THE PROMISE OF AI & ML IN HEALTHCARE

OPPORTUNITIES, CHALLENGES, AND VENDOR LANDSCAPE

CHILMARK RESEARCH

MARKET SCAN REPORT

ABOUT **CHILMARK RESEARCH**

OUR MISSION

Improve the delivery *and* experience *of care through the effective adoption and use of information technology.*

Chilmark Research is a global research and advisory firm whose sole focus is the market for healthcare IT solutions. This focus allows us to provide our clients with the most in-depth, objective research on the critical technology and adoption trends occurring throughout the healthcare sector. Areas of current research focus include among others: Analytics (including AI/ML), Clinician Network Management, Cloud-computing for Healthcare, Care Management & Coordination, Community Engagement, Consumerism and Patient Engagement, Interoperability and Platforms, Payer-Provider Convergence, Population Health Management, and Value-based Care.

Using a pragmatic, evidence-based research methodology with a strong emphasis on primary research, Chilmark structures its reports to serve the needs of technology adopters, consultants, investors and technology vendors. In addition to reports for the general market, Chilmark Research conducts research for clients based on their specific needs as they align with our own research themes. Such research has included competitive analyses, market opportunity assessments, strategic evaluations of vendors for partnership and/or acquisition, and more.

In 2012, Chilmark Research launched the Chilmark Advisory Service (CAS) in direct response to clients' request for a continuous feed of research on the most pertinent trends in the adoption and use of healthcare IT. This annual subscription service provides varying access to our research reports as well as direct interactions with Chilmark analysts to address specific client needs throughout the year. Partly in response to changing industry dynamics, in 2020 we introduced new tiers to offer access to this service to a broader assortment of organizations than the enterprise vendor or buyer that have been our traditional clients. Please contact us directly for further information about CAS.

Chilmark Research is proud of the clients it has had the pleasure to serve including Abbott Labs, Allscripts, Anthem, athenahealth, Bain, Cerner, Cleveland Clinic, Epic, HCA, Highmark, IBM, Watson Health, Kaiser-Permanente, Mayo Clinic, McKinsey, Medtronic, Merck, Microsoft, and Verizon to name a few. It is our hope that at some future date we will have the pleasure to serve you as well.

CHILMARK RESEARCH LLC

One Beacon Street. 15th Floor
Boston, MA 02108
www.ChilmarkResearch.com
info@chilmarkresearch.com
617.615.9344

TABLE OF CONTENTS

TABLES AND FIGURES

Executive Summary

Artificial intelligence (AI) and machine learning (ML) have become focal points of healthcare innovation in recent years. A generally accepted definition is:

> *The simulation of human intelligence processes by machines, especially computer systems. ... Particular applications of **AI** include expert systems, speech recognition and machine vision.*[1]

In this report we begin with early AI work, such as medical imaging and cardiology, and review different types of AI and ML. We also discuss the challenges that health data create, the growing concern over bias in algorithms and data, and related risks they raise in healthcare settings.

We focus on AI use in healthcare and in vendor solutions, exploring challenges and opportunities in areas of:

- Business Operations, including revenue cycle management, claims management, fraud and payment integrity, hospital operations, supply chains, risk stratification
- Clinical Decision Support, including clinical documentation, medical imaging and pathology
- Population Health Management, which includes risk stratification, population/care management toold, and patient engagement
- Research, Drug Development, and Discovery, including precision medicine
- Patient-Facing Applications, including bots, symptom checkers, laboratory testing, substance abuse, voice assistants, wearables

We also examine the evolution of the market and current approaches for validating the underlying algorithms and impact of AI, data markets, emerging market dynamics, and the cloud. Lastly, we address ethical concerns, human-computer interactions in healthcare, governance issues, and business model implications.

1 https://searchenterpriseai.techtarget.com/definition/AI-Artificial-Intelligence

Key Takeaways

Mainstream AI use in healthcare is still five to seven years away. Substantial bottlenecks in data access and data quality are the norm and will take several or more years to address at scale.

Substantial work in curating, labeling, and cleaning data is required to make datasets market-ready for healthcare applications. Nearly 80% of the work to develop and test an AI/ML algorithm is preparing health data for use in training algorithms.

Concerns for bias in AI are real and challenging to address. Given the lack of inclusion of minorities, women, and children in clinical research, the bias can be amplified if not addressed early in AI system development. Black box algorithms heighten the risks for users, so we will likely see more white box algorithms providing transparency. Regulatory drivers will demand transparency as well.

The most advanced areas for AI/ML in the current healthcare marketplace include medical imaging and business operations. AI's recent evolution in healthcare has been heavily focused on image processing but is expanding to address other clinical use cases. Business operation applications generate less patient risk and have already demonstrated results in cost savings and efficiencies.

A great deal of social and business innovation will be required in the areas of data sharing, governance, and public-private partnerships for AI to scale. Data blocking and well-known data silos in healthcare are significant roadblocks that must be addressed for the market to mature.

Underlying ML models will need to become stronger in causal inference for relevance in clinical care. Most machine learning focuses on associations of large numbers of variables in large datasets, but most medical or clinical use cases require an understanding of more complex causal mechanisms.

Cyber-security and ethics will need to be "baked in" from the start. Neglecting cyber-security and ethical issues in AI will create substantial reputational risks and damage trust on the part of both patients and providers. These issues should be addressed early in the development of solutions.

Why AI/ML Now?

AI and ML have become buzzwords across the economy in recent years amid a great deal of activity across all of the major healthcare segments, from revenue cycle management to precision medicine. A seminal paper on image recognition published in 2012 significantly accelerated innovation in deep learning.[2] Early adoption has focused on imaging because much early AI work centered on databases of images such as Image Net that were created to train deep learning algorithms. Following the trend of early work based on image data, initial AI application in healthcare was in areas such as dermatology and radiology. General interest in AI is driven by the need to analyze large datasets at scale, improve accuracy in diagnosis, improve feedback mechanisms, and reduce clinical and administrative errors.

A recent World Intellectual Property Organization (WIPO) study on AI notes that over half of AI patents have been published since 2013.[3] The ratio of theory to invention skyrocketed from 8:1 in 2010 to 3:1 by 2016. Machine learning has accounted for a third of all patents, and the life sciences and healthcare rank third (12%) of industry sectors with the most AI patents. IBM has received over 8,000 patents and Microsoft is second with nearly 6,000.

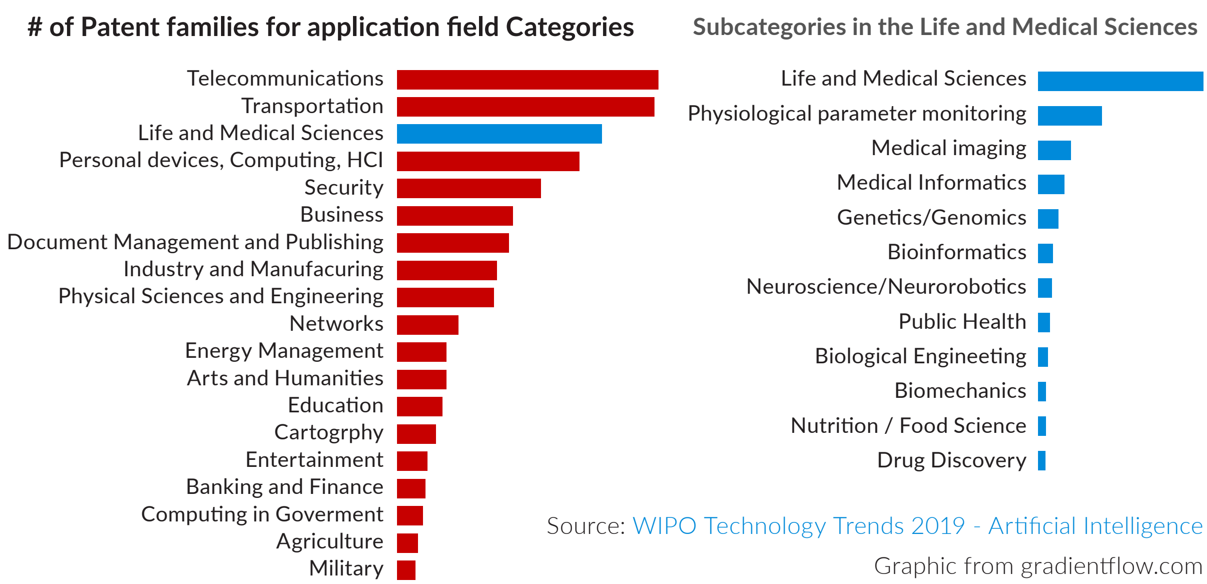

Figure 1: Patent applications by field

At the processing level, we have also seen dramatic improvements in both central processing units (CPU) and graphics processing units (GPU). The vast amounts of unstructured data that AI algorithms can analyze required dramatic improvements in GPUs that are capable of performing the same processing procedures repeatedly on large batches of data in ways that a traditional CPU cannot. Innovation in GPUs has enabled more scalable AI applications capable of processing these large data sets. While CPUs and GPUs are designed for different processing applications, the two are used in tandem in many AI applications. This is part of the technology stack that has enabled AI/ML to scale and enable some cloud-based AI engines to provide analytics services at lower cost.

2 See Eric Topol (2019). Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again. The paper of interest: Krizhevsky, A, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," ACM Digital Library. 2012: NIPS'12 Proceedings of the 2thth International Conference on Neural Information Processing Systems, pp. 1097-1105.

3 https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf

HEALTH DATA AND AI

Much larger volumes of data can be analyzed with AI/ML tools than with traditional statistics. Applying AI/ML to larger longitudinal datasets in healthcare would enable discernment of more nuanced changes in health conditions and discovery of new biomarkers, molecular pathways, and anomalies unobservable with traditional analytics. However, using past data for predictive purposes is problematic. For many conditions, treatment patterns shift every two to three years, quickly rendering past patterns antiquated. Indeed, recent research has found that the **half-life of health data is just four months**,[4] leading to their conclusion that small, more recent datasets can have more predictive power than larger volumes of data.

Claims data are of great interest due to their longitudinal nature and large amount of data which allow data scientists to see adherence patterns and outcomes. To unlock the most value, claims data need to be linked to EHR and social determinants data such as census data. One can utilize far more variables when applying AI to these linked datasets rather than using standard statistics, which can shed light on non-linear interactions between these variables that standard statistics may miss. This is also where we see how AI can be applied **to detect bias** in treatment patterns or reduce bias in randomized trials.[5]

It's worth noting that a great deal of healthcare AI work is based on a small number of datasets. MIMIC[6] is a set of 46,000 de-identified medical records from Boston's Beth Israel Deaconess Medical Center used by more than 12,000 researchers since 2009. The only free dataset sufficiently detailed to be readily used in AI research, MIMIC has played an equivalent role in health and medicine to ImageNet's in the image recognition space. While the large dataset stems from an ethnically diverse area and population, practice patterns of physicians at that hospital could introduce bias, a problem that has plagued IBM's Watson for Oncology which trained algorithms on data from Memorial Sloan Kettering.[7] (In 2018, Philips, too, opened up a dataset of 140,000 patients treated across 200 hospitals in 2014-15. Philips's dataset has fewer progress notes, less unstructured text, and variable quality of data across hospitals, so it's considered less useful than MIMIC despite its larger population size.)

These examples highlight a major challenge to implementing AI/ML in healthcare: the substantial amount of work required to clean data. All of these datasets have substantial problems with missing data, corrupt values, and typographic errors and require a great deal of work to label. Tools developed to address some of these issues include Holoclean, for automating error detection, and Snorkel, for automating training set creation and data labeling.[8]

TYPES OF AI/ML

AI/ML utilize a number of different types of algorithms and models for analyzing data. Below we provide a brief overview of the algorithms and models included in the general field. Machine learning is a method that utilizes analytical algorithms to extract insights or features from data inputs. These algorithms must be trained on datasets that include demographic data, medical histories, disease specific data, laboratory data, diagnostic imaging, and clinical symptoms. Increasingly, behavioral data from wearables, social determinants, and even -omic data (genomic, microbiomic, transcriptomic, etc.) data is also used. These background data can then be combined with outcomes data or real world data over time to show survival times, progression of disease (quantitative), and other correlations for predictive analytics. Various types of AI models and algorithms can then be applied to the data.

4 https://www.sciencedirect.com/science/article/pii/S138650561730059X
5 Thesmar, D et al.2019. Combining the Power of Artificial Intelligence with the Richness of Healthcare Claims Data: Opportunities and Challenges. PharmacoEconomics, 37:745-752.
6 https://mimic.physionet.org (see discussion in Rebecca Robbins, "How patient records from one Boston hospital fueled an explosion in AI research in medicine." STAT+, July 12, 2019.)
7 See Robbins (ibid)
8 https://www.oreilly.com/ideas/how-new-tools-in-data-and-ai-are-being-used-in-health-care-and-medicine

Appendix: Patient Safety Issues

Distributional shift

- Has the system been tested in diverse locations, underlying software architectures (such as electronic health reccords), and populations?
- How can we be sure the training data matches what we expect to see in real life and does not contain bias?
    - How can we be confident of the quality of the 'labels' the system is trained on?
    - Do the 'labels' represent a concrete outcome ('ground truth') or a clinical opinion?
    - How has imbalance in the training set been addresed?
- How is the system going to be monitored and maintained over time to adjust for prediction drift?

Insensitivity to impact

- Does the system adjust its behavior ('err on the side of caution') where there are high impact negative outcomes?
- Can the system identify 'out of sample' input and adjust its confidence accordingly?

Black box decision-making, unsafe failure and automation complacency

- Are the system's predictions interpretable?
- Does it produce an estimate of confidence?
- How is the centainty of prediction communicated to clinicians to avoid automation bias?

Reinforcement of outmoded practice and self-fulfilling predictions

- How can it accomodate breaking changes to clinical practica?
- What aspects of existing clinical practice does this system reinforce?

About the Author

Dr. Ranck has nearly 30 years of experience working in the global health arena and has helped lead a number of major health technology initiatives throughout his career. Author of two books on digital health, he is a globally recognized thought leader on digital health and has been listed in the "Always On" top 100 minds in Global mHealth (2013). His past clients have included Humana, TM Forum, CLSA, T-Systems, Stanford University's School of Medicine, UC Berkeley, the UN, and ARM to name a few. He has been a frequent advisor to large healthcare companies and startups focused on providing more patient-centric care and transitioning to value-based care. In the past he has been appointed as a member of an Institute of Medicine Committee on ICTs in global health/violence prevention and helped launch a major global eHealth initiative with the Rockefeller Foundation. He has been a frequent keynote speaker at health IT conferences and recently organized and chaired the Healthcare Blockchain Summit (2017-18).

Jody has written and worked extensively on mobile innovations, the Internet of Things (IoT), wearables, blockchain and the analytics market in healthcare. He is also working with cutting edge startups on next generation biosensor platforms, patient generated data for clinical research, and emerging blockchain applications in healthcare. His education includes a Doctorate in Public Health (University of California, Berkeley), MA in International Relations and Economics (Johns Hopkins University) and a BA in Biology (Ithaca College).

CHILMARK
RESEARCH

ONE BEACON STREET

BOSTON, MA 0210

WWW.CHILMARKRESEARCH.COM

INFO@CHILMARKRESEARCH.COM

PH: 617.615.9344