bleeding, incontinence, and erectile problems.

AEHRC has developed software that integrates with existing radiation oncology workflows and automatically generates CT information and organ boundaries from standard MRI scans (Figure 33). This removes the need for both CT and MRI scans and potential errors from combining the two scans. This will reduce side effects and increase the efficiency of prostate radiotherapy. The software uses a suite of AI techniques to solve different tasks at each step through the process – the process for developing this algorithm has been explained in full in Box 6.

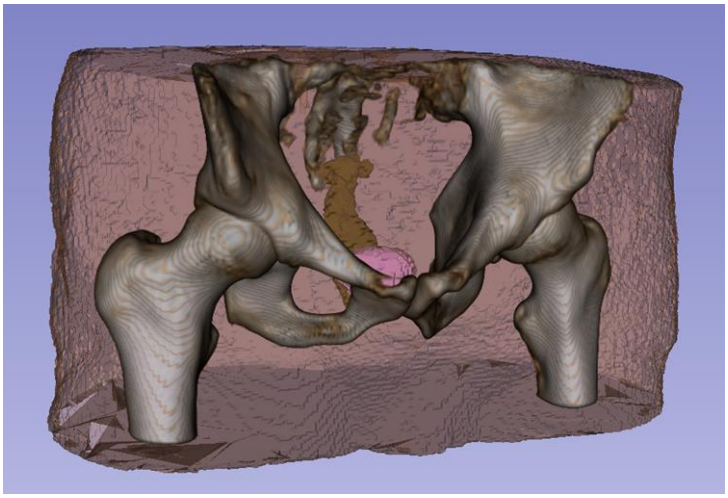


Figure 33: 3D rendered view showing automatically segmented skin, bones, rectum and prostate from an MRI of a patient with localised prostate cancer.

References: Case Study 28: AI Driven Prostate Radiation Therapy, page 79.

Box 6: From Image to Radiation Therapy Plan: AI steps

A number of AI and ML techniques are involved in this project to generate CT scans from MRI for prostate cancer radiation therapy. Here we give a detailed description of the process of implementing the algorithm and the different AI techniques used. Full details are given in the papers in the publication list for Case Study 28.

1. Initially 3D multi-modal structure-guided image registration (CT and MRI) is used to develop an offline training atlas set from a range of patient treatment images.
2. Image registration is an AI/computer vision method that involves mapping information from one image to another. It is an iterative process where one image is distorted to match another (in this case using free form deformation) and a metric between the two images (in this case normalised mutual information) is minimised (usually through a gradient descent optimiser). In this work there can be large differences in bladder and rectum filling between the CT and MRI scans, and we developed methods to accurately map these organs without distorting rigid structures (such as bone).
3. The main conversion method used in this work is based on multi-atlas fusion, where each MRI from the offline atlas set is registered to a new patient's MRI (using the diffeomorphic demons algorithm). Each of these registrations output a rigid transform and a deformation field. These two files can be used to propagate structure labels and co-registered CT-MR information to the new MRI.
4. Once this is performed a patch based local weighted voting procedure is applied where the best matching patches from the registered MRI atlases are calculated and used to select the corresponding Hounsfield units from the propagated co-registered CT-MR images. The same weightings can also be applied to generate automatic organ segmentations from the binary structure labels.
5. We have also investigated ML and deep learning approaches to generating CT information, and these methods can result in improved execution time with similar results. However, an advantage of the registration-based approaches is that they produce explainable results and work with a smaller training set (which are more common in radiation oncology clinical trials). We have also developed methods for segmenting pelvic organs using active appearance models, which extend the registration atlas approach to generate shape and appearance priors to initialise a local search for organ boundaries.

Case Study 29: Automated Analysis of PET Images for Neurodegenerative Diseases

Partner: Austin Health, The Australian Imaging, Biomarker & Lifestyle Study of Ageing

Alzheimer's disease is characterized by an accumulation of Amyloid and Tau proteins years before the onset of the disease. While novel PET tracers can measure these proteins *in vivo,* early detection can be difficult as subtle changes can often only be recognised by expertly trained doctors. To aid quantification, an MRI scan is typically aligned to the PET scan and used to measure the Amyloid and Tau burden in key areas of the brain. However, an MRI is not always available and can be problematic to acquire in older and less compliant populations.

CSIRO has developed a novel PET-only quantification method which does not require a matching MRI to obtain accurate quantification. This technology, which we refer to as CapAIBL (Computational Analysis of PET for the Australian Imaging, Biomarker & Lifestyle Study of Ageing), allows us to generate accurate 3D rendering of the PET retention (Figure 34), and produces clinician-friendly reports (Figure 35).

An adaptive atlas, which can dynamically simulate the pattern of PET tracer retention of a new image, was developed to aid the spatial normalisation to a standard template of the brain. For each subject, the parameters of the adaptive atlas are optimised to best match the target image. The image is then non-rigidly aligned to the atlas. Advanced pattern matching and ML algorithms are then used to map the PET tracer retention onto a template brain surface. Since the PET image lacks the fine anatomical details required to build a surface mesh, we instead rely on a template surface mesh and a database of pairs of PET images (atlases) with their corresponding surface mesh built from MRI. For each vertex on the target mesh, a local patch centred on that point is extracted on the PET image and the PET atlases. The atlas patches that are the most similar to the PET image patch are selected. The local PET uptake on their corresponding surface mesh are then weighted in a Bayesian fusion framework and the result mapped onto the template cortical surface mesh. This is repeated for all vertices on the template mesh.

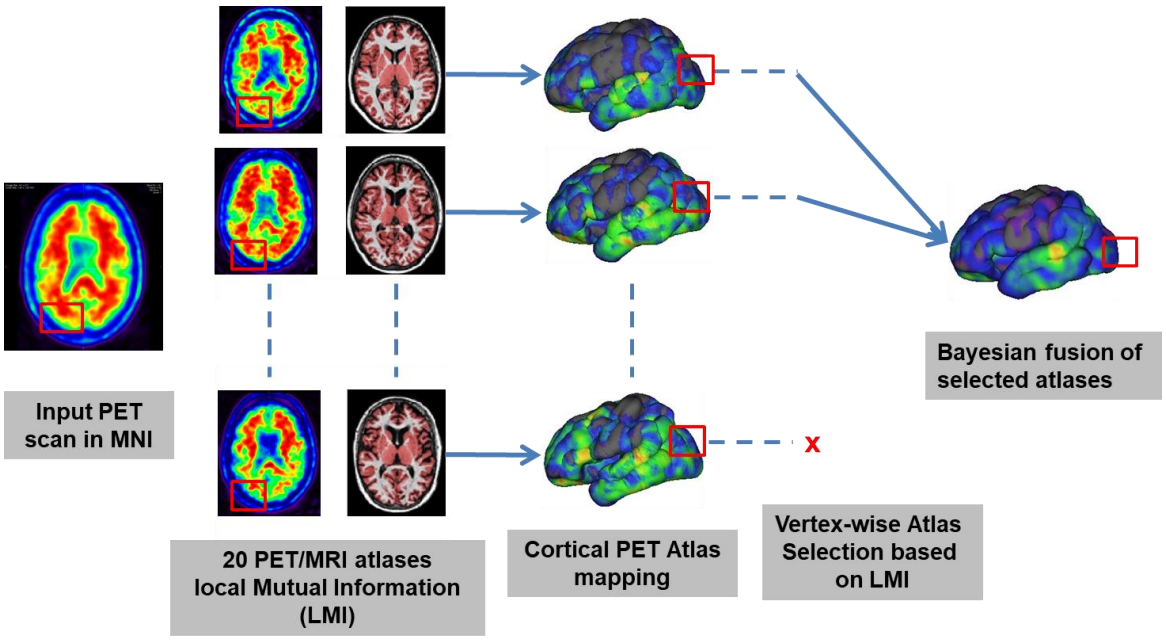


Figure 34: Mapping of the PET uptake onto a surface mesh without a structural MRI.

CapAIBL has been extensively validated (with our clinical partners) against clinical expert reading and histology and has also been calibrated to produce Centiloids (a measure of amyloid protein). While originally developed solely for amyloid imaging, CapAIBL has more recently been extended to FDG and tau imaging. This technology is patented and undergoing commercialization.

References: Case Study 29: Automated Analysis of PET Images for Neurodegenerative Diseases, page 80.

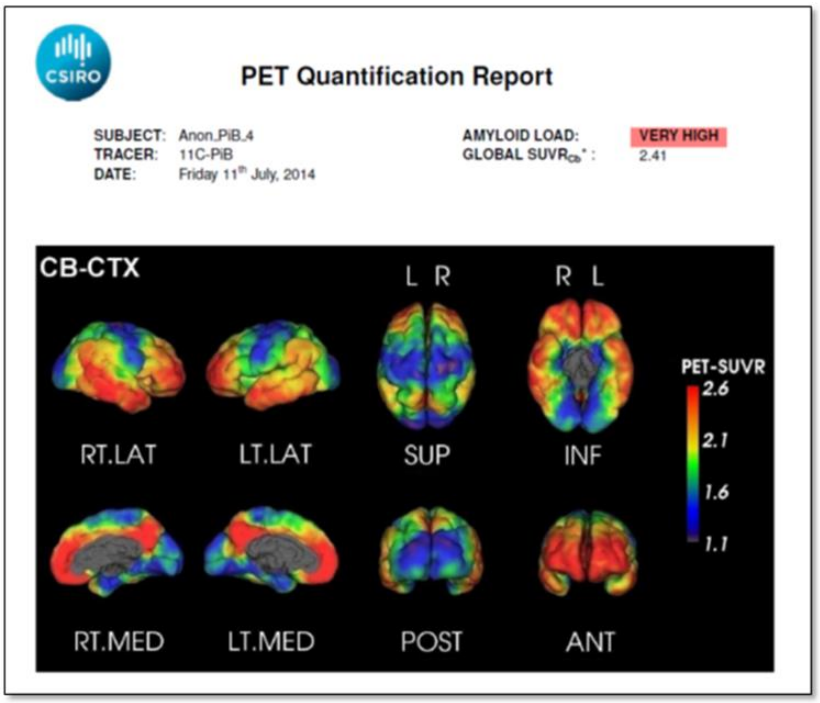


Figure 35: A PET Quantification report showing Amyloid load on the brain surface.

Box 7: Store and Forward Tele-health with AI Assistive Technologies

Australia's sparse population provides many challenges to the health system, as well as many opportunities for tele-health to provide both economic and health improvements.

We have developed two technologies to support tele-health services which need to capture and exchange various images of the patient. Our Remote-I technology is a cloud-based system to support screening programs through a "store and forward" mechanism. The Remote-I platform provides a user-friendly interface for clinicians, nurses and health workers to take an image of the patient, store it locally and then forward it to a specialist for review. Remote-I has been used extensively for tele-ophthalmology to support screening for diseases such as Diabetic Retinopathy.

The second technology is our Medical Imaging Communication & Exchange (MICE) platform. This is a mobile phone app-based platform for the remote monitoring and management of wound, burns and electronic consent capture processes.

Since both platforms capture images to forward to specialists for review, they provide an ideal opportunity for AI algorithms to do an initial review to support the clinical workflow – for instance to flag cases for urgent review. These AI algorithms use various types of AI and Deep Learning to support rapid and accurate image interpretation and improved health system workflow. Clinical applications include ophthalmology, diabetes, hypertension, and neuro-degenerative diseases such as Alzheimer's and Stroke. The goal is to improve health outcomes and produce cost savings to the healthcare system.

See Case Studies 30, 31 and 32 for examples of these types of AI and tele-health applications.

Case Study 30: Automated Detection and Classification of Early Age-related Macular Degeneration Biomarkers

Partners: Doheny Eye Institute, University of California, Los Angeles, USA

Age-related macular degeneration (AMD) affects millions of people and is a leading cause of blindness throughout the world. Ideally, affected individuals would be identified at an early stage before further progression produces irreversible loss of vision. However, accurate and precise staging of AMD, particularly using newer optical coherence tomography (OCT)-based biomarkers is time-intensive and requires expert training, which is not feasible in many circumstances including in screening settings.

CSIRO and our partner have developed a method for the automated detection and classification of an early AMD OCT biomarker.

Deep learning models (convolutional neural networks, CNN), were trained for automated detection and classification of hyperreflective foci (Figure 36), hyporeflective foci within the drusen, and subretinal drusenoid deposits from OCT B-scans.

For each of the pathology types we trained a different CNN, namely Inception-v3, ResNet50, and InceptionresNet50. Pre-segmentation of the retinal layers was performed using the ReLayNet CNN and then 11 different CNN training approaches were tested, each fine-tuning different proportions of the CNN layers. A total of 19584 Spectral domain (SD)-OCT B-scans of 153 patients, who were diagnosed with early or intermediate AMD in at least one eye, were used for training and testing the CNNs (random 10% selected for testing).

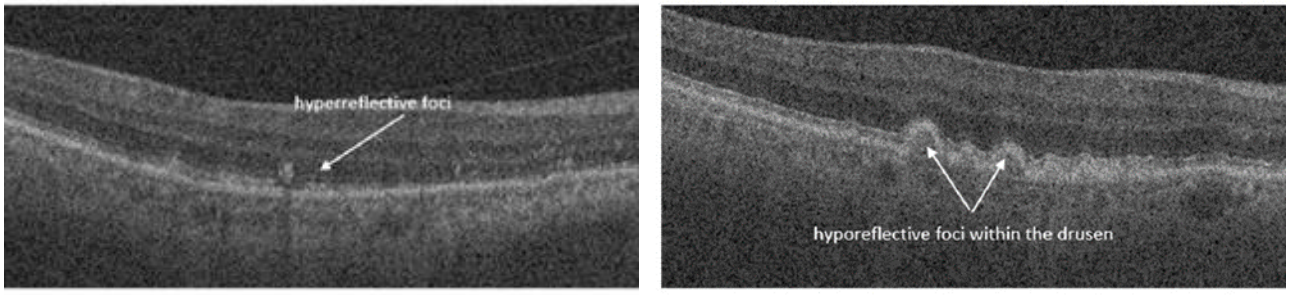


Figure 36: Examples of hyperreflective foci (A) and hyporeflective foci in drusen (B), from optical coherence tomography imaging.

An overall accuracy of 87% for identifying the presence of early AMD biomarkers was achieved. Given the increasing burden of AMD on the healthcare system, the proposed automated system is highly likely to perform a vital role in decision support systems for patient management and in population and primary care-based screening approaches for AMD.

Case Study 31: Segmentation of Hyperautofluorescent Fleck Lesions in Stargardt Disease

Partners: Lions Eye Institute

Stargardt disease is the most common form of inherited retinal disease that leads to permanent vision loss. A diagnostic feature of the disease is flecks, which appear as bright fluorescent spots in fundus autofluorescence (FAF) imaging. The size and number of these flecks increase with disease progression. Manual segmentation of flecks allows monitoring of disease but is time-consuming.

CSIRO has developed and validated a deep learning approach for segmenting these Stargardt flecks and tracking disease progress through longitudinal images. A U-Net convolutional neural network architecture was constructed with a residual neural network (ResNet) encoder for the deep learning fleck segmentation. The number of flecks and total area of flecks are calculated and compared longitudinally. Outputs were compared to expert clinician manual grading (Figure 37). The data included 1750 training and 100 validation FAF patches from 37 eyes with Stargardt disease.

Testing was done in 10 separate Stargardt FAF images and we observed a good overall agreement between manual counting and deep learning in both fleck count (mean difference 0.6 spots) and fleck area (mean difference -0.03 mm^2). Longitudinal data were available in both eyes from six patients (average total follow-up time 4.2 y), with both manual and deep learning segmentation performed on all (n = 82) images. Both methods detected a similar upward trend in fleck number and area over time.

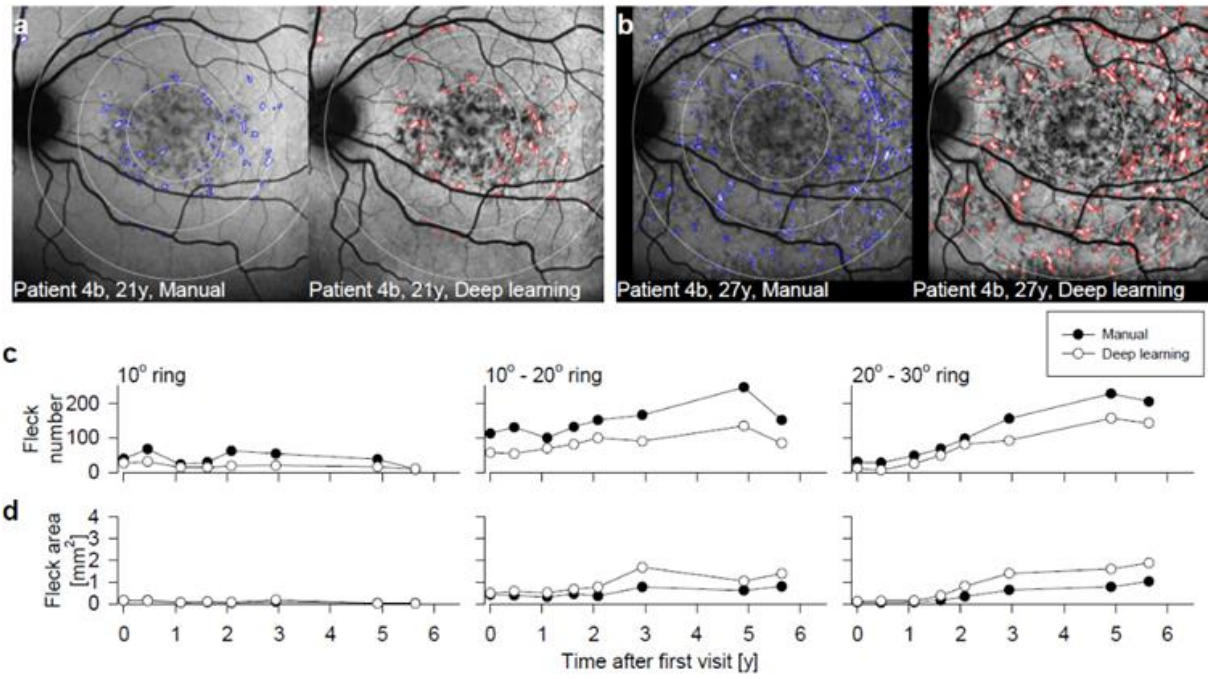


Figure 37: (a) Manual (blue outlines) and deep learning (red outlines) segmentation of the hyperautofluorescent flecks of Patient 4b, 21 years old, at the first visit. Each image is sub-divided into three rings, centred on the fovea. (b) Manual and deep learning segmentation of the same eye as Panel a six years later at 27 years old. All other details as per Panel a. Fleck number (c) and fleck area (d) plotted against time after first visit using manual (filled) and deep learning (unfilled) segmentation in the left, middle and right rings.

We demonstrated the feasibility of utilizing deep learning to segment and quantify FAF lesions, laying the foundation for future studies using fleck parameters as a trial outcome measure in Stargardt disease. More importantly, we trained and put forward a deep learning- based fleck segmentation method which is less time- consuming than manual marking.

Case Study 32: Using AI and Tele-health to Diagnose Oral Cancers

Partners: School & Hospital of Stomatology, Wuhan University, China

An important part of tele-health is to increase early diagnosis and timely treatment of a range of diseases. This is particularly true for regional and rural areas where specialists are limited to metropolitan hospitals. With oral cancer cases worldwide increasing, remote consultations by specialists from metropolitan hospitals would enable early diagnosis of oral cancer through remote screening and timely intervention through tele-consultation.

CSIRO and our partners have developed a tele-oral care system, which provides AI-driven oral mucosal disease classification and specialist-based clinical decision support. The system aims to provide oral medicine service to patients with oral mucosal diseases remotely and avoid unnecessary patient travel.

The oral images taken by clinicians from remote hospitals are sent via store-and-forward tele-medicine to a cloud-based AI-agent. The AI agent uses a clustering-based image segmentation to automatically extract lesion segments from oral images and then a deep convolutional neural network to perform image classification. This allows the potential diagnosis of various oral diseases to be made automatically and then reviewed and confirmed by the Oral Specialist**.** This workflow is shown in Figure 38.

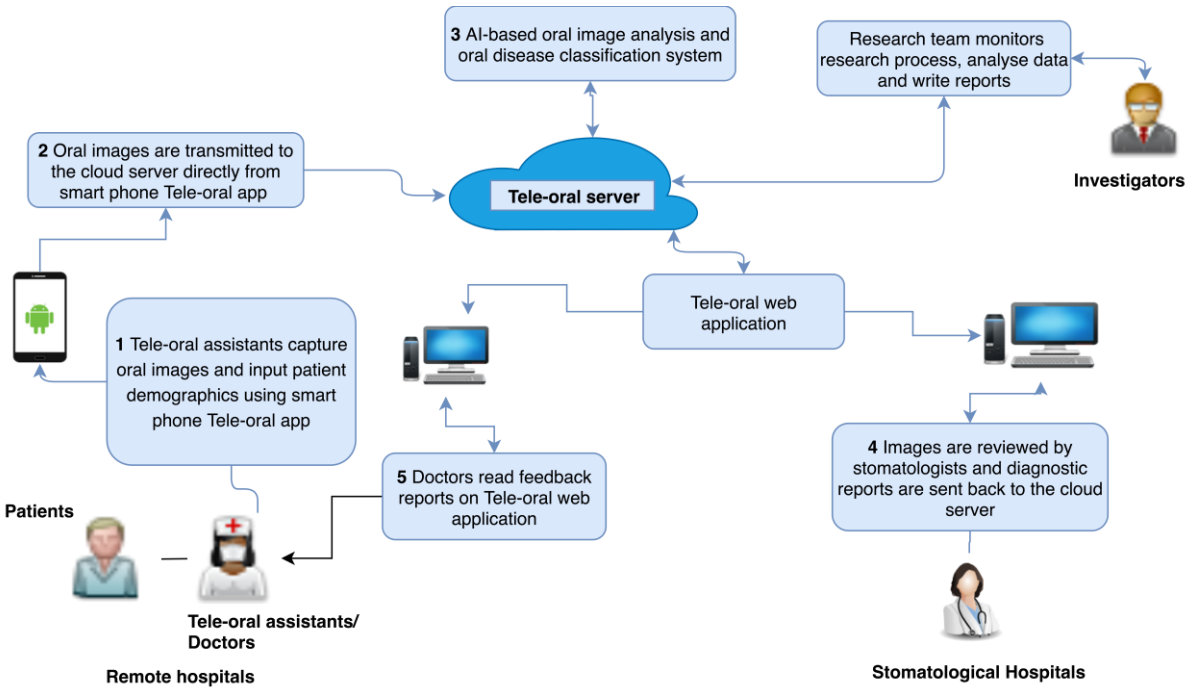


Figure 38: Workflow of Tele-oral care system.

The cloud-based AI agent was trained using more than 2000 oral mucosa images taken from the Wuhan university hospital. Currently the algorithm achieves 90% accuracy in distinguishing between ulcers and oral cancers. The technology is being evaluated by our collaborators at the Hospital of Stomatology of Wuhan University.

Case Study 33: Emotion Monitoring System for Tele-health and Robot-assisted Consultations

As we move more into tele-health and robot-assisted consultations (e.g. tele-presence robots), it is important that we don't compromise the clinician's ability to determine the mood, reactions and emotional states of a person due to a slow or patchy internet connection.

This project developed a system that provides clinicians with real-time information about a patient's emotions using audio-visual data. Bandwidth requirements of the system are lower than those needed by a video conference, allowing it to function even when the video quality is poor.

The project aimed to achieve face detection and emotion classification. Face detection was implemented using a histogram of oriented gradients (HOG) method. The HOG method involves extracting the contours of an image to be used as part of the detection mechanism. The HOG features were then used with a support vector machine algorithm to identify the location of faces.

Facial emotion recognition was then implemented through an Xception model, a convolutional neural network algorithm commonly used to classify a range of images. In our case, the model was trained to differentiate between images of people expressing Ekman's six basic emotions: anger, disgust, fear, happiness, sadness, and surprise. Audio classification was achieved using a long short-term memory recurrent neural network (LSTM-RNN), a deep learning algorithm used to classify between different sequences of data, such as speech. The final outcome used a decision level data fusion to combine image and audio classifiers. The accuracy of the model is approximately 73%.

Finally, a user interface was developed to present the information in two different ways: (1) the discrete output from the combined classifier based on the basic emotions, (2) a representation of the output mapped into Russel's Circumplex model of emotions, that distributes emotions in a two-dimensional circular space, containing different levels of valence and arousal (Figure 39). A timeline that displays how patient's changed over-time is also available.

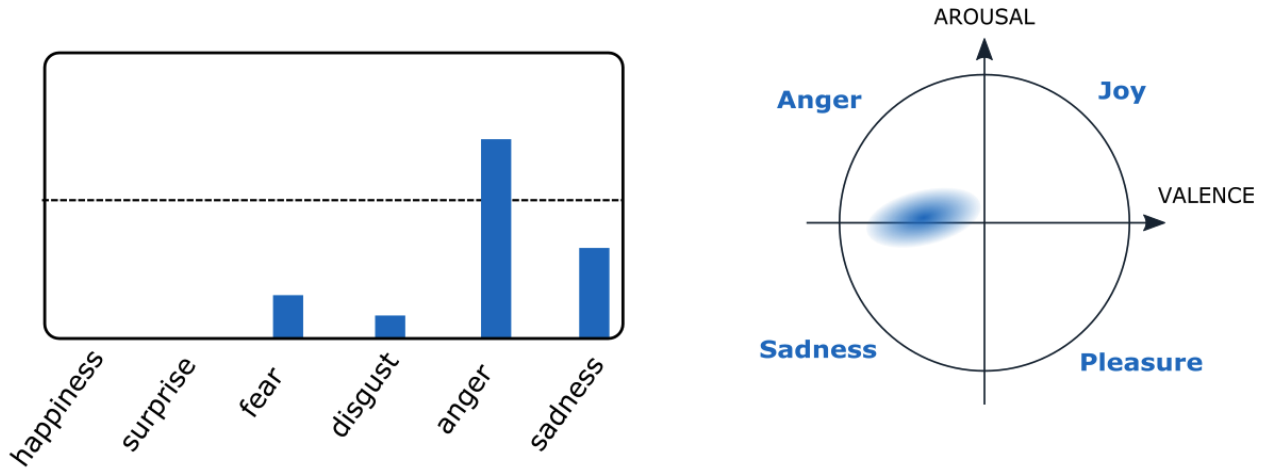


Figure 39: User interface can display emotions based on Ekman's six basic emotions (left) and Russel's Circumplex model of emotions (right).

In future work, this system will be incorporated into a tele-health/tele-presence setting. Clinician perception of the user interface will be evaluated, as well as the efficacy of the full system in the support of remote sessions.

Case Study 34: Autonomous Behaviour Planning for Socially-assistive Robots

Most individuals on the autism spectrum have a preference for predictable systems. In response to that, socially-assistive robots have been proposed as tools to supplement traditional therapy and education for children with a diagnosis of autism spectrum disorder. One major challenge in socially-assistive robots is the substantial labour required to program robot behaviours that are appropriate for therapy and education.

We have developed a system that allows robots to learn how to behave during a conversation, by 'observing' how people interact with each other. By using this system, robots can create appropriate behaviours to accompany their speech, without the need for manual programming.

Behaviours are, in essence, a time series of poses. To create a behaviour database, human poses were extracted from videos using the open source computer vision library 'FastPose' (Figure 40). By using context-appropriate videos (i.e. educators interacting with groups of children on the autism spectrum), this stage engineered the system toward a use case in the education for children on the autism spectrum.

A regression network was then used to capture the correlation between transcribed text in the videos, and the extracted behaviours. Essentially, the regression algorithm identifies which words in a sentence are more important in deciding a behaviour. During behaviour selection, the model compares a new input sentence with all sentences in the database using the similarities between word vectors quantified using the global vectors for word representation (GloVe) algorithm. The sentence in the database with the highest similarity score is selected, and the behaviour corresponding to that sentence is used to accompany the new input sentence.

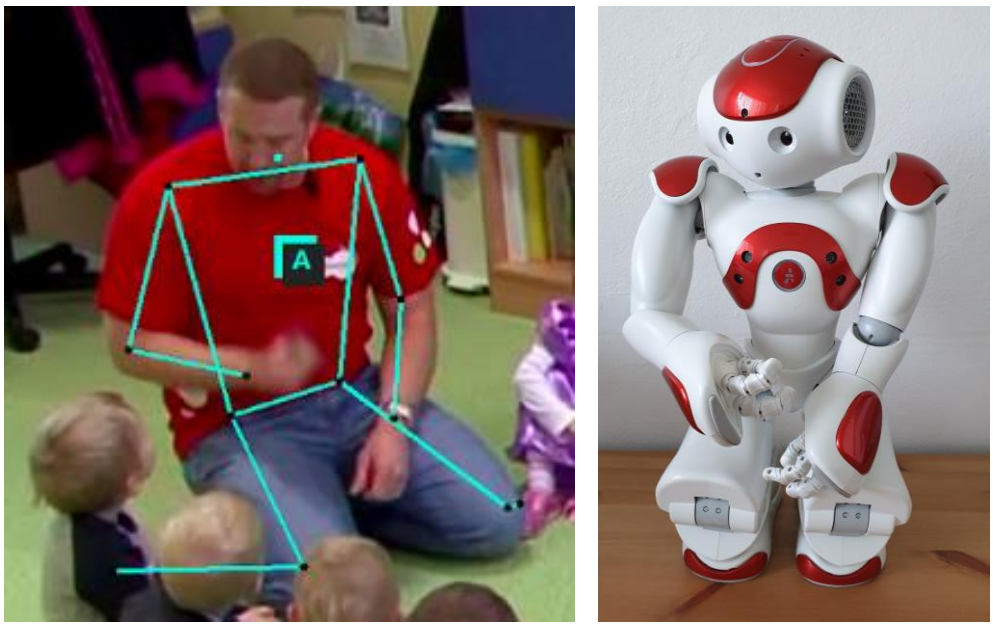


Figure 40: Pose recognition (left) and robot behaviour (right). Video used to extract this pose sourced from https://www.youtube.com/user/Wattsenglish.

References: Case Study 34: Autonomous Behaviour Planning for Socially-Assistive Robots, page 80.

The proposed system successfully captures metaphoric and iconic gestures, with behaviours generated showing context-specific characteristics. A user study showed comparative performance between the autonomously generated behaviours and a set of industrial-level pre-programmed behaviours. The autonomous system, however, provides added flexibility to learn an unlimited number of behaviours from visual observation. The new system will be incorporated into our current robot-assistive education program for children and adolescent students on the autism spectrum, to facilitate planning and delivery of robot-assisted lessons by educators.

9 Harnessing the full value of AI and ML in Healthcare

A number of recent reports from Australian organisations, such as the Australian Council of Learned Academies[7], the Australian Academy of Health and Medical Sciences[8], our Data61 Business Unit[9], as well as international organisations such as the UK National Health Service[10], have provided significant overviews of the likely impact of AI on society. These reports provide an excellent analysis of three key areas of AI implementation: the promise of the technology; the legal, ethical and social concerns; and the workforce issues that will confront healthcare, and in fact all industries, as they embrace AI.

Our report complements these high-level reports by providing a significant number of Case Studies demonstrating a range of AI and ML approaches in use across a wide selection of application areas.

In this Chapter we consider the key practical challenges in harnessing the full value of AI and ML in healthcare for Australia and provide some thoughts on what AI might look like in the future.

9.1 Who needs to be involved?

As we implement AI and ML in healthcare it will be important to engage fully with everyone that the technology will impact – including the patients, carers, clinicians, technicians, and health service and IT managers. This will ensure a comprehensive understanding of the promise, and the limits, of the new technology.

Embracing AI is likely to lead to significantly better outcomes for all concerned. Clinicians will be able to benefit from the ability of AI to undertake some of the time-consuming, mundane tasks they currently do – and be provided with information to which they may not have otherwise had access. Clinicians who use AI will have the capacity to provide better patient care than those who do not.

For patients and their families, understanding the tools which the clinical workforce is using should instil in them a confidence that the health service is using the best possible technology to give them the best possible treatment. For complex diseases or illnesses that extend over time such as cancer, patients and families may also undertake their own research into the disease, and may find AI tools that support them along their healthcare journey.

---

[7] Walsh, et al. (2019) *The effective and ethical development of artificial intelligence: An opportunity to improve our wellbeing*. Report for the Australian Council of Learned Academies, www.acola.org.

[8] AAHMS (2020) Artificial Intelligence in Health: Exploring the Opportunities and Challenges. A Roundtable Report.

[9] Hajkowicz et al. (2019) Artificial intelligence: Solving problems, growing the economy and improving our quality of life. CSIRO Data61, Australia

[10] Harwich & Laycock (2019) Thinking on its Own. AI in the NHS. #reformhealth.

For all involved in the health service – executives, IT and technical workforce – AI will provide an opportunity for new technologies to be used to make service provision more efficient and provide better services to their patients.

AI developers must be able to provide information along with their tools that allows all users to understand both the promise and limits of the proposed technology.

AAAIH: A national health AI community

The Australian Alliance for Artificial Intelligence in Healthcare (AAAIH) was formed in 2018 to start the process of bringing together those organisations with an interest in the introduction of AI into Australia's healthcare system. Led by Macquarie University's Professor Enrico Coiera, Australia's foremost Health Informatician, the AAAIH brings together more than 90 national and international partners and engaged stakeholders in academia, government, consumer, clinical, industry organisations, and peak bodies to translate frontier AI technologies into real-world health services.

The members of the Alliance see the need for a coordinated approach to the introduction of AI in Australia, building on the standards being introduced into our healthcare system and the fantastic AI technologies being developed and trialled all around Australia.

9.2 What is the problem the AI is solving?

There are a range of issues to be considered in the development of AI solutions. As demonstrated in the Case Studies in this report, AI underpins a diverse range of technologies and it is not simply a matter of installing an AI platform from a vendor and all your problems are solved.

The first step is to define the real-world problem to be solved. Many AI projects start from an idealised place – perfect data or without fully considering how it will be used – and while the results may be impressive, the lack of consideration of real-life constraints will hamper adoption. Defining the problem should be undertaken strategically, by measuring the impact and collecting the data needed to evaluate the nature and extent of the issue at hand. This might include variables such as who has the problem, where does the problem lie, and what resource or temporal barriers need to be overcome? Once the problem has been defined, interim and long-term goals need to be determined and monitored, such as defining an outcome within a specified time frame, measures of performance toward that outcome and criteria for failure.

The next stage is of course to develop the solution. AI tools can also be developed and optimized in many ways, but it is almost always an iterative tactic of trying different approaches, revisiting the data, adding additional features, considering the uncertainty introduced, etc. So, the solution is a result of iterative implementation and experimentation. Ultimately, the proposed solution needs to be economically desirable, technically possible, practically feasible, and socially acceptable.

At AEHRC, we use a wide range of AI techniques and will consider the best AI approach(es) before undertaking any digital health project. Our primary focus is to move healthcare innovation towards precision and personalised medicine and to make health systems more efficient and accessible.

The development and implementation of the AI is not the end. The way the AI is used and the answers it gives should be monitored and evaluated, and where appropriate, revised to take into account changes in the real-world situation which will inevitably occur.

The nature of the use of AI will also change as it becomes more advanced and as we undertake more implementations. Currently, the use of AI is often about humans making decisions with information. New systems will likely provide clinicians and patients with considerably more tailored and personalised information than is currently the case.

This change in the use of AI is giving rise to a greater focus on areas such as 'explainability' that enable the users to understand how the AI reached a decision and 'sentiment' style methods that are able to discern contradicting evidence and different points of view.

9.3 Data, Data, Data

It has been argued that "*Almost all conversations on AI quickly go back to data*". Our succinct summary of those conversations is:

> High quality AI solutions are dependent on high quality data – and lots of it.
>
> *(We did say succinct…, Eds.)*

9.4 Implementation Challenges – 'the Last Mile'

A critical aspect to the success of AI in health is the final step of integrating AI into the health system. This is often called the 'the last mile'. Some of the best AI solutions are often not the ones to be first implemented – rather, understanding how the AI solutions fit in and support the work of the people concerned to ensure success is the last mile to achieve a result. This integration often needs an implementation science approach to developing the AI from the start.

Implementing AI will require the usual process of ethically approved clinical trials, evaluations and regulatory approval. At the AEHRC we work within these guidelines to ensure that the digital health solutions we provide – many of which use AI – are effective and can be trusted.

While there is much hype around AI in healthcare at the moment, AI is not a silver bullet to the problems in the healthcare system. The above-mentioned reports provide excellent overviews of the issues which must be addressed across ethical, societal and regulatory fields. AI has been through a number of changes of emphasis over the past 75 years – the technology will continue to evolve as new mathematics, statistics and information science is developed.

To fully take advantage of AI in Australia, we need to consider the 'infrastructure' that will support it. This is the technical platforms, regulatory and approval processes that enable AI to be implemented and monitored in a way that enables the innovation it can deliver.

9.5 What does success look like for Australia?

The measure of success for AI in Australian healthcare is likely to be more than just improvements to health outcomes, efficiency or other empirical measures.

Rather, success for AI in health in Australia will be measured when we have a workforce and population which understands the role of AI in healthcare and is confident that Australia has the technical, regulatory and societal infrastructure to support the safe implementation and use of AI in healthcare. This will ensure that the technologies that rely on AI are safe, effective and an improvement on the status quo.

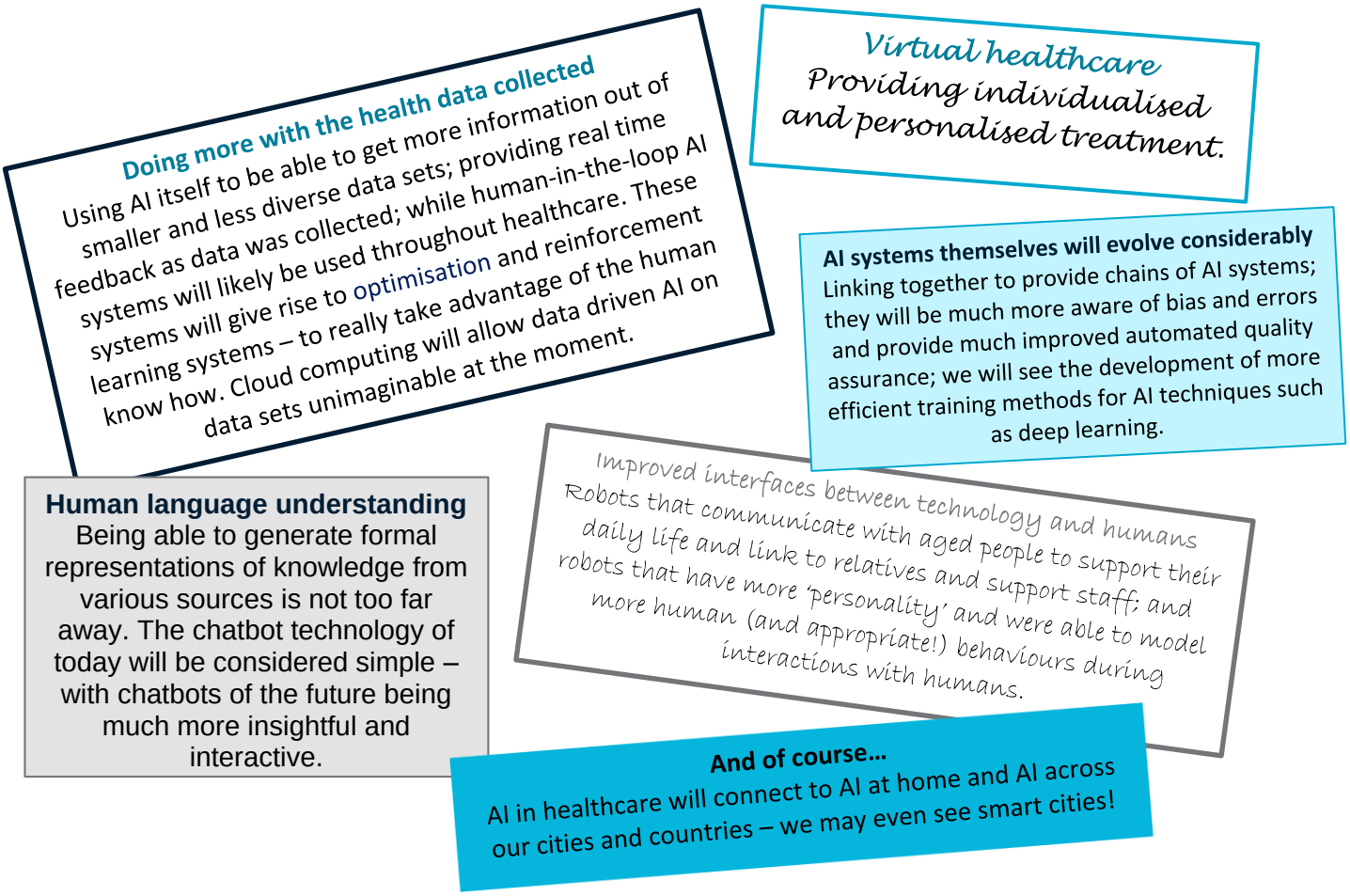
This means not taking a 'cookie cutter' approach to the use of AI; accepting that AI will not be perfect but having the knowledge and systems to deal with the errors; and of course knowing that the AI will reflect the biases in the data on which it was based (See Section 9.3 above).

9.6 What does the future hold?

The Case Studies in this report are very embedded in the present – a view of how AI and ML is being used right now. And this final Chapter until now has sought to temper the enthusiasm we sometimes see in the community, that AI will solve all the challenges our healthcare faces. This may give the idea that we at AEHRC don't share in the enthusiasm for these technologies.

Nothing could be further from the truth.

In putting this document together, we asked our scientists about some of the immediate challenges they were working on and where they thought AI would have a big impact over the next 10 years. The answers they provided give just a snapshot of how AI will change healthcare in Australia.



All our scientists recognised that AI algorithms implemented ethically, with social responsibility and ensuring best outcomes, was key to a successful AI-driven future for our healthcare system.

Looking 10 years ahead, we would hope to see that AI has been central in building a truly learning healthcare system in Australia. That AI solutions will be enabling more targeted, tailored and timely interventions for patients. And that AI will contribute to a more efficient health system that is able to translate chronic and post-operative care to home and community, such that hospital services are predominantly focused on their original purpose, and what they do best – acute care.

10 References

Box 1: Virtual Care: From Mobile Health to Wearable Devices

Varnfield, M., Karunanithi, M., Lee, C. K., Honeyman, E., Arnold, D., Ding, H., ... & Walters, D. L. Smartphone-based home care model improved use of cardiac rehabilitation in postmyocardial infarction patients: results from a randomised controlled trial. Heart, 2014, 100(22), 1770-1779.

Zhang, Q., Karunanithi, M., Rana, R., & Liu, J. Determination of activities of daily living of independent living older people using environmentally placed sensors. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2013 pp. 7044-7047.

Box 3: Cloud Computing for Machine Learning

Bayat, A. et al (2020) VariantSpark: Cloud-Based Machine Learning for Association Study of Complex Phenotype and Large-Scale Genomic Data" GigaScience.

Case Study 1: Risk Stratification for Identifying Patients at Risk of Hospitalisation

Khanna S, Rolls DA, Boyle J, Xie Y, Jayasena R, Hibbert M, et al. A risk stratification tool for hospitalisation in Australia using primary care data. Scientific Reports. 2019 Mar 21;9(1):1–14.

Khanna S, Good N, Boyle J. Predicting Unplanned Return to Hospital for Chronic Disease Patients. Stud Health Technol Inform. 2016; 227:67–73.

Khanna S, Boyle J, Good N. Precise prediction for managing chronic disease readmissions. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2014. p. 2734–7.

Case Study 2: Improving Scheduling of Elective Surgery

Khanna S, Sattar A, Hansen D, Stantic B. An Efficient Algorithm for Solving Dynamic Complex DCOP Problems. In: 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. 2009. p. 339–46.

Khanna S, Sattar A, Boyle J, Hansen D, Stantic B. An Intelligent Approach to Surgery Scheduling. In: Desai N, Liu A, Winikoff M, editors. Principles and Practice of Multi-Agent Systems. Berlin, Heidelberg: Springer; 2012. p. 535–50. (Lecture Notes in Computer Science).

Khanna S, Cleaver T, Sattar A, Hansen D, Stantic B. Multiagent Based Scheduling of Elective Surgery. In: Desai N, Liu A, Winikoff M, editors. Principles and Practice of Multi-Agent Systems. Berlin, Heidelberg: Springer; 2012. p. 74–89. (Lecture Notes in Computer Science).

ShahabiKargar Z, Khanna S, Good N, Sattar A, Lind J, O'Dwyer J. Predicting Procedure Duration to Improve Scheduling of Elective Surgery. In: Pham D-N, Park S-B, editors. PRICAI 2014: Trends in Artificial Intelligence.

Case Study 4: Predicting Demand for Hospital Services

Boyle, Justin; Jessup, Melanie; Crilly, Julia; Green, David; Lind, James; Wallis, Marianne; et al. Predicting Emergency Department Admissions. Emergency Medicine Journal. 2012; 29(5):358-365. https://doi.org/10.1136/emj.2010.103531

Crilly, Julia; Boyle, Justin; Jessup, Melanie; Wallis, Marianne; Lind, James; Green, David; et al. The Implementation and Evaluation of the Patient Admission Prediction Tool: Assessing Its Impact on Decision-Making Strategies and Patient Flow Outcomes in 2 Australian Hospitals. Quality Management in Health Care. 2015; 24(4):169-176. https://doi.org/10.1097/QMH.0000000000000070

Boyle, Justin; Ireland, Derek; Webster, Fiona; O'Sullivan, Kim. Predicting Demand for Hospital Capacity Planning. In: BHI-2016 International Conference on Biomedical and Health Informatics; 24th-27th Feb 2016; Las Vegas. IEEE; 2016. 328-331. https://ieeexplore.ieee.org/document/7455901

Jessup, Melanie; Crilly, Julia; Boyle, Justin; Wallis, Marianne; Lind, James; Green, David; et al. Users' experiences of an Emergency Department patient admission predictive tool: a qualitative evaluation. Health Informatics Journal. 2016; 22(3):618-632. https://doi.org/10.1177/1460458215577993

Boyle, Justin; Sparks, Ross; Keijzers, Gerben; Crilly, Julia; Lind, James; Ryan, Louise. Prediction and surveillance of influenza epidemics. MJA. 2011; 194(4 Supp):S28-33. https://onlinelibrary.wiley.com/doi/epdf/10.5694/j.1326-5377.2011.tb02940.x

Jessup, Melanie; Wallis, Marianne; Boyle, Justin; Crilly, Julia; Lind, James; Green, David; et al. Implementing an emergency department patient admission predictive tool: insights from practice. Journal of Health Organization and Management. 2010; 24(3):306-318. https://doi.org/10.1108/14777261011054635

Boyle, Justin; Le Padellec, Remy; Ireland, Derek. Statewide Validation of a Patient Admissions Prediction Tool. In: EMBC 2010 - IEEE EMBS Conference: 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 31 - September 4, 2010; August 31 - September 4, 2010. IEEE; 2010. 3887-3890. https://www.ncbi.nlm.nih.gov/pubmed/21097075

Boyle, J.; Wallis, M.; Jessup, M.; Crilly, J. Regression forecasting of patient admission data. In: 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBC 2008); 2008; Vancouver, BC. IEEE; 2008. 3819-3822. https://ieeexplore.ieee.org/document/4650041

Case Study 5: Predicting Patient Deterioration

Mann, Kay; Good, Norm; Khanna, Sankalp. Predicting Patient Deterioration - Initial Pilot Activities. Client Report EP194012. CSIRO; 2019.

Case Study 6: Discovery Genetic Drivers for Amyotrophic Lateral Sclerosis

Bayat, A. et al (2020) VariantSpark: Cloud-Based Machine Learning for Association Study of Complex Phenotype and Large-Scale Genomic Data" GigaScience

Nicolas A, et al (2018) Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. Neuron.

Wolf, J., Safer, A., Wöhrle, J.C. et al. Factors predicting one-year mortality in amyotrophic lateral sclerosis patients - data from a population-based registry. BMC Neurol 14, 197 (2014).

Case Study 8: Digitally Guided Genome Editing for Gene Therapy

Wilson, L.O.W., et al (2018) "High activity target-site identification using phenotypic independent CRISPR-Cas9 core functionality" The CRISPR Journal 1(2).

Wilson L.O.W., et al (2019) "VARSCOT: variant-aware detection and scoring enables sensitive and personalized off-target detection for CRISPR-Cas9" BMC Biotechnology 19:40.

O'Brien A.R., et al (2019) "Unlocking HDR-mediated nucleotide editing by identifying high-efficiency target sites using machine learning" Scientific Reports 9:2788.

O'Brien A.R., et al (2020) "Domain-specific introduction to machine learning terminology, pitfalls and opportunitites in CRISPR-based gene editing" Briefings in Bioinformatics.

Wilson L.O.W., et al (2018) "The current state and future of CRISPR-Cas9 gRNA design tools" Frontiers in pharmacology 9:749.

Case Study 9: Informing Public Health Responses for Infectious Disease Outbreaks

Bauer D.C., et al, (2020) "Supporting pandemic response using genomics and bioinformatics: a case study on the emergent SARS-CoV-2 outbreak" Transboundary and Emerging Diseases.

Case Study 10: Aging in Place and Objective Activities of Daily Living

Tran SN, AA Garcez, T Weyde, J Yin, Q Zhang, M Karunanithi, Sequence classification restricted boltzmann machines with gated units, IEEE Transactions on Neural Networks and Learning Systems.

Tran SN, D Nguyen, TS Ngo, XS Vu, L Hoang, Q Zhang, M Karunanithi, On Multi-resident activity recognition in ambient smart-homes, Artificial Intelligence Review.

Case Study 11: Using Deep Learning to Identify Individuals in a Multi-resident House

Tran SN, Q Zhang, V Smallbon, M Karunanithi, Multi-Resident Activity Monitoring in Smart Homes: A Case Study IEEE International Conference on Pervasive Computing and Communications.

Yin J, SN Tran, Q Zhang, Human identification via unsupervised feature learning from UWB radar data Pacific-Asia Conference on Knowledge Discovery and Data Mining.

Case Study 12: Movement Classification for Early Diagnosis of Cerebral Palsy

Redd CB, Barber LA, Boyd RN, Varnfield M, Karunanithi MK. Development of a Wearable Sensor Network for Quantification of Infant General Movements for the Diagnosis of Cerebral Palsy. In2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2019 Jul 23 (pp. 7134-7139). IEEE.

Murad A, Pyun JY. Deep recurrent neural networks for human activity recognition. Sensors. 2017 Nov;17(11):2556.

Marcroft C, Khan A, Embleton ND, Trenell M, Plötz T. Movement recognition technology as a method of assessing spontaneous general movements in high risk infants. Frontiers in Neurology. 2015 Jan 9;5:284.

Case Study 13: “Reasoning” over Medical Knowledge – the Snorocket Classifier and SNOMED CT

Lawley M. Exploiting Fast Classification of SNOMED CT for Query and Integration of Health Data. Proceedings of the 3rd International Conference on Knowledge Representation in Medicine, 2008; 8-14.

Lawley M, Vickers D, and Hansen D. Converting Ad Hoc terminologies to SNOMED CT extensions. Health Informatics Conference, 2008 Proceedings.

Lawley M, Bousquet C. Fast classification in Protégé: Snorocket as an OWL 2 EL Reasoner. Proc. 6th Australasian Ontology Workshop (IAOA’10).

Michel J, Lawley M, Chu A, Barned J. Mapping the Queensland Health iPharmacy Medication File to the Australian Medicines Terminology Using Snapper. Studies in health technology and informatics 168, 104-116.

Case Study 16: Post-coordination and the SNOMED CT Expression Constraint Language

Lawley M. Exploiting Fast Classification of SNOMED CT for Query and Integration of Health Data. Proceedings of the 3rd International Conference on Knowledge Representation in Medicine, 2008; 8-14.

Lawley M, Bousquet C. Fast classification in Protégé: Snorocket as an OWL 2 EL Reasoner. Proc. 6th Australasian Ontology Workshop (IAOA’10).

Case Study 19: Automating Cancer Registry Tasks to Enhance Clinical Data Quality

Nguyen AN, Lawley MJ, Hansen DP, Bowman RV, Clarke BE, Duhig EE, Colquist S. Symbolic rule-based classification of lung cancer stages from free-text pathology reports. Journal of the American Medical Informatics Association. 2010 Jul 1;17(4):440-5.

Nguyen AN, Moore J, O’Dwyer J, Philpot S. Assessing the utility of automatic cancer registry notifications data extraction from free-text pathology reports. In AMIA Annual Symposium Proceedings 2015 (Vol. 2015, p. 953). American Medical Informatics Association

Nguyen AN, Moore J, O'Dwyer J, Philpot S. Automated cancer registry notifications: validation of a medical text analytics system for identifying patients with cancer from a state-wide pathology repository. In AMIA Annual Symposium Proceedings 2016 (Vol. 2016, p. 964). American Medical Informatics Association.

Koopman B, Nguyen A, Cossio D, Courage MJ, Francois G. Extracting Cancer Mortality Statistics from Free-text Death Certificates: A View from the Trenches. In Proceedings of the 23rd Australasian Document Computing Symposium 2018 Dec 11 (p. 6). ACM.

Case Study 20: Checking Radiology Reports to Prevent Missed Fractures

Zuccon, A. Wagholikar, A. Nguyen, L. Butt, K. Chu, J. Greenslade, S. Martin, "Automatic Classification of Free-Text Radiology Reports to Identify Limb Fractures using Machine Learning Algorithms and SNOMED CT" AMIA CRI, 2013

Koopman B, Zuccon G, Wagholikar A, Chu K, O'Dwyer J, Nguyen A, Keijzers G. Automated Reconciliation of Radiology Reports and Discharge Summaries. AMIA 2015 Annual Symposium, 2015.

Hassanzadeh H, Nguyen A, Karimi S, Chu K. Transferability of artificial neural networks for clinical document classification across hospitals: A case study on abnormality detection from radiology reports. Journal of biomedical informatics. 2018 Sep 1;85:68-79.

Case Study 21: Tackling Antimicrobial Resistance with Test Result Reviews

Nguyen A, Hassanzadeh H, Zhang Y, O'Dwyer J, Conlan D, Lawley M, Steel J, Loi K, Rizzo P. A Decision Support System for Pathology Test Result Reviews in an Emergency Department to Support Patient Safety and Increase Efficiency. Studies in health technology and informatics. 2019; 264:729-33.

Case Study 22: Searching Engines for Compiling Systematic Reviews

Scells, H., L. Azzopardi, G. Zuccon, and B. Koopman. Query variation performance prediction for systematic reviews. In Proc of SIGIR, 2018.

Scells, H., G. Zuccon, and B. Koopman. Automatic boolean query refinement for systematic review literature search. In The World Wide Web Conference, 1646–1656, 2019.

Scells, H., G. Zuccon, B. Koopman, and J. Clark. Automatic boolean query formulation for systematic review literature search. In The World Wide Web Conference, 2020.

Scells, H., G. Zuccon, B. Koopman, and J. Clark. A computational approach for objectively derived systematic review search strategies. In European Conference on Information Retrieval. Springer, 2020.

Case Study 23: Quantifying Semantic Similarity of Clinical Evidence in the Biomedical Literature

Hassanzadeh H, Nguyen A, Verspoor K. Quantifying semantic similarity of clinical evidence in the biomedical literature to facilitate related evidence synthesis. Journal of biomedical informatics. 2019 Dec 1;100:103321.

Case Study 25: Early Prediction of Neurodevelopmental Outcomes of Preterm Infants

Saha, S., Pagnozzi, A.M., Bourgeat, P., et al. Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model. Neuroimage (2020). https://www.sciencedirect.com/science/article/pii/S1053811920302949

Saha, S., Pagnozzi, A.M., George, J., et al. Investigating Brain Age Deviation in Preterm Infants: A Deep Learning Approach. MICCAI (2018). http://link.springer.com/10.1007/978-3-030-00807-9_9

Case Study 26: Clinical Decision Support Tool for Children

Pagnozzi, A.M., Dowson, N., Bradley, A.P., et al. Expectation-Maximization with Image-Weighted Markov Random Fields to Handle Severe Pathology. Digit Image Comput Tech Appl DICTA 2015.

Pagnozzi, A.M., Shen, K., Doecke, J.D., et al. Using ventricular modeling to robustly probe significant deep gray matter pathologies: Application to Cerebral Palsy. Hum Brain Mapp 2016. https://doi.org/10.1002/hbm.23276

Pagnozzi, A.M., Dowson, N., Fiori, S., et al. Alterations in regional shape on ipsilateral and contralateral cortex contrast in children with unilateral Cerebral Palsy and are predictive of multiple outcomes. Hum Brain Mapp (2016). http://doi.wiley.com/10.1002/hbm.23262

Pagnozzi, A.M., Dowson, N., Doecke, J., et al. Automated, quantitative measures of grey and white matter lesion burden correlates with motor and cognitive function in children with unilateral Cerebral Palsy. NeuroImage (2016). https://doi.org/10.1016/j.nicl.2016.05.018

Case Study 27: AI for the Diagnosis of Musculoskeletal Disorders

Neubert A, Bourgeat P, Wood J, et al. Applying simultaneous super-resolution and contrast synthesis to routine clinical magnetic resonance images for automated segmentation of knee joint cartilage. In: International Conference on Medical Imaging with Deep Learning -- Extended Abstract Track. London, United Kingdom; 2019. https://openreview.net/forum?id=S1xKZwFpKN.

Chandra SS, Dowling JA, Engstrom C, et al. A lightweight rapid application development framework for biomedical image analysis. Comput Methods Programs Biomed. 2018;164:193-205. doi:10.1016/j.cmpb.2018.07.011

Xia Y, Fripp J, Chandra SS, Schwarz R, Engstrom C, Crozier S. Automated bone segmentation from large field of view 3D MR images of the hip joint. Phys Med Biol. 2013;58(20):7375-7390. doi:10.1088/0031-9155/58/20/7375

Neubert A, Fripp J, Engstrom C, et al. Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys Med Biol. 2012;57(24):8357-8376. doi:10.1088/0031-9155/57/24/8357

Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. IEEE Trans Med Imaging. 2010;29(1):55-64. doi:10.1109/TMI.2009.2024743

Paproki A, Engstrom C, Strudwick M, et al. Automated T2-mapping of the Menisci From Magnetic Resonance Images in Patients with Acute Knee Injury. Acad Radiol. 2017;24(10):1295-1304. doi:10.1016/j.acra.2017.03.025

Case Study 28: AI Driven Prostate Radiation Therapy

Rivest-Hénault D, Dowson N, Greer PB, Fripp J, Dowling JA. Robust inverse-consistent affine CT-MR registration in MRI-assisted and MRI-alone prostate radiation therapy. Med Image Anal. 2015;23(1). doi:10.1016/j.media.2015.04.014.

Ghose S, Greer PB, Sun J, et al. Regression and statistical shape model based substitute CT generation for MRI alone external beam radiation therapy from standard clinical MRI sequences. Phys Med Biol. 2017. doi:10.1088/1361-6560/aa9104.

Largent A, Barateau A, Nunes JC, et al. Pseudo-CT Generation for MRI-Only Radiation Therapy Treatment Planning: Comparison Among Patch-Based, Atlas-Based, and Bulk Density Methods. Int J Radiat Oncol Biol Phys. 2019;103(2):479-490. doi:10.1016/j.ijrobp.2018.10.002.

Arabi H, Dowling JA, Burgos N, et al. Comparative study of algorithms for synthetic CT generation from MRI: Consequences for MRI -guided radiation planning in the pelvic region. Med Phys. 2018;45(11):5218-5233. doi:10.1002/mp.13187.

Largent A, Barateau A, Nunes JC, et al. Comparison of Deep Learning-Based and Patch-Based Methods for Pseudo-CT Generation in MRI-Based Prostate Dose Planning. Int J Radiat Oncol Biol Phys. 2019;105(5):1137-1150. doi:10.1016/j.ijrobp.2019.08.049.

Chandra SS, Dowling JA, Shen K-K, et al. Patient specific prostate segmentation in 3-D magnetic resonance images. IEEE Trans Med Imaging. 2012;31(10). doi:10.1109/TMI.2012.2211377.

Dowling JA, Lambert J, Parker J, et al. An atlas-based electron density mapping method for magnetic resonance imaging (MRI)-alone treatment planning and adaptive MRI-based prostate radiation therapy. Int J Radiat Oncol Biol Phys. 2012;83(1):e5-11.

Dowling JA, Sun J, Pichler P, et al. Automatic substitute computed tomography generation and contouring for magnetic resonance imaging (MRI)-alone external beam radiation therapy from standard MRI sequences. Int J Radiat Oncol Biol Phys. 2015;93(5). doi:10.1016/j.ijrobp.2015.08.045.

Greer P, Martin J, Sidhom M, et al. A multi-center prospective study for implementation of an MRI-only prostate treatment planning workflow. Front Oncol. 2019;9(AUG). doi:10.3389/fonc.2019.00826.

Case Study 29: Automated Analysis of PET Images for Neurodegenerative Diseases

Bourgeat P, Villemagne VL, Dore V, et al. Comparison of MR-less PiB SUVR quantification methods. Neurobiol Aging. 2015;36 Suppl 1:S159-166. doi:10.1016/j.neurobiolaging.2014.04.033

Zhou L, Salvado O, Dore V, et al. MR-less surface-based amyloid assessment based on 11C PiB PET. PLoS ONE. 2014;9(1):e84777. doi:10.1371/journal.pone.0084777

Bourgeat P, Dore V, Fripp J, Villemagne VL, Rowe CC, Salvado O. Computational analysis of PET by AIBL (CapAIBL): a cloud-based processing pipeline for the quantification of PET images. In: Ourselin S, Styner MA, eds. ; 2015:94132V. doi:10.1117/12.2082492

Dore V, Bourgeat P, Burnham SC, et al. Automated reporting of tau pet quantification on the brain surface. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 2019;15(7):P1269. doi:10.1016/j.jalz.2019.06.4811

Doré V, Bullich S, Rowe CC, et al. Comparison of 18F-florbetaben quantification results using the standard Centiloid, MR-based, and MR-less CapAIBL® approaches: Validation against histopathology. Alzheimers Dement. 2019;15(6):807-816. doi:10.1016/j.jalz.2019.02.005

Bourgeat P, Doré V, Fripp J, et al. Implementing the Centiloid transformation for 11C-PiB and β-amyloid 18F-PET tracers using CapAIBL. Neuroimage. 2018;183:387-393. doi:10.1016/j.neuroimage.2018.08.044

Dore V, Bourgeat P, Cummins TL, et al. VISUAL ASSESSMENT OF β-AMYLOID PET SCAN IS IMPROVED BY CAPAIBL. Alzheimer's & Dementia: The Journal of the Alzheimer's Association. 2018;14(7):P18-P19. doi:10.1016/j.jalz.2018.06.2072.

Case Study 34: Autonomous Behaviour Planning for Socially-Assistive Robots

Xiao, Y. Autonomous Behaviour Planning of an Assistive Social Robot for Autism (Unpublished honours thesis). University on New South Wales, Australia.

Glossary

Active Appearance Models	Extend the registration atlas approach to generate shape and appearance priors to initialise a local search for organ boundaries
Advanced Pattern Matching	Maps the PET tracer retention onto a template brain surface
Apache Spark	A distributed computing framework
Artificial Neural Networks	Artificial neural networks or connectionist systems are computing systems inspired by biological neural networks. Such systems "learn" to perform tasks by considering examples, generally without being programmed with task-specific rules.
Autoregressive Integrated Moving Averages	A statistical analysis model that uses time series data to either better understand the data set or to predict future trends. (See also exponential smoothing)
Basic Splines	Smoothing splines are function estimates, obtained from a set of noisy observations, in order to balance a measure of goodness of fit. They provide a means for smoothing noisy data.
Bayes Network	A probabilistic graphical model that represents a set of variables and their conditional dependencies via a directed acyclic graph.
Biomarker	A naturally occurring molecule, gene, or characteristic by which a particular pathological or physiological process, disease, etc. Can be identified.
Centiloid	A 100-point scale which is an average value of zero in “high certainty” amyloid negative subjects and an average of 100 in “typical” Alzheimer's disease patients.
Computer Vision	Artificial intelligence that trains computers to interpret and understand the visual world.
Convolutional Neural Networks	Convolutional neural networks are simply neural networks that consist of a series of convolutional layers that employ a mathematical operation called convolution. They have applications in image and video recognition, recommender systems, image classification, medical image analysis, natural language processing, and financial time series.
Decision Tree	A decision support tool that uses a tree-like model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. It is one way to display an algorithm that only contains conditional control statements. An optimal decision tree is a tree that accounts for most of the data, while minimizing the number of levels (or "questions").
Deep Learning	Part of a broader family of machine learning methods based on artificial neural networks with representation learning. Learning can be supervised, semi-supervised or unsupervised.
Deep Neural Networks	An artificial neural network with multiple layers between the input and output layers. The dnn finds the correct mathematical manipulation to turn the input into the output, whether it be a linear relationship or a non-linear relationship.
Description Logic	A family of knowledge representation languages with (usually) decidable semantics and efficient reasoning algorithms.
Diffeomorphic Demons Algorithm	A voxel-wise deformable image registration algorithm. The diffeomorphic property means topology is preserved by the displacement field and the image mapping is invertable.
Dimensionality Reduction Methods	Enable the visualisation of objects by a two- or three-dimensional graph so similar objects are represented by nearby points and dissimilar objects by distant points
Explainability, Explainable ML	Explainable AI refers to methods and techniques in the application of artificial intelligence technology (AI) such that the results of the solution can be understood by humans.
Exponential Smoothing	A technique for smoothing time series data using the exponential window function. Whereas in the simple moving average the past observations are weighted equally, exponential functions are used to assign exponentially decreasing weights over time. (See autoregressive integrated moving averages)

Exponentially Smoothed State Space Models	A family of forecasting methods which are based on the key property that forecasts are weighted combinations of past observations
Formal Logic	A form of symbolic reasoning that represents knowledge in declarative form as a set of facts and rules.
Free Form Deformation	A geometric technique commonly used in deformable image registration which involves the optimization of an image similarity metric based on changes in a grid of control points overlaid on an image.
Generalised Additive Models	A generalized linear model in which the linear predictor depends linearly on unknown smooth functions of some predictor variables, and interest focuses on inference about these smooth functions.
Generalised Boosting	Repeatedly fit many decision trees to improve the accuracy of the model. For each new tree in the model, a random subset of all the data is selected using the boosting method.
Genome Engineering	Precise editing of the genome of a living cell, thereby altering its function.
Hidden Markov Models	Probabilistic frameworks where the observed data are modeled as a series of outputs (or emissions) generated by one of several (hidden) internal states.
Hounsfield Units	Make up the grayscale in medical CT imaging. It is a scale from black to white of 4096 values (12 bit) and ranges from -1024 HU to 3071 HU (zero is also a value).
Inertial Measurement Unit	An electronic device that measures and reports a body's specific force, angular rate, and sometimes the orientation of the body, using a combination of accelerometers, gyroscopes, and sometimes magnetometers.
K-Nearest Neighbour	A simple algorithm that stores all available cases and classifies new cases based on a similarity measure.
Knowledge Graphs	Acquires and integrates information into an ontology and applies a reasoner to derive new knowledge. In other words, a programmatic way to model a knowledge domain with the help of subject-matter experts, data interlinking, and machine learning algorithms.
Knowledge Representation	Representing information about the world in a form that a computer system can utilize to solve complex tasks such as diagnosing a medical condition or having a dialog in a natural language.
Logistic Regression	A statistical model that in its basic form uses a logistic function to model a binary dependent variable. Used when the response variable is categorical in nature.
Machine Learning	Machine learning builds a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed.
Manifold learning	An approach to non-linear dimensionality reduction. Algorithms for this task are based on the idea that the dimensionality of many data sets is only artificially high.
Naïve Bayes	A family of simple "probabilistic classifiers" based on applying Bayes' theorem with strong independence assumptions between the features.
Natural Language Processing	A subfield of linguistics, computer science, information engineering, and artificial intelligence concerned with the interactions between computers and human (natural) languages, in particular how to program computers to process and analyse large amounts of natural language data.
Neural Network	See artificial neural networks, convolutional neural networks.
Ontology	A collection of concepts with properties including the relationships between the concepts

Optimisation	The process of setting decision variable values in such a way that the objective in question is optimized. The optimal solution is a set of decision variables that maximizes or minimizes the objective function while satisfying the constraints.
Patch Based Local Weighted Voting Procedure	Image segmentation method which combines information from multiple labelled atlas images which have been deformably mapped to a target image. The method selects patches in the same location in each mapped image, compares their voxel intensities to the same patch in the target image, and then weighs those patches by similarity. These weightings are then used to propagate and fuse tissue labels from the atlas images.
Pathling	An advanced analytics service that exploits the standardised medical data https://pathling.csiro.au.
Poisson Regression	A generalized linear model form of regression analysis used to model count data and contingency tables.
Predictive Analytics	A variety of statistical techniques from data mining, predictive modelling, and machine learning, that analyse current and historical facts to make predictions about future or otherwise unknown events.
Principal Component Analysis	A dimensionality-reduction method that is often used to reduce the dimensionality of large data sets, by transforming a large set of variables into a smaller one that still contains most of the information in the large set.
Query Chain Transformation Model	A model that takes a user's search query, transforms it through a chain of operations, into a more effective query which is then sent to the search engine.
Query Performance Prediction Model	A model that predicts the effectiveness a user's search query, when explicit relevance information is not available.
Random Forest	An ensemble learning method for classification, regression and other tasks that operate by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes or mean prediction of the individual trees. (See also decision trees)
Recurrent Neural Network	A class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence. This allows it to exhibit temporal dynamic behaviour.
Regression	A set of statistical processes for estimating the relationships between a dependent variable and one or more independent variables.
Regression Trees	Decision trees where the target variable can take continuous values (typically real numbers).
Semantic Representation	An abstract (formal) language in which meanings can be represented.
Serverless Computing	A cloud computing execution model in which the cloud provider runs the server, and dynamically manages the allocation of machine resources.
Shrimp	A free SNOMED CT search, visualisation, and exploration tool.
Statistical AI	'Learns' relationships from modelling existing data sets.
Support Vector Machine	Supervised learning models with associated learning algorithms that analyse data used for classification and regression analysis.
Symbolic AI	The representation or encoding of human knowledge into a form of known facts and/or rules.
Time Series	A time series is a series of data points indexed in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time, forming a sequence of discrete-time data.
T-distributed Stochastic Neighbour Embedding	A technique for dimensionality reduction that is particularly well suited for the visualization of high-dimensional datasets.
Unsupervised Learning	Looks for previously undetected patterns in a data set with no pre-existing labels and with a minimum of human supervision.
Variant Effect Predictor	An annotation tool to determine functional effects of gene variants, CSIRO has implemented a cloud-native adaptation: serverless VEP (sVEP).
VariantSpark	An adaptation of the Random Forest implementation using Apache Spark to scale to high-dimensional data, such as genomics.

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

Contact us

1300 363 400
+61 3 9545 2176
csiroenquiries@csiro.au
csiro.au

For further information

A**ustralian e-Health Research Centre**
David Hansen
+61 0419 646 108
David.Hansen@csiro.au
aehrc.com