DATA & AI MAGAZINE

ISSUE 8

SPECIAL FOCUS:
LIFE SCIENCES

CREATING VALUE FROM AI IN PHARMA

WITH

DR TOMISLAV ILICIC

STEVE ORRIN:
ENTERPRISE
AI SECURITY

REX VANHORN:
SEMANTICS & GLOBAL
QUALITY DATA

NICOLE JANEWAY BILLS:
16 BOOKS TO TRANSFORM
DATA INTO WISDOM

INSIDE ISSUE #8

DATA & AI MAGAZINE
ISSUE 8
SPECIAL FOCUS: LIFE SCIENCES
CREATING VALUE FROM AI IN PHARMA
WITH
DR TOMISLAV ILICIC
STEVE ORRIN: ENTERPRISE AI SECURITY
REX VANHORN: SEMANTICS & GLOBAL QUALITY DATA
NICOLE JANEWAY BILLS: 16 BOOKS TO TRANSFORM DATA INTO WISDOM

CONTRIBUTORS
Tomislav Ilicic
Rex VanHorn
Philipp Diesinger
Gabriell Fritsche-Máté
Andreas Thomik
Stefan Stefanov
John O’Gorman
Walid Mehanna
Nicole Janeway Bills
Steve Orrin
Daniel Martinez
RuZe Liang
Kirk Marple
Anthony Alcaraz
Frank Stadler
Ziad Al-Ziadi
Dr Ana Moya
Francesco Gadaleta

EDITOR
Damien Deighan

DESIGN
Imtiaz Deighan
imtiaz@datasciencetalent.co.uk

DATA SCIENCE TALENT

Data & AI Magazine is published quarterly by Data Science Talent Ltd, Whitebridge Estate, Whitebridge Lane, Stone, Staffordshire, ST15 8LQ, UK. Access a digital copy of the magazine at datasciencetalent.co.uk/media.

DISCLAIMER

The views and content expressed in *Data & AI Magazine* reflect the opinions of the author(s) and do not necessarily reflect the views of the magazine, Data Science Talent Ltd, or its staff. All published material is done so in good faith.

All rights reserved, product, logo, brands and any other trademarks featured within *Data & AI Magazine* are the property of their respective trademark holders. No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form by means of mechanical, electronic, photocopying, recording or otherwise without prior written permission. Data Science Talent Ltd cannot guarantee and accepts no liability for any loss or damage of any kind caused by this magazine for the accuracy of claims made by the advertisers.

WELCOME TO ISSUE 8 OF *DATA & AI MAGAZINE*

As we enter our third year we have decided to rename our publication ‘*Data & AI Magazine*’, a deliberate evolution from our previous title, ‘The Data Scientist.’ This change reflects the diverse community that shapes our readership, encompassing not only data scientists (who form 60% of our audience) but also an interesting mix of business leaders and many other kinds of data and AI practitioners.

Our commitment to a wide spectrum of content remains unchanged. We will continue to explore a broad array of topics, from traditional data analytics to cutting-edge generative AI and everything in between. This blend of different technical and non-technical insight will ensure that we remain a vital resource for all professionals in the data and AI community, regardless of their specialty and job role.

AI in Life Sciences Focus

In this issue, we spotlight the transformative impact of AI in the life sciences, a sector well poised to capitalise on the rapid era of technological advancement we are currently experiencing. Our cover story features Tomislav Ilicic from Accenture, with his superb piece on ‘Creating Value at Scale with AI in Pharmaceuticals,’ where he explores how AI-driven strategies are impacting the entire pharma life cycle. Rex VanHorn of Boehringer Ingelheim highlights innovative approaches to enhancing data interpretation in his article ‘Using Semantics to Find Trends in Global Quality Data’.

Additionally, John O’Gorman examines ‘The Role of Data Science, Machine Learning, and AI in Reducing the Clinical Burden’, showcasing AI’s potential to streamline clinical trials and improve outcomes. Each article underscores the critical role of AI in pushing the boundaries of what’s possible in healthcare.

Harnessing Advanced Data & AI Approaches

We have several contributions that look at the practical challenges and strategic implementations of cutting-edge AI technologies. Ziad Al-Ziadi articulates the complexities of transitioning large language models from theoretical constructs to valuable products in ‘LLMs Aren’t Products’. Kirk Marple explores the pivotal role of knowledge graphs and GraphRAG in enhancing model performance in his insightful article. Ana Moya offers a deep dive into ‘Comprehensive Approaches and Methodologies for Modern and Strategic Content Publishing,’ addressing the evolving landscape of digital content. Meanwhile, Steve Orrin from Intel confronts the pressing issue of ‘Understanding and Mitigating AI Security Threats in the Enterprise,’ spotlighting the critical need for robust security measures in AI deployments. Each piece underscores the necessity of bridging theory with practical application in the realm of artificial intelligence.

Q4 Conferences to Watch Out For

In the upcoming fourth quarter, I am looking forward to speaking at two significant events in Germany and the UK.

First, the ai:NOW conference in Munich on October 9th and 10th will assemble over 80 participants and 10 sponsors for a comprehensive exploration of AI innovations. This event uniquely focuses on practical outcomes, enabling attendees to develop strategic AI implementation tools such as playbooks and roadmaps.

You can find out more about that conference here:
cxoportals.com/event/ainow-conference-dach/

Following this, I will be at the AI World Congress in London on November 27th and 28th with over 250 participants. This premier gathering is a nexus for AI professionals worldwide, offering a great opportunity to network, share knowledge, and discover the latest advancements in AI. For more details on the AI World Congress, visit aiconference.london

The value of attending these events in person lies in the irreplaceable interactions and collaborations that only face-to-face meetings can foster. The more compact size of these events is ideally suited for networking, offering each participant direct access to industry leaders, speakers and peers, allowing stronger collaboration and exchange of ideas.

Damien Deighan
Editor

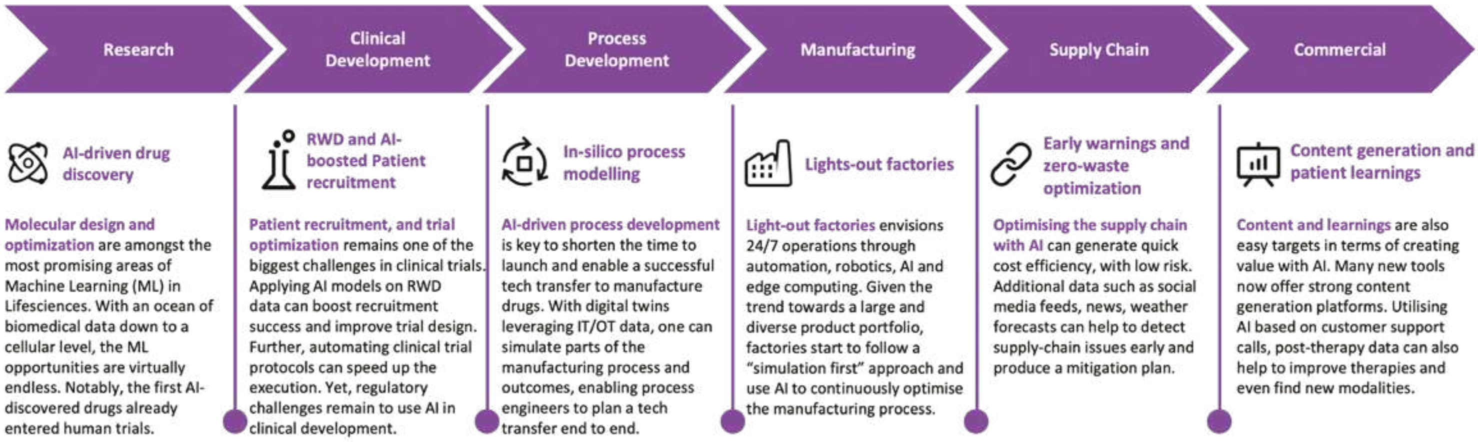
DR TOMISLAV ILICIC (TOMI) is a director at Accenture, leads the ASG Life Sciences Data & AI practice, and is a top voice for pharma and technology companies. Over the last 10 years, he has held multiple roles, including Data Engineer, Machine Learning Engineer, Data Scientist, and Product Owner. He also holds a Bioinformatics degree from the University of Cambridge, UK. Tomi's mission is to accelerate the AI transformation journey of life science companies by leveraging Accenture's powerhouse, a thought-out and C-level prioritised data strategy, as well as rallying excitement and motivation amongst life science experts to adopt new innovative tools.

CREATING VALUE AT SCALE WITH AI IN PHARMACEUTICALS

In recent years, artificial intelligence (AI), particularly generative AI models like GPT-4, PaLM2 and LLaMA captured the interest of virtually any CEO and spiked expectations to transform entire industries. Within a year, ChatGPT became mainstream, and is already used in everyday tasks, making the entire world appreciate the immense potential and power of AI.

This global awakening to AI's capabilities has set the stage for transformative changes across all industries, but probably the most disruptive will be within banking, insurance and pharmaceuticals. Big Pharma companies carry a 'big bag'. This bag contains outdated IT systems, complex and bespoke data, strict regulations, millions of manually drafted reports, and increasing pressure not just to identify and produce new groundbreaking medicine but to do so cheaper and faster.

In seeking to serve society by enabling better patient outcomes, Big Pharma is embracing an AI-enabled transformation (Figure 1). The AI-enabled transformation extends beyond a digital core but will bring about a total enterprise reinvention and enable data-driven insights to be gained by the wealth of bespoke data residing in the 'big bag'. One key difference is this transformation will (sometimes) be enthusiastically driven both top-down and bottom-up, compared to introducing other tools, moving to the cloud and various other technology transformations.


Research
Clinical Development
Process Development
Manufacturing
Supply Chain
Commercial
AI-driven drug discovery
RWD and AI-boosted Patient recruitment
In-silico process modelling
Lights-out factories
Early warnings and zero-waste optimization
Content generation and patient learnings

Molecular design and optimization are amongst the most promising areas of Machine Learning (ML) in Lifesciences. With an ocean of biomedical data down to a cellular level, the ML opportunities are virtually endless. Notably, the first AI-discovered drugs already entered human trials.

Patient recruitment, and trial optimization remains one of the biggest challenges in clinical trials. Applying AI models on RWD data can boost recruitment success and improve trial design. Further, automating clinical trial protocols can speed up the execution. Yet, regulatory challenges remain to use AI in clinical development.

AI-driven process development is key to shorten the time to launch and enable a successful tech transfer to manufacture drugs. With digital twins leveraging IT/OT data, one can simulate parts of the manufacturing process and outcomes, enabling process engineers to plan a tech transfer end to end.

Light-out factories envisions 24/7 operations through automation, robotics, AI and edge computing. Given the trend towards a large and diverse product portfolio, factories start to follow a "simulation first" approach and use AI to continuously optimise the manufacturing process.

Optimising the supply chain with AI can generate quick cost efficiency, with low risk. Additional data such as social media feeds, news, weather forecasts can help to detect supply-chain issues early and produce a mitigation plan.

Content and learnings are also easy targets in terms of creating value with AI. Many new tools now offer strong content generation platforms. Utilising AI based on customer support calls, post-therapy data can also help to improve therapies and even find new modalities.

FIGURE 1: AI-enabled transformations across the pharmaceutical value chain to bring new innovative medicine - faster and cheaper.

AI IMPLEMENTATION ACROSS THE PHARMACEUTICAL VALUE CHAIN

The pharmaceutical industry presents a compelling example of AI's impact in 2024. Use cases across the pharma value chain demonstrate AI's versatility and potential for innovation (Table 1). From drug discovery and patient diagnostics to supply chain optimisation and personalised medicine, AI is reshaping how the industry operates. These use cases not only enhance efficiency and reduce costs but also pioneer novel solutions to long-standing challenges, thereby creating substantial value for the industry and its consumers. Below is a high-level summary of the most common use cases across the pharma value chain.

EARLY RESEARCH AND DRUG DISCOVERY

Traditionally, drug discovery is done by utilising prior knowledge and laboratory experimentation to understand the underlying biological process of the diseases. This is done to identify new targets and perform high-throughput screenings (HTS) of large chemical libraries to select lead compounds and optimise those to satisfy a set of specific properties, such as solubility, toxicity, pharmacokinetics and other effects. It can take 3-5 years until a lead compound reaches Phase I of clinical trials. Although this approach led to many successful new drugs, it becomes increasingly difficult to find and develop innovative medicine faster and without exploding costs. Therefore, pharma companies have started to invest heavily in AI-driven drug discovery to boost their early-stage drug pipeline. Several biotech companies have already embraced AI at their core, and by doing so, they have managed to shorten the time it takes to reach clinical trials. AI-driven drug discovery can have many forms, and primarily aims to generate molecular designs, training on public and proprietary chemical data, omics, clinical trial outcomes, literature, and other third-party data. AlphaFold, for example, can predict the 3D folding of a protein to help understand its biological function, which used to be a very experiment-heavy and expensive exercise. Notably, there are advancements in leveraging GenAI to enable researchers with little computer science background to query their data or summarise publications using natural language. It is still unclear when the first AI-discovered drug will be approved, but the time is ripe, and pharma companies could develop innovative medicine for patients at a fraction of the current cost and time.

CLINICAL OPERATIONS

For drugs that have reached clinical trials, other challenges emerge, mostly around completing the clinical stages more quickly, cheaply, and with higher success ratios.

Pharmaceuticals struggle to reduce the time-to-market for a new drug because clinical trials are very complex, often spanning multiple countries, relying on third parties such as clinical research organisations (intermediary between pharma and clinics), clinical sites and the struggle to find (and retain) the right subjects and execute the clinical protocol as intended. Pharma companies are turning to AI support, and several tools already exist: some are built on large language models to identify safety and efficacy information from clinical trial data or summarise them to ease the understanding of both doctors and subjects. Many efforts also go into predicting the eligibility of potential subjects and their risk of dropping out of the trial to increase the overall recruitment success. Drafting high-quality trial protocols enabling real-time monitoring using wearables is also a high-potential area to cut both time and costs and ultimately help market new treatments.

TECHNICAL OPERATIONS

Not unique to the pharmaceutical industry, manufacturing and supply chains are rich in opportunities as they come with a lot of data and challenges. One of the major challenges is the trend towards personalised treatments (such as cell and gene therapies), which results in lower production volumes as they are more targeted. This is a big shift from so-called blockbuster drugs, which can be administered to millions of patients. That's why manufacturing sites need to become more agile and modular to produce different drugs quickly without much effort. By training models on data from laboratory equipment, metadata of the batches, and technical documentation of the process, they can learn to predict the optimal parameters of new batches and make sure that the operators can generate more drug substance output, with less cost.

Adding a semantic layer on top of IT/OT data allows to simulate the manufacturing process, and to fine-tune a GenAI chatbot that can interact with instruments in real time. This can help the operators to quickly fix quality problems, and help them to plan a tech transfer from development to manufacturing, which is a very manual and difficult task.

COMMERCIAL

With the advent of generative AI, many commercial functions such as branding, marketing and content production will be completely disrupted. This is mostly because of the strength of GenAI in being able to create new content, such as images and text. This is already affecting the many designer and marketing agencies that pharma companies are leveraging, as at least part of their work can be automated. A new trend is emerging, which is around creating country-specific product campaigns entirely driven by AI automate, including the Medical Legal Review. But, it can also be used to better understand the customer segments in the first place, and even make informed investment decisions about the early research pipeline. Further, clustering can be used to determine customer archetypes' behaviour and optimise distribution channels.

POST-MARKET SURVEILLANCE

Patients and subjects are monitored after they have received a treatment, and so-called real-world data is collected, which can be biomarkers such as blood pressure, heart rate, glucose levels, DNA tests and many others. Such data is already being used to identify patient factors that might trigger an adverse reaction and, in turn, can be used to predict adverse events for new drugs and ultimately make them safer for consumption. It can also be used for apps on wearable devices to trigger notifications to healthcare professionals, informing them about the potentially dangerous and life-threatening health status of the patient.

> *From drug discovery and patient diagnostics to supply chain optimisation and personalised medicine, AI is reshaping how the industry operates.*

STAGE	USE CASE	DESCRIPTION
Research & Development	Molecular Design and Optimisation	Generative machine learning is used to design and identify novel molecular structures, leveraging public and proprietary chemical structures, and data from bioassays.
		Using chemogenomics and bioassays data can train models to predict Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADME/T) properties of lead compounds and design the formulation.
	Predictive Toxicology	Leveraging a range of chemical data, one can predict the toxicity of compounds and boost safety and decrease risks during clinical trials.
	Regulatory Compliance Reporting	Auto-generate and pre-review regulatory documents, leveraging historic filing submissions, clinical data and Structured Content Management Platforms.
	Subject Recruitment	Predict success rates of subjects during clinical trials by training models on clinical history, demographics, genomics, wearables sensor and other data.
	Real-Time Data Monitoring	Monitoring sensor data in real-time to ensure the safety and efficacy of trials and predict adverse events in advance.
	Adaptive Trial Design	Using standard operating procedures, historic clinical trial protocols, outcomes and other data to auto-generate trial protocols dynamically for efficiency and effectiveness, and reduce the overall trial risk.
Manufacturing & Supply Chain	Deviation Detection	Leverage IT/OT data, standard operating procedures and historic deviations to identify trends and potential issues during manufacturing.
	Process Optimisation	Predict yield or other desired properties using IT/OT data, standard operating procedures and quality attributes of historic batch runs.
	Tech Transfer Optimisation	Ease the transfer of technologies with 3D visualisations of factory floors and embedded digital twins of all instruments (leveraging IT/OT data), and technical documentation. This enables planning and simulation of a manufacturing process in detail, accelerating the overall setup and saving resources.
	Demand Forecasting	Predicting drug demand and supply shortages, to optimise inventory for clinical and commercial batches, using internal Enterprise Resources Planning data, but also public data such as news.
Commercial	Customer Behaviour and Segmentation	Understand the needs and behaviours of customers and end-consumers (patients) to make informed decisions about investments and targeted campaigns.
	Targeted Marketing	Analysing customer data and purchasing behaviours to create personalised marketing strategies.
	Price Optimisation	Optimise gross-to-net leveraging forecasting models, competitors pricing, cost, regulatory and patient data.
Post-Market Surveillance	Adverse Event Monitoring	Continuously monitoring drug safety in the market using AI-driven pharmacovigilance.
	Real-World Evidence Generation	Analysing real-world data to gain insights into drug performance and patient outcomes.

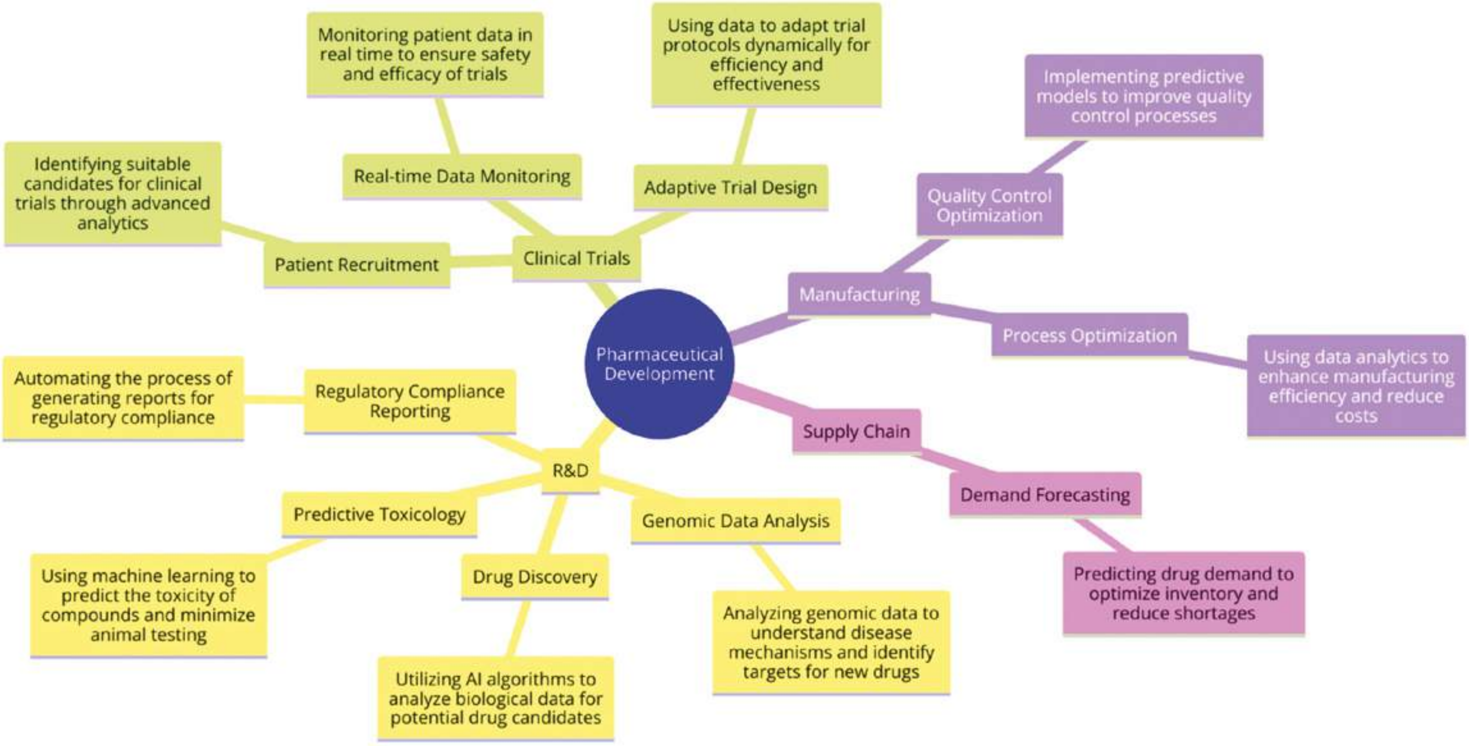

Pharmaceutical Development
Clinical Trials
Real-time Data Monitoring
Monitoring patient data in real time to ensure safety and efficacy of trials
Adaptive Trial Design
Using data to adapt trial protocols dynamically for efficiency and effectiveness
Patient Recruitment
Identifying suitable candidates for clinical trials through advanced analytics
Manufacturing
Quality Control Optimization
Implementing predictive models to improve quality control processes
Process Optimization
Using data analytics to enhance manufacturing efficiency and reduce costs
Supply Chain
Demand Forecasting
Predicting drug demand to optimize inventory and reduce shortages
R&D
Regulatory Compliance Reporting
Automating the process of generating reports for regulatory compliance
Predictive Toxicology
Using machine learning to predict the toxicity of compounds and minimize animal testing
Drug Discovery
Utilizing AI algorithms to analyze biological data for potential drug candidates
Genomic Data Analysis
Analyzing genomic data to understand disease mechanisms and identify targets for new drugs

TABLE 1: A high-level view of the most common AI use cases across the pharma value chain (not exhaustive).

'SCALING AI': BEYOND TECHNOLOGY

Although there is a wealth of use cases, each of them has one critical challenge: to generate measurable or observable value at scale. Scaling AI transcends mere technological expansion. It involves integrating AI into the core operational, strategic, and decision-making processes of an organisation. This scaling is not just a matter of deploying more algorithms or accumulating data. It's about embedding AI into the fabric of the organisation, influencing everything from business strategies to employee workflows and customer interactions.

To effectively scale AI, organisations must consider a spectrum of aspects, not just technological. This includes developing a robust data infrastructure, fostering a culture of data literacy, ensuring responsible AI practices, and aligning AI initiatives with business objectives. It's about creating an ecosystem where AI can thrive, supported by skilled personnel, clear governance, and strategic vision. Scaling AI, therefore, is as much about people and processes as it is about technology. Let's talk about the six key factors that prevent scaling AI.

1. Data, Data and Again Data: The Cornerstone of Pharma Innovation

Data, the most critical element in the pharmaceutical industry, presents significant challenges due to its historical accumulation and complex system landscapes. Pharma companies have an ocean of data, but as you can imagine, finding the treasure in the ocean is a difficult task. The data ecosystem in the pharmaceutical industry is complex and rapidly evolving, driven by the increasing volume and variety of data from research, clinical trials, and real-world evidence (common data systems in Table 2). It can cover hundreds of different data systems, each with global and local instances and rollouts, and sometimes outdated technologies, which not only creates a complex web but also poses significant challenges to standardisation and modularity – a nightmare for IT, labour-intensive work for data engineers, value stagnation for data scientists and a box of pandora for executives. Many pharma companies are undergoing major transformations to simplify their system landscape, and continuous investments will be needed to identify the treasure and piles of gold while not drowning in data. It's a paradox of not having enough data but also too much data. Companies didn't do this on purpose but rather as a result of decades of changes and accumulation of systems. Removing what's bad whilst ensuring that everything else runs is a very challenging task. It goes beyond technology and includes having the end-users follow FAIR (findability, accessibility, interoperability, and reusability) data principles and implementing them bottom-up everywhere and every time. Investments into modernising the IT and data landscape will yield high returns when coupled with the most strategic use cases/business problems (e.g. AI-driven drug discovery).

VALUE CHAIN STAGE	COMMON IT AND DATA SYSTEMS
Research & Development	Laboratory Information Management Systems (LIMS)
	Electronic Lab Notebooks (ELN)
	Clinical Trial Management Systems (CTMS)
	Bioinformatics Systems
	Cheminformatics Systems
	Regulatory Information Management System (RIM)
Manufacturing	Manufacturing Execution Systems (MES)
	Enterprise Resource Planning (ERP)
	Quality Management Systems (QMS)
	Process Control Systems
Supply Chain	Supply Chain Management Systems (SCMS)
	Inventory Management Systems
	Distribution Management Systems
	Enterprise Resource Planning (ERP)
Sales and Marketing	Customer Relationship Management (CRM)
	Sales Force Automation Tools
	Digital Marketing Platforms
	Market Research and Data Analytics Tools
Post-Marketing Surveillance	Pharmacovigilance Systems
	Patient Registry Databases
	Adverse Event Reporting Systems
	Real World Evidence (RWE) Data Analytics Tools

TABLE 2: non-exhaustive view of most critical IT systems in life science companies

2. Platforms: The Ecosystem of Data Management

Another key challenge is to manage data integration from these disparate sources, ensuring data quality and compliance with stringent regulatory standards and addressing data privacy concerns. The proliferation of diverse data platforms creates another layer of complexity. This multiplicity not only fragments the data landscape, creating silos and complicating the integration process, but also imposes significant costs and maintenance burdens. The key to navigating this complexity lies in adopting a holistic data management strategy. This strategy should prioritise robust data integration tools capable of harmonising disparate data sources, thereby ensuring data integrity and facilitating comprehensive analysis. Additionally, a shift towards cloud-based unified platforms could offer more seamless integration and scalability. It's essential to involve data scientists and IT experts in this process, leveraging their insights for optimal platform integration and utilisation. Regular audits and a focus on scalable architectures can further streamline operations. Ultimately, the goal is to transform this multitude of individual platforms from a fragmented puzzle into a coherent, efficient, and insightful data ecosystem, driving informed decision-making and innovation in the pharmaceutical sector.

3. Data Products and the Semantic Layer

Just because data sits in a data lake, it doesn't mean that it can be used to solve business problems. Data engineers face the monumental task of creating 'consumable' or business-relevant data, by cleansing, normalising and transforming data, and describing the context found from the tangled web of platforms and raw sources. To do so more systematically, data mesh established itself as a popular concept, comprising a decentralised data architecture that treats data as a product, managed by domain-oriented teams who own and operate their data independently, while adhering to common standards and governance. It's like a well-documented storage of goods grouped together (e.g. parts of a car engine). But even if a large and well-maintained storage of data products does exist, solving complex problems (e.g. building a car) requires knowing the relationship between data points (e.g. what parts are required to build the car). To do so, a so-called semantic layer can be added on top, which allows models to learn relationships between data products and create accurate responses/predictions. This makes the difference between AI being able to generate emails, versus regulatory filings containing confidential and complex information. Scaling the creation of business-relevant data products and a semantic layer requires a strong emphasis on involving key SMEs who can help the data teams in identifying and understanding the raw data. Their involvement is critical, yet constrained by their own priorities and workload. With the right team, clear focus and a strong data platform in place, one can implement data products in a factory approach, creating large amounts of high-quality data at a fast pace.

4. MLOps: Optimising Model Deployment and Evaluation

Hundreds of new foundation models have emerged, most of which are specialised for specific use cases. Assuming data products are available, it now requires a highly parallelised evaluation approach to pick out the best-performing model for a specific use case. This is where MLOps become another critical component of the end-to-end architecture of data analytics at scale. Choosing one MLOps platform and training data scientists and engineers on it can truly accelerate building, deploying, and maintaining models, often referred to as 'industrialised' models. Such tools also help to be on top of costs and consumption, which for large organisations can become substantial. Pick one MLOps platform and train data scientists and engineers in using it, to truly build models at scale.

5. Leading by Value: From End-User to Data Source

Many organisations focus on all the technical aspects above and invest a lot of money into platforms, data curation and models. However, one important aspect often gets neglected: the involvement of end-users very early on and having a detailed understanding of the most valuable projects, capabilities and the estimated impact of the AI solution. So rather than starting with development right away, it can help to take a step back and think critically about what use cases (or groups of use cases, or capabilities) will provide the most value. Many clients have a long backlog of use cases (sometimes in the hundreds) they want to implement, but without a clear focus one can run into the danger of getting stuck in PoC land. There, it can help to estimate ROI and the technical feasibility of each use case early on, in order to better prioritise. In addition, holding design-thinking workshops with a few end-users to create paper and Figma prototypes can speed development massively and create a first solid product fast. This approach is similar to building a house, where seeing the final 3D design helps with planning and optimising the building blocks within. Once it's clear what needs to be built, it's much easier to navigate the complex web mentioned above. This methodology ensures that the development of AI models and data products is aligned with the most critical use cases and end-user needs, ensuring a return on investments, and hence moving away from PoC land.

6. The Human Component in Scaling AI

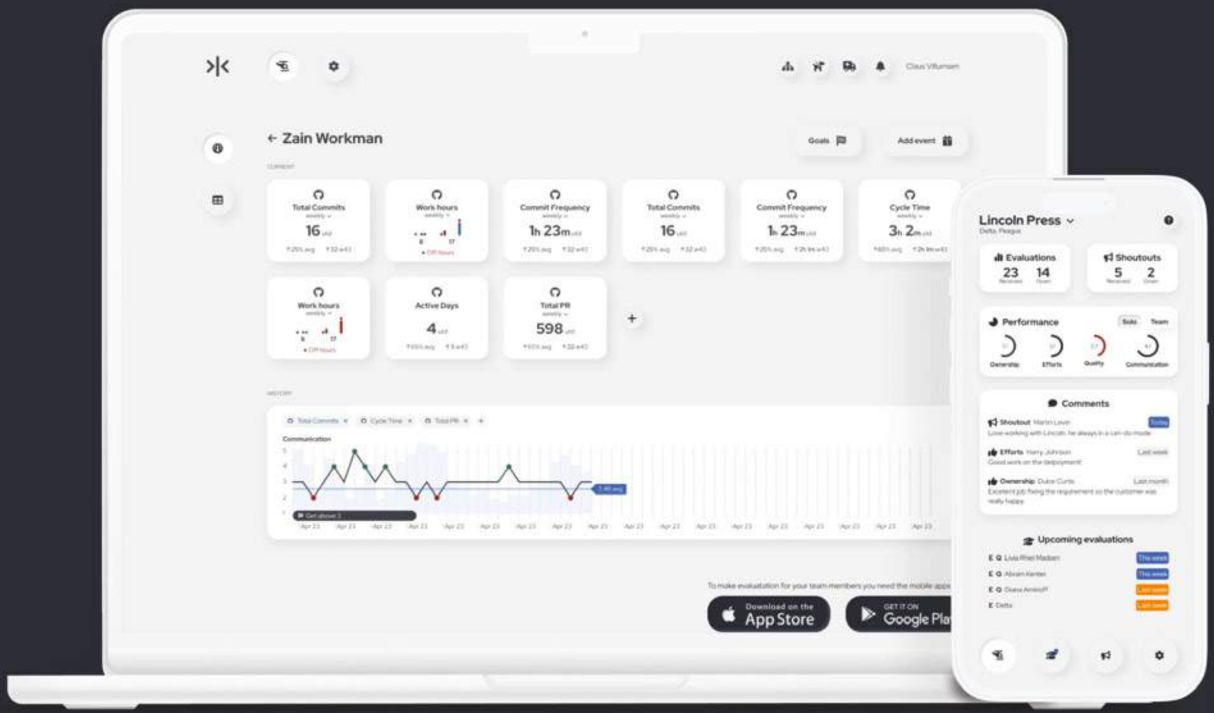
Finally, an important puzzle piece to scale AI in organisations is not only the technology but us humans. It starts with employees entering data in a standardised way, helping data teams to resolve issues, and lastly, defining how the solution can be used effectively day-to-day and training their co-workers to adopt it. Doing this will move the needle. Whoever performs a core-business job, such as researcher, clinician, manufacturing operator, or regulatory/quality expert, and is interested in AI, should invest time to work with data teams. Although it will take time for the two worlds to speak one common language, over time, it can become a win-win situation for both worlds, and will moon-shoot careers once high-quality data products have been built and are used to create measurable value and impact across the organisation.

BOTTOM LINE

In synthesising the future of AI in pharmaceuticals, scaling is the pivotal chapter. It's not merely about enlarging the role of AI but fundamentally transforming how we think about and employ this technology. Scaling AI transcends the technical – it's about weaving intelligent algorithms into the very DNA of pharmaceutical operations and decision-making frameworks. It means moving beyond pilot programs and isolated successes to a state where AI-enabled insights and automation are as ubiquitous and essential as electricity in the industrial age.

To truly scale AI is to embed it into the day-to-day rhythm of the pharma industry, ensuring that from the laboratory bench to the patient's bedside, every process is enhanced by data-driven decision-making. It's the seamless convergence of AI with the core strategic goals of the industry – accelerating drug discovery, personalising patient care, optimising supply chains, and ensuring the safety and efficacy of medications in real-world scenarios.

As we advance, scaling AI will be the barometer of innovation and the catalyst for a new paradigm in healthcare. It promises a future where treatments are not only discovered and developed faster but are also more in tune with the needs of diverse populations. In this future, AI will not just be an adjunct but a fundamental engine of growth, driving the industry towards unprecedented efficiency and effectiveness in serving the global community. This is the essence of scaling AI in pharmaceuticals: it's a journey of transforming potential into real-world impact, ensuring that the power of AI fully realises its promise to revolutionise healthcare for all.


Zain Workman
Goals
Add event
Total Commits
16
Work hours
Commit Frequency
1h 23m
Cycle Time
3h 2m
Active Days
4
Total PR
598
Communication
To make evaluation for your team members you need the mobile app
Download on the App Store
GET IT ON Google Pla
Lincoln Press
Evaluations
23
14
Shoutouts
5
2
Performance
Solo
Team
Ownership
Efforts
Quality
Communication
Comments
Shoutout
Efforts
Ownership
Upcoming evaluations

USING SEMANTICS TO FIND TRENDS IN GLOBAL QUALITY DATA

REX VANHORN has been employed by Boehringer Ingelheim Pharmaceuticals for more than 25 years, having worked as a software developer by trade for more than 20 of those years, specialising in automation through API interfacing and advanced applications. Against his father's wishes, Rex earned a bachelor's degree in Spanish but reentered his dad's good graces shortly thereafter by earning an MBA in Finance, both from The Ohio State University. Rex recently graduated from the University of Georgia with a MSc degree in Artificial Intelligence, where he researched the impact of fine tuning and retrieval-augmented generation on ambiguous question answering using large language models. He is currently pursuing a PhD in AI at UGA, focusing on advanced natural language processing techniques. Parents to four rambunctious but adorable children, Rex and his wife Jackie spend their days on their urban farm and rooting for The Ohio State Buckeyes (while wishing the Georgia Bulldogs well).

In 2022, our organisation was challenged to find an innovative way to find existing trends in our Quality data. After running through the usual statistical suspects – think regression analysis and business intelligence dashboards – we were encouraged to think bigger. The system should be able to group similar, existing records together regardless of who entered them, where they were entered, and the language in which they were entered. Just as importantly, the system must be able to evaluate a new record and predict the type of trend it might fall into, as well as offer a potential path towards investigation and resolution. And of course, the application also needed to be compliant with the regulatory rules for applications assisting with low-risk processes in the pharmaceutical manufacturing space (i.e., GxP applications).

THE TRENDING TOOL

The Trending Tool (TT) is backwards-facing and looks over the sea of existing information, and clusters them into groups by the type of record they represent. These collections of records are then the trends.

There are four main processes that underpin the Trending Tool:

1. The records are queried from the system.
2. The queried records are translated into English.
3. The records are vectorised and clustered using HDBSCAN.
4. The records are then presented for visualisation.

We focus on the unstructured text in the records, such as the short and detailed descriptions. The goal is to find every major trend hiding in the data, which we do by clustering the similar records together. The process begins by transforming the unstructured text into structured data through the creation of embeddings. With embeddings, we can identify significant words and ideas that carry the most weight and relevance within the given context. First, we translate every record into a common language (i.e., English), then comes the crucial step of clustering. Clustering algorithms analyse the embeddings and group similar pieces of text into clusters. In the final stage, we label the clusters/trends by automatically assigning some meaningful phrase to each cluster, which briefly describes how the records in the cluster are related. These labels attempt to provide a summary of the content within each cluster, making it easier for users to navigate and understand large volumes of unstructured data. With the records now assigned to their trend/cluster, they can be further segmented, and their occurrence can be graphed over time, revealing important information about our efforts to manage those trends. Our (human) employees then evaluate each trend, looking for new, emerging trends, or if known trends are being adequately mitigated. For example, if we find more deviations, month over month, then that strongly suggests that our mitigation efforts need to be reviewed. If, in contrast, we find that the number of deviations drops month per month, then that strongly supports the conclusion that our mitigation efforts are succeeding. The TT therefore provides both a broad, yet focused view into our quality processes, and does so regardless of the records' input language or location.

The Trending Tool's sister application is the Recurrence Check. The TT looks back over time, while the Recurrence Check Tool (RCT) looks towards the future. The purpose of the RCT is to take a brand new (and unclustered) record, read the unstructured text fields, find existing records similar to this new record, and predict to which trend this new record might belong. The RCT is an AI-enabled extension of the previous, manual recurrence check process, in which we perform a query-based recurrence check, whereby individuals manually query the GOTrack database looking for existing records whose field data match the new deviation record. The RCT still performs this task but does so at the push of a button, eliminating the need for any manual querying.

In addition to the automated database query, the RCT employs two new semantic search operations, looking first at individual record similarity, as well as at cluster similarity. The first operation, individual similarity, semantically compares the current record's text data to every other record's text data, independent of language or origination location, and returns the list of existing records that are most closely semantically related to the new record. The second operation again uses semantic similarity to find the record that is most semantically similar, and then determines if that record belongs to a cluster. If so, the RCT uses additional information to compare new and existing clustered records for potential similarity. The application of individual – and cluster – similarity

provides an AI-enabled 'four eyes' principle to finding semantically similar records, while using these three methods together gives us an automated recurrence check that stays faithful to the original search, but also extends the search beyond keyword/field matching and includes semantically related records. Records whose fields/values match through querying are presented to the user as a SQL-based query match. In contrast, those records whose fields/values do not match the current record, but whose text fields are nonetheless semantically similar in meaning, independent of language, will have a similarity score between 0.0 and 1.0. The similarity score is the calculated cosine similarity between the new and existing records' vector embedding. A similarity score of 1.0 means that the records are completely similar. A score of 0 means the records are completely dissimilar, while a similarity score of 0.5 indicates that the records are partially related. Finally, those existing records that are part of a cluster, and are therefore associated with an existing quality trend, are presented to the system user with the trend/cluster number next to the similarity score. Users can then use this information to complete the regulatorily-required recurrence check.

Our long-term vision with the TT and RCT is to create an autonomously intelligent framework for trend and recurrence management, whereby the system alone can find trends using a myriad of information and intelligence perspectives, and then manage those trends without human interaction. For many obvious reasons, we did not pursue that vision in the initial release of the applications. Instead, releasing the TT and RCT productively meant introducing them with the requirement that a human review the information provided by the tools and take ultimate responsibility for the quality and accuracy of the recurrence check. This approach, using a framework of assistive intelligence, will give us the necessary data to verify that the intelligence within the applications is operating as expected, while guaranteeing the continued stability and oversight required of a regulatorily-required process.

The AI team consisted of five internal employees (Kriti, Pascal, Tobias, Dr Jonas and me) and one external consultant, all of whom had various development experience, and all of whom had deep AI or data science experience. This team worked with the project team members to build the necessary data models, and then integrate them into our Quality application platform. In the end, we spent more than 4,000 hours building approximately 30 different models to deliver the clustering and similarity results that most closely aligned with the opinions of the subject matter experts. Ultimately, we found that HDBSCAN with a TF-IDF vectorizer and special preprocessing returned the best results for the TT. We were able to demonstrate that TF-IDF even outperformed most of the newer, transformer-based models. Note that one BERT-based model was shown to match TF-IDF's vectorization performance but was shown to be computationally more expensive. On the other hand, we found that a fine-tuned version of the MS MARCO embedding model provided the best vectorization output for calculating semantic similarity with the RCT. A serendipitous side effect of this 'four-eye' application is that we cast a slightly wider net, achieving better results together than would be provided by using just one embedding model for both the TT and RCT.

We employ two different embedding models, TF-IDF and MS MARCO. MS MARCO is a neural network-based, semantic embedding model that produces dense vectors while TF-IDF is a syntactically based algorithm that simply uses word frequencies to calculate sparse embedding vectors. The inexorable march of technology would entice one to believe that MS MARCO should be superior to TF-IDF, especially considering the MS MARCO model was fine-tuned on record text to further learn semantic associations within the records' information. Nonetheless, TF-IDF proved to be a worthy weapon in our toolbox. We attribute its performance partially to the preprocessing we performed on the record text (e.g., stemming and lemmatization) tailored to the type of text in this domain. While we need to do more testing to verify or refute our hypothesis, we believe

that TF-IDF has an advantage over open-source, trained embedding models when the record text to embed contains proprietary data such as product names, which would not likely appear in an open-source embedding model's pretraining corpus. Because TF-IDF does not 'care' about the 'meanings' of such words, but rather focuses on its frequencies, it's arguably better able to handle these kinds of proprietary words.

As mentioned, the TT and RCT were designed to be used as part of a regulated process, which meant that the tools needed to be validated. While there is guidance from the various regulatory agencies on how to build, use and maintain AI applications in a regulated space, the prevailing regulations still lean heavily on 'classic validation' techniques, which prove the function's reliability through repeated demonstration that entering a specific input yields a known, expected output. That paradigm has worked wonders over the past decades, but it doesn't naturally apply to situations in which the output is not known, or where valid variations among the 'correct' answers could exist. For example, imagine you gave a list of 20 popular, temporally diverse books to a group of friends, and asked them to classify them into two groups: Classics and non-Classics. Unless your list consists of timeless, universally loved masterpieces along with absolute dreck, it is incredibly unlikely that your collection of friends will generate the exact same groupings. And even more unlikely if you ask them to group the stories into lists by 'type.' Let's assume there is at least some agreement. Are some friends right, and the rest unread rubes? Perhaps. I don't know your friends. But it's more likely that the subjective nature of the task leaves room for varying degrees of correctness.

More data about the things to be clustered may not necessarily lead to universal cluster agreement, either. Every year as the weather turns cold, American college football fans turn their attention to the annual celebration of creating two groups of college football teams: those who have earned the right to play for the national championship, and those who get to watch it at home. Despite having every datapoint that could ever conceivably lead to the objective determination of the best four teams in the country, there are as many clusters as there are people opining. In fact, even the experts, whose entire professional purpose revolves around this singular purpose, rarely agree.

These contentions demonstrate the inherent difficulty in applying classic validation to an AI-enabled application, whose purpose very well may be to deliver the answers we couldn't calculate or anticipate on our own. Indeed, even relatively proscriptive AI such as HDBSCAN and TF-IDF depend on the complete set of datapoints to calculate the output. As the ebbs and flows of data push semantic associations one way one month and the other way the next, one college football team may be worthy of the playoffs today but may be pushed out tomorrow when another team's resume shines slightly brighter in the light of another day. This shift may not come through any of the team's actions, but rather result from the gravitational forces exerted through the actions of the other teams. For this reason, we employed a validation approach that recognised the inherent differences between a normal, deterministic operation and an intelligent one, staying faithful to both the letter and spirit of the validation process, while demonstrating the operation's correctness even despite not knowing exactly what the output would be. We did this primarily by demonstrating longitudinal stability through human review. Specifically, we created 12 test sets, slicing the data into 2-year windows that progressively slid over the dataset, greeting the new datapoints while forgetting the old. Each slice of data was reviewed to determine how much change occurred between and within the groupings, and if the groups were correct. We established an accepted threshold for the variation in size in the main clusters and showed that the composition of these clusters was always within the threshold, thereby demonstrating stability. More importantly, a human expert judged each slice's composition, and certified that the clusters were consistent with the expected quality of human clustering.

Having successfully released our company's first GxP AI application into production, we have turned our sights on the future, and intend to deliver three important updates. First and foremost, we are working with human experts to understand how the clustering and semantic search operations can be improved. This may include offering multiple clustering operations (e.g., one for big sites and one for smaller sites) or performing the clustering operation with multiple models and algorithms and using a mix-of-experts approach to return the optimal results.

Creating and releasing Boehringer's first fully validated, AI-enabled application in the GxP space was a challenging and rewarding experience.

Similarly, we are considering our opportunities for continuously fine-tuning our MS MARCO embedding model with new and extended data. We are particularly excited about the possibility to further fine-tune the MS MARCO model with human feedback using triplet loss, which is a loss function that takes positively and negatively correlated examples and seeks to push embedding results towards the positive examples and away from negative examples. Fortunately, we already have the data to perform this operation without human intervention; we know the records that the human experts indicated were related (the ones they chose) and the records that the human experts felt were not related (the ones they were recommended but did not choose). With this information we can automatically incorporate triplet loss to improve our model's performance.

Finally, we are investigating the possibility of using tools and techniques like density-based clustering validation (DBCV) and other validational metrics to further demonstrate stability over time. We believe that through consistency and both system-based and human-based oversight, we can automate the process of updating and validating our models, providing increasing performance over time.

Creating and releasing Boehringer's first fully validated, AI-enabled application in the GxP space was a challenging and rewarding experience. Our team is delighted that we could deliver such a valuable, cutting-edge application, which has already delivered numerous insights and time savings for the company, and more importantly, has helped us better supply our patients – both humans and animals – with the medicines they need, when they need them. Looking ahead, we are excited about the significant advancements we can make and the similar applications we can develop, using the TT and RCT as models. Pun intended.

CAN LARGE DATA REVEAL HOW SPONSOR COMPETITION AFFECTS PARTICIPANT RECRUITMENT IN CLINICAL TRIALS?

By PHILIPP DIESINGER, GABRIELL FRITSCHE-MÁTÉ, ANDREAS THOMIK AND STEFAN STEFANOV

PHILIPP DIESINGER
is a data and AI enthusiast, having built a career dedicated to serving clients in the life science sector across diverse roles. Philipp is driven by his passion for leveraging data-driven decision-making to deliver tangible results with real-world impact.

GABRIELL FRITSCHE-MÁTÉ
works as a data and technology expert in the pharma industry, solving problems across the pharma value chain. He has a PhD in Theoretical Physics and a background in physics, computer science and mathematics.

STEFAN STEFANOV
is a senior software engineer who brings 7 years of experience in developing software solutions in the domain of life science. His passion lies in perfecting UI/UX design and transforming intricate data into user-friendly, insightful visualisations.

ANDREAS THOMIK
is a senior data scientist driven by a passion for leveraging data and AI to generate business value, and has done so across multiple industries for close to a decade, with a particular focus on the life sciences field.

EFFICIENT CLINICAL TRIALS: MANAGING COSTS AND RECRUITMENT CHALLENGES

Clinical trials are integral to the development of new medical treatments, but they come with significant financial burdens. The costs associated with clinical trials often reach tens or hundreds of millions of dollars [1] .

Several factors contribute to these high expenses. Clinical trials require meticulous planning and design to ensure they meet regulatory standards and scientific rigour. This involves extensive preclinical research, protocol development, and obtaining necessary approvals. Ensuring compliance with regulatory requirements is essential and trials must adhere to strict guidelines set by regulatory entities like the FDA or EMA, necessitating detailed documentation, monitoring, and reporting. Non-compliance can result in costly delays or trial termination. In addition, conducting a trial involves significant operational expenses, including site management, patient monitoring, data collection, and analysis. These operations require specialised personnel, sophisticated technology, and robust infrastructure. Trials that necessitate hospitalisation or frequent clinic visits further increase the financial burden.

Given the high costs, it is crucial to minimise risks during the design and setup of clinical trials. Accurate and sufficiently conclusive data is essential for the approval of new treatments and for advancing medical knowledge [2,3]

Effective risk minimisation can lead to more efficient trials, faster approvals, and reduced costs. Efficient trials maximise the return on investment and ensure that resources are used effectively.

One of the major risks in clinical trial operations is participant recruitment [4]. Participant recruitment involves finding and enrolling individuals who meet the trial's eligibility criteria and are willing to participate in the study. Recruiting participants is critical for the success of a trial, as it directly affects the validity and reliability of the results. It is crucial to achieve adequate sample sizes, which are necessary to ensure statistical power and reliable results. Insufficient recruitment can lead to inconclusive or invalid findings.

In addition, trials often need diverse participants to ensure that findings are generalisable to the broader population. Recruiting participants from various backgrounds and demographics helps achieve this diversity. Time-efficient recruitment is essential to adhere to trial timelines. Delays in recruitment can prolong trial duration, increasing costs and delaying the availability of new treatments.

Participant recruitment risks include factors such as insufficient enrolment, participant dropout and ethical concerns:

1. Failing to enrol enough participants can invalidate a trial, wasting time and significant resources. This can happen due to overly stringent eligibility criteria, lack of awareness, participant reluctance or other factors.
2. Retaining participants throughout the trial can be challenging. High dropout rates can lead to incomplete data and affect the study's outcome.
3. Ensuring informed consent and voluntary participation is crucial.

The risk of recruitment can be exacerbated when multiple sponsors directly compete at the same clinical sites or hospitals for participants [5]. Multiple trials vying for the same participants can deplete the available – and often limited – participant pool, making recruitment even more challenging. In addition, overburdening participants with multiple trial options can lead to confusion, reduced participation rates, and increased dropout rates. Clinical sites may become overwhelmed by managing multiple trials simultaneously, leading to operational inefficiencies and increased errors.

AN ATTEMPT AT QUANTIFYING THE RELATIONSHIP BETWEEN RECRUITMENT RISK AND SPONSOR COMPETITION AT CLINICAL SITES

In an attempt to investigate whether recruitment risk measurably changes when multiple sponsors compete for participants at the same sites, we used a data set of almost one million clinical studies across various countries, clinical sites and therapeutic areas.

We focus our analysis on clinical studies of breast neoplasms, a well-known medical condition with a significant amount of reported data. We did so by filtering the original data on the condition 'breast neoplasm' and its synonyms defined by UMLS [6].

We excluded studies that started before the year 2000. Studies that started later have been reported with more rigorous data standards through online study registries. The website clinicaltrials.gov is the first registry that went online on February 29, 2000 [7].

In the end, 9336 studies remain in the filtered data set.

For each of the selected breast neoplasm studies, we attempt to quantify the number of competing studies.

We consider two studies to be 'competing' if they are set up to recruit participants for the same medical condition from at least one identical clinical site at the same time, i.e., they run in at least one common site and they 'overlap' in time at least for one day.

By this definition, we find that, on average, around 80% of the sites a given study recruits from, also run competing studies.

We denote $n_h^d(s)$ as the number of studies that compete with a given study s at site h on day d. We calculate the average number of competing studies $\bar{n}_s$ as the average of the number of competing studies $n_h^d(s)$ across sites and the duration of study s:

$$\bar{n}_s = \frac{1}{\Delta t_s} \sum_{d=date_{start}}^{date_{completion}} \frac{1}{|H_s|} \sum_{h \in H_s} n_h^d(s),$$

H_s denotes the set of all clinical sites study s was conducted at and Δt_s is the duration of study s (counted in days).

For example, let us assume that study s runs in two sites (*'site 1', 'site 2'*), starting on January 1, 2021 and is concluded on December 31th, 2021. Study s is investigating breast neoplasms.

While in *site 1* there are no competing studies, at *site 2* there are three competing studies z_1, z_2 and z_3. All three studies z_1, z_2 and z_3 investigate breast neoplasms as well. All three studies z_1, z_2 and z_3 start January 1, 2020 and run for over five years. Then,

$$\Delta t_s = 365, |H_s| = 2, n_{site\ 1}^d(s) = 0 \text{ and } n_{site\ 2}^d(s) = 3$$

for all days d during which study s is active. Thus,

$$\bar{n}_s = \frac{1}{365} \sum_{d=01.01.2021}^{31.12.2021} \frac{1}{2}(0 + 3) = \frac{365*1.5}{365} = 1.5.$$

To simplify our analysis further, we group studies based on the average number of competing studies $\bar{n}_s$ into five categories:

(i) $n_s = 0,$

(ii) $\bar{n}_s \in (0, 1],$

(iii) $\bar{n}_s \in (1, 3],$

(iv) $\bar{n}_s \in (3, 7],$

We differentiate between studies based on the type of enrolment they report, i,e., 'estimated enrolment' or 'actual enrolment': Ongoing studies typically report planned ('estimated') number of participants. Concluded studies. on the other hand, typically report the number of participants that have actually enrolled in the study.

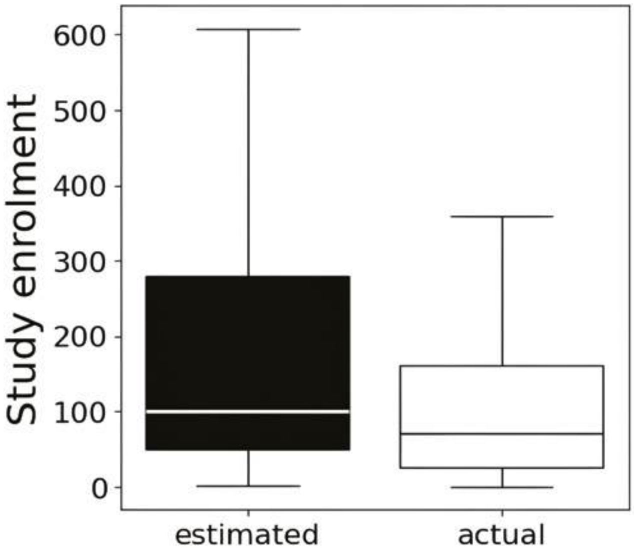
In an attempt to quantify the impact of sponsor competition on participant recruitment we define a new KPI. We name it *recruitment discrepancy* D, defined as the relative discrepancy between the median of estimated (planned) enrolment $med(E_e)$ and the median of actual enrolment $med(E_a)$

$$D := 1 - med(E_a)/med(E_e).$$

To perform our data analysis, first, we calculate the *recruitment discrepancy* D of all studies in set *(i)*, i.e. studies without any competing studies. We do so to establish a baseline.

We find that even in the case of no competing studies, there is a *recruitment discrepancy* of 30%, potentially suggesting that studies often enrol 30% less patients than they plan to. The distributions of estimated and actual enrolments, respectively, show significant overlap as demonstrated in the box-plot (Figure 1). This indicates that the effect we are attempting to measure is deeply statistical in nature.

FIGURE 1:
Comparison of actual and estimated enrolment across breast neoplasm studies without directly competing studies.



Next, we calculate *recruitment discrepancy* D for all remaining sets *(ii) – (v)* to find out how D might change with an increasing number of competing studies. We find that the results are again deeply statistical with significant overlap between distributions (Figure 2). However, we find that *recruitment discrepancy* D increases with an increasing number of competing studies.

FIGURE 2:
Box-plot comparison of median of estimated (planned) enrolment and median of actual enrolment for varying numbers of competing studies. The numbe of competing studies is indicated by set notation at the x-axis.

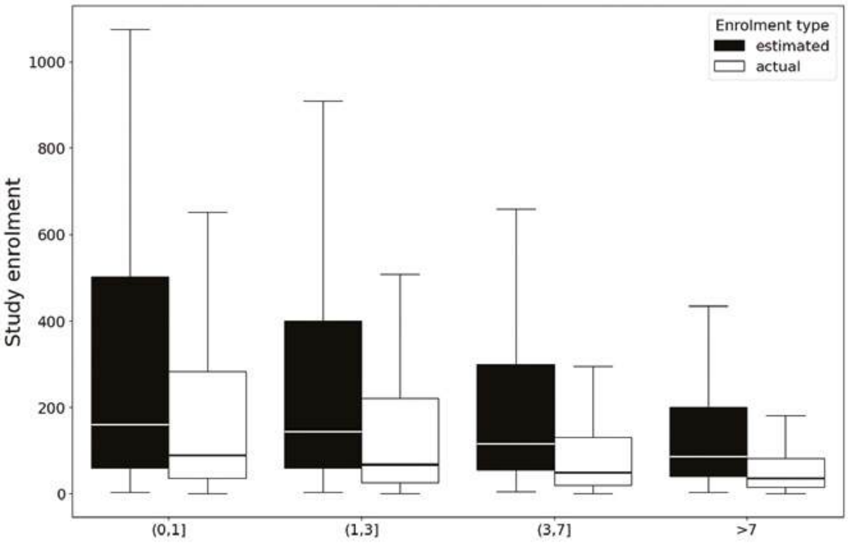
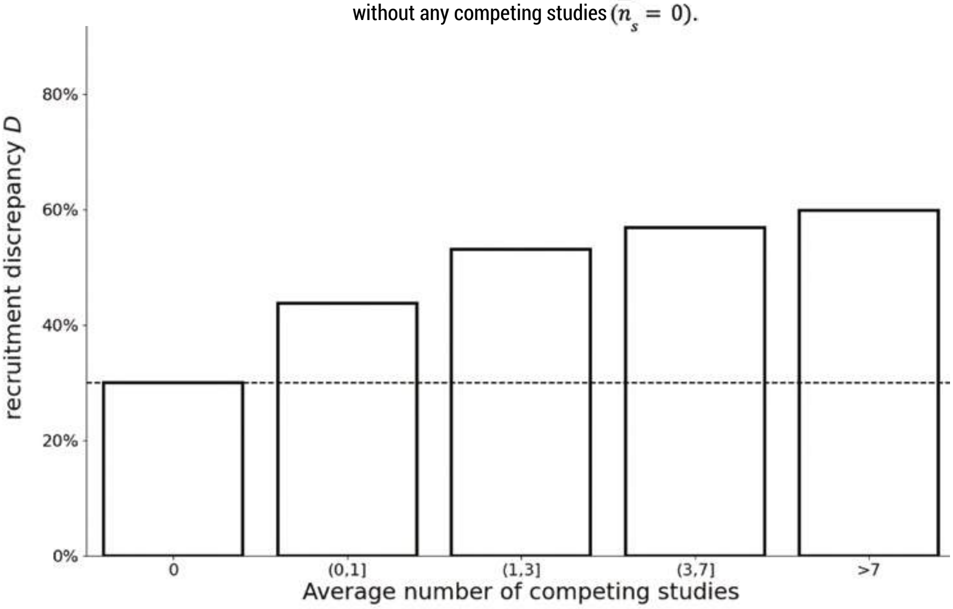
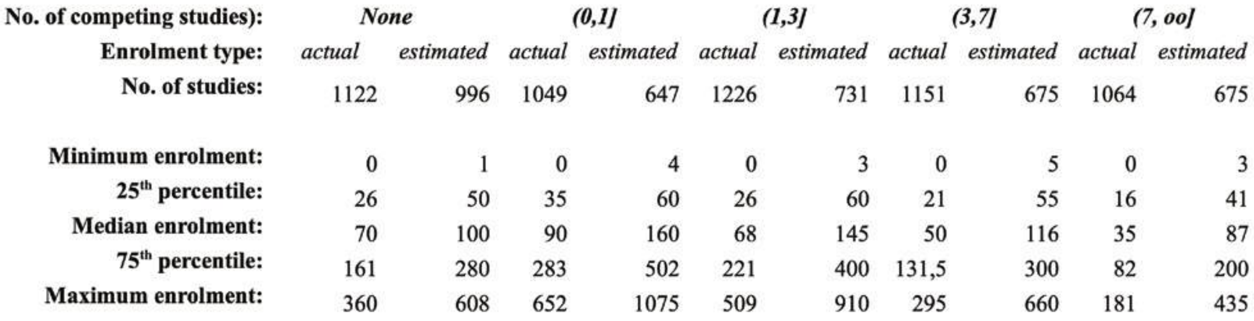


FIGURE 3: *Recruitment discrepancy* D as defined above for varying numbers of competing studies (as indicated by set notation at x-axis). The horizontal dashed bar indicates the recruitment discrepancy baseline for studies without any competing studies $(\bar{n}_s = 0)$.



The effects we find are highly statistical in nature, motivating the need for large amounts of data sets to achieve robust insights. We limited the data set to a narrow subset of one medical condition and an advantageous time period, starting from the year 2000. However, the anecdotal analysis presented above might suggest that a statistical relationship between the number of competing clinical trials and elevated risk of participant recruitment exists.



No. of competing studies):	*None*		*(0,1]*		*(1,3]*		*(3,7]*		*(7, oo]*	
Enrolment type:	*actual*	*estimated*	*actual*	*estimated*	*actual*	*estimated*	*actual*	*estimated*	*actual*	*estimated*
No. of studies:	1122	996	1049	647	1226	731	1151	675	1064	675
Minimum enrolment:	0	1	0	4	0	3	0	5	0	3
25th percentile:	26	50	35	60	26	60	21	55	16	41
Median enrolment:	70	100	90	160	68	145	50	116	35	87
75th percentile:	161	280	283	502	221	400	131,5	300	82	200
Maximum enrolment:	360	608	652	1075	509	910	295	660	181	435

TABLE: Data overview by set of competing studies.

REFERENCES

[1] Joseph A. DiMasi, Henry G. Grabowski, Ronald W. Hansen, Innovation in the pharmaceutical industry: New estimates of R&D costs, Journal of Health Economics, Volume 47, 2016, Pages 20-33, doi.org/10.1016/j.jhealeco.2016.01.012

[2] FDA Approval: What it means, fda.gov/drugs/development-approval-process-drugs#FDA

[3] The evaluation of medicines, step-by-step, ema.europa.eu/en/human-regulatory-overview/marketing-authorisation/evaluation-medicines-step-step

[4] Desai M. Recruitment and retention of participants in clinical studies: Critical issues and challenges. Perspect Clin Res. 2020 Apr-Jun;11(2):51-53. doi: 10.4103/picr.PICR_6_20. Epub 2020 May 6. PMID: 32670827; PMCID: PMC7342339.

[5] Gelinas L, Lynch HF, Bierer BE, Cohen IG. When clinical trials compete: prioritising study recruitment. J Med Ethics. 2017 Dec;43(12):803-809. doi: 10.1136/medethics-2016-103680. Epub 2017 Jan 20. PMID: 28108613; PMCID: PMC5519451.

[6] Unified Medical Language System (UMLS) nlm.nih.gov/research/umls/index.html

[7] "Fact Sheet, ClinicalTrials.gov". U.S. National Library of Medicine. May 3, 2011. Archived from the original on September 27, 2011. Retrieved September 16, 2011.

JOHN has a strong background in software development and has worked for Microsoft, Symantec, and multiple Medtech companies over the years. Recently, his focus has been on technologies that will drive improvements in patient care. He runs active projects in machine vision, augmented reality, data connectivity, security for medical devices, and machine learning.

ROLE OF DATA SCIENCE, MACHINE LEARNING, AND AI IN REDUCING THE CLINICAL BURDEN

BY **JOHN O'GORMAN**, PRINCIPAL TECHNOLOGY OFFICER – CYBERSECURITY & DATA ANALYTICS,
S3 CONNECTED HEALTH

Heralded as one of humanity's most significant steps forward, perhaps there's nothing so frequently discussed as AI and machine learning. Many industries are scrambling to implement regulations and guidelines to handle the mass amounts of data inevitably flooding systems. How can we adapt quickly enough? What exactly are the possibilities? What are the risks involved? These are just some of the questions in the slew of thought leadership pieces that pepper our daily feeds.

Healthcare is no different. Yet, the industry's regulated nature tends to amplify the challenges. Given the sensitive nature of their work, healthcare companies must take adequate precautions when investing in new technologies or using AI and machine learning in their systems.

As most of you will know, healthcare is facing a crisis. A study by the World Economic Forum found that we will face a global shortage of 10 million healthcare providers (HCPs) by 2030, and a survey by the American Medical Association claims that 40% of doctors in the US have considered leaving the field. How will we provide adequate care in the face of such numbers? Many are turning to AI and machine learning, seeing these as tools to help reduce the growing clinical burden.

In doing so, many of the questions we must ask are unprecedented. We can improve healthcare and give patients and clinicians access to the best data and answers. What remains unknown is the best way to go about it.

Creating the best possible path forward is no easy task, and the guidelines are ever-changing, so healthcare professionals must lean on experts and partners who understand the complexities inherent in data management and security and can help address potential issues with rigorous risk assessments.

The entire field must remain vigilant about the evolving role of data science, machine learning, and AI in healthcare and the responsibilities of implementing such innovations in conjunction with pharmaceutical treatments, medical devices, or digital health platforms. AI has tremendous potential when applied safely and with a rigorous approach to risk assessment.

TWO AREAS WHERE AI IS ALREADY DELIVERING ON ITS POTENTIAL IN HEALTHCARE

AI and machine learning's benefits will likely touch every part of the healthcare system in the future. For this article, let's focus on the two areas delivering on their potential to reduce the clinical burden.

AI is particularly useful in clinical decision support (CDS) and decision support interventions (DSI). By mining and making sense of large volumes of data collected by digital health solutions and medical devices, AI efficiently spots patterns to support healthcare professionals.

Algorithms can sort and pattern large amounts of data into accessible insights in minutes, and these tools have been used effectively in several healthcare areas. From managing global pandemics to early cancer detection, AI has proven that it can lessen some of the growing burden on the industry, reducing clinical burden and the burnout of healthcare professionals.

SOME OF THE PROVEN BENEFITS OF CDS AND DSI SOLUTIONS

According to the National Library of Medicine, some of the potential advantages of CDS/DSI systems when implemented well:

- **Improves patient safety** by reducing the incidence of medication/prescription errors and adverse events.
- **Increases patients' adherence** to treatment by providing clinical guidelines, follow-ups, and treatment reminders.
- **Contains costs** by reducing test and order duplication, by automating tedious steps and reducing provider workload.
- **Supports administrative function** by automating the selection of diagnostic codes through streamlined documentation and note autofill.
- **Provides accurate diagnostic suggestions** to clinicians based on patient data and test results.
- **Makes medical images and tests more accessible** by augmenting the extraction, visualisation, and interpretation of such images.
- **Aids patients with decision support** through transparent access to personal health records so they can make informed decisions.
- **Reduces clinical burden** by improving workflows, documentation, and access to accurate information, expediting previously arduous tasks that only exacerbated the clinician's workload.

However, developing effective CDS and DSI solutions in the face of a multitude of regulations and a rapidly changing technological landscape can be challenging.

THE KEY GUIDELINES GOVERNING HEALTHCARE DATA GOVERNANCE AND AI

Below are four sets of guidelines or governance that currently impact the healthcare industry's adoption of AI and machine learning initiatives.

EU AI ACT

This act is a comprehensive regulatory framework for the global governance of AI and is relevant to multiple sectors. Its primary concern is that AI systems are safe, transparent, and respect fundamental rights. It categorises systems into four risk areas: minimal, limited, high, and unacceptable – and the obligations of those developing the AI increase based on the level of risk. The EU AI board oversees it, and penalties for non-compliance are up to 35 million EUR or fines of up to 7% of global turnover.

What this act means for healthcare: Healthcare solutions are likely to be classified as high-risk, given the vulnerable nature of handling so much patient data. Already, systems used for biometric categorisation based on sensitive attributes and those evaluating eligibility for healthcare services fall under high risk.

The compliance requirements include:

- Establishing a comprehensive risk framework.
- Ensuring data quality, integrity, and security.
- Maintaining detailed technical documentation of AI's function and compliance.
- Providing transparent information about the AI system's capabilities and functions.
- Implementing measures to allow for appropriate human oversight and intervention.

The adoption of these measures will present challenges:

- The EU AI Act mandates specific systems requirements (e.g. Risk Management, Cybersecurity, QMS) for high-risk AI systems. The harmonisation of these requirements with the Medical Device Regulation will be needed to allow medical device companies to meet these requirements efficiently for notified bodies.
- Guidance on a clear pathway for clinical and performance evaluation of AI technologies in the clinical setting. It is unknown yet if the notified bodies will specify test techniques for specific types of AI techniques.
- Standardisation of the types of technical documentation to be presented. The cybersecurity guidance has done a lot to standardise what is required for secure medical device submissions (e.g., threat models and trust boundaries). Clear guidance for AIs will be needed in the future.

As the AI Act moves towards becoming law, clarifying these issues for industry will become critical to a smooth transition.

IEC'S GENERAL TECHNOLOGY STANDARDS

Below are some that recently emerged to support this transition to AI and Machine Learning. Two of the standards which are of particular interest to medical device companies include:

ISO/IEC 23894: Guidance on AI Risk Management

This new framework focuses on risk management in AI systems. It will be familiar to medical device companies as it lays out the principles of risk assessment and deals with the specifics of risk identification, analysis, and evaluation.

Following the identification of these risks, a strategy to manage and reduce them is laid out. Within the medical context, these can then be evaluated as AI risks and fed into the safety risk process to identify potential harm to patients.

INTERNATIONAL STANDARD

ISO/IEC 23894

First edition 2023-02

Information technology — Artificial intelligence — Guidance on risk management

Technologies de l'information — Intelligence artificielle — Recommandations relatives au management du risque

Reference number ISO/IEC 23894:2023(E)

© ISO/IEC 2023

INTERNATIONAL STANDARD

ISO/IEC 42001

First edition 2023-12

Information technology — Artificial intelligence — Management system

Technologies de l'information — Intelligence artificielle — Système de management

Reference number ISO/IEC 42001:2023(E)

© ISO/IEC 2023

ISO/IEC 42001:2023 Guidance for Management Systems

This framework outlines the principles of a QMS suited for AI systems. It is similar to the already required ISO/IEC 13485 in the medical device world. It lays out the fundamental pillars of running a quality system:

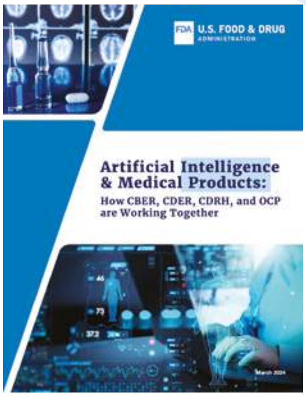
- Operational management within the context of risk management strategy.
- Performance evaluation of the system.
- Identification of opportunities for improvement.
- Support and management of the system.

Tracking this information and meeting these requirements will require a robust QMS that comprehensively ensures risk is adequately managed. It will become a critical component of risk management as we move forward with AI and machine learning in medical devices.

FDA ARTIFICIAL INTELLIGENCE/MACHINE LEARNING GUIDANCE FOR SAMD

The FDA is responsible for ensuring the safety and effectiveness of many medical products that use AI technology, regulating solutions based on their intended use and the level of risk to patients if they are inaccurate. The guidance mainly governs Software as a Medical Device (SaMD); examples include software that helps detect and diagnose a stroke by analysing MRI images or computer-aided detection (CAD) software that processes images to aid in detecting breast cancer.

AI-enabled software, like any other medical device, is subject to FDA review based on its risk classification. Devices are categorised into three classes based on their risk level.



FDA U.S. FOOD & DRUG ADMINISTRATION

Artificial Intelligence & Medical Products:

How CBER, CDER, CDRH, and OCP are Working Together

March 2024

- **Class I devices**, for example, software that only displays readings from a continuous glucose monitor, pose the lowest risk.
- **Class II devices** are considered to be moderate to high risk and may include AI software tools that analyse medical images, such as mammograms, and flag suspicious findings for a radiologist to review. Most Class II devices undergo a 510(k) review, where the manufacturer demonstrates that its device is 'substantially equivalent' to an existing device on the market with the same intended use and technological characteristics.
- **Class III devices** pose the highest risk and include products that are life-supporting or substantially important in preventing impairment of human

health. In the full premarket approval process, developers of these devices must submit clinical evidence that the benefits of the product outweigh the risks.

In March 2024, the FDA's Centers jointly published a paper detailing the Centers' four high-level priorities for a patient-centred, risk-based regulatory approach that strikes a balance between fostering responsible and ethical innovation and upholding quality, safety, and effectiveness.

- **Foster collaboration to safeguard public health:** The FDA is committed to working with various stakeholders to establish consistent standards and guidelines. It plans to seek input from global regulators, developers, patient groups, and academics to shape important regulatory aspects. Additionally, the FDA aims to support stakeholders in their efforts by promoting educational initiatives related to AI in medical products.
- **Advance the development of regulatory approaches and support innovation:** The FDA intends to bring predictability and clarity to the regulation of AI in medical products and work towards issuing comprehensive guidance to support these critical endeavours. This will involve closely monitoring trends, proactively identifying knowledge gaps, and seizing opportunities to enhance and streamline regulatory efforts.
- **Promote the development of harmonised standards, guidelines, best practices, and tools:** Building on their previously issued Good Machine Learning Practice Guiding Principles, the FDA plans to take a number of proactive steps to establish standards, guidelines, and best practices across the medical product life cycle. This includes a comprehensive evaluation of best practices for safety and performance monitoring, encompassing ethics, representativeness, bias, transparency, safety, cybersecurity, quality assurance, and risk mitigation dimensions.
- **Support research related to the evaluation and monitoring of AI performance:** Subject to the availability of resources, the Centers also plan to support demonstration projects. These include projects focusing on managing and mitigating potential bias, addressing health inequities, promoting equity, ensuring representative data, and leveraging diversity, equity, and inclusion efforts. Additionally, there will be projects aimed at ensuring adherence to standards, performance, and reliability throughout the product life cycle.

HEALTH GOV IT HTI-1 RULE

In early 2024, The Department of Health and Human Services and Office of the National Coordinator for Health Information Technology (ONC) published the Health Data, Technology, and Interoperability: Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule. The rule represents a significant step in regulating AI in healthcare IT and ensuring that AI technologies are safe, effective, and provide value for money, all while protecting patient interests.

The framework requires developers and providers to demonstrate transparency and risk management to ensure the safe and effective use of predictive models and algorithms. It also mandates that decision support systems (DSIs), like clinical decision support tools and other AI-driven applications, provide in-workflow access to supporting evidence for impartial decision-making. This leading-edge regulatory approach will promote responsible AI and allow clinical users to access a consistent baseline set of information about the algorithms they use to support their decision-making and to assess such algorithms for fairness, appropriateness, validity, effectiveness, and safety (FAVES).

ONC
Office of the National Coordinator for Health Information Technology

Health Data, Technology, and Interoperability

Certification Program Updates, Algorithm Transparency, and Information Sharing (HTI-1) Final Rule

FAVES is the five key characteristics high-quality predictive algorithms and decision-support interventions ought to have:

Fair: The model must not exhibit prejudice or favouritism toward an individual or group based on their inherent or acquired characteristics.
Appropriate: The model must be well-matched to specific contexts and populations to which it is applied.
Valid: The model has been shown to estimate targeted values accurately and as expected in both internal and external data.
Effective: The model has demonstrated benefit in real-world conditions.
Safe: The model is free from any unacceptable risks and for which the probable benefits outweigh any probable risk.

WHERE DO THESE REGULATIONS POINT US IN THE FUTURE?

When taken as a whole, it's evident that the principles used to regulate healthcare solutions protect against many of the same risks. They remind companies to remain vigilant in a fast-paced environment and indicate that as healthcare technology rapidly evolves, so will the guidelines and regulations that safeguard us. The next decade will see rapid changes, and the parameters will likely change as time progresses. We will learn which restrictions are necessary and those that may be a hindrance to moving forward. What will remain important is managing risk.

Building CDS Tools that Incorporate AI and Machine Learning Safely

When assessing the requirements for future medical devices, developers must be pragmatic about risks and consider the full impact of AI or machine learning on safety. The sensitivity of the data involved means that there will always be a need for rigorous risk assessment and a need to understand any risks to the patient of harm, any risks to patient data, and any risks to the integrity of treatment.

CDS tools must be built with the encryption and safe storage of vulnerable data always at the forefront. As we move forward, it will be critical that patients and clinicians can rely on solutions to protect and manage data while mitigating risk.

Best practices for stakeholders in the health sector

- **Assess AI systems for risk level:** Determine the risk level classification and compliance requirements for existing and planned AI systems.
- **Regular engagement with regulatory bodies:** Stay informed of guidance or changes to regulations issued by relevant regulatory bodies.
- **Plan for compliance obligations and timelines:** Awareness of critical dates and obligations will be vital in achieving regulatory compliance for medical devices or CDS/DSI tools.
- **Develop comprehensive compliance frameworks:** Establish, regularly maintain, and update protocols and guidelines for risk management, data governance, and technical documentation.
- **Conduct third-party conformity assessments:** Planning for third-party conformity assessments will become critical, particularly for any CDS tools using AI systems that might be classified as high-risk.
- **Build clinician panels with regular feedback loops:** This will ensure you meet regulatory obligations for post-market surveillance and support your team in building devices correctly.
- **Follow a quality-based approach:** Using standard quality management principles based on standards ISO 23894 and ISO 42001.

IN CONCLUSION

Regulatory and notified bodies need to know that companies are building safe systems that mitigate harm to their citizens. As shown in this article, both Europe and America have laid out quality-based guidelines for AI and machine learning that can be used to ensure the risk of harm is appropriately managed when building medical devices. Likely, the evolution of AI and machine learning technologies will outpace the regulatory bodies. A pragmatic and cautious approach to risk that upholds the principles around quality, safety, and effectiveness is the only way forward when looking at the future development of CDS and DSI tools.

> *Likely, the evolution of AI and machine learning technologies will outpace the regulatory bodies. A pragmatic and cautious approach to risk that upholds the principles around quality, safety, and effectiveness is the only way forward when looking at the future development of CDS and DSI tools.*

KEY PRINCIPLES FOR DELIVERING AI OUTCOMES IN LARGE ENTERPRISES

WALID MEHANNA is Chief Data & AI Officer the Merck Group headquartered in Germany, where he also chairs the company's Digital Ethics advisory panel. At Merck, he leads and helps deliver data & AI strategy, architecture, governance and engineering, across the entire global organisation.

Born in Egypt and raised in three different states in Germany, Walid has harnessed his multicultural background to inform his commitment to diversity, equity, and inclusion in the workplace. Walid is a senior corporate executive with the soul of a startup entrepreneur who is very focused on integrating data and AI in classic company environments.

You have first-hand experience of working through various tech hype cycles – the internet, distributed computing, blockchain and now GenAI. What's your take on the GenAI hype cycle?

All hype cycles occur because of human nature: we get overly excited, and then we get overly critical. And often the truth is somewhere in the middle. So what tends to happen with every cycle is there is always a period of hype followed by a period of awakening. And the question is: how significant is the awakening; how useful, and how timely?

So blockchain, for example, is a very exciting technology but I believe its time hasn't come yet. Maybe it will be mainstream one day, but maybe it'll remain niche. We'll eventually find out. But when we look at GenAI, there will always be substance that's brought in, and there will be waste in terms of attention, money and investment that went into the technology but didn't lead anywhere.

So the first key challenge for me and my teams is to identify the right inception point. When does a technology become relevant for us, and could it be a competitive differentiator? We need to be ready to leverage it at that point. For example, in the case of quantum computing, there's a lot to observe. We're actively being involved in discussions and we have a solid foundation of understanding what works, what doesn't work, and what's happening in the market where we have investments. So we double-down when it becomes relevant for our operations and business models.

For GenAI, after ChatGPT in November 2022 – that was the inception point for us, because it was becoming mainstream.

It's always a question of finding a good balance between the over-selling and the over-hyping, versus leveraging the potential and the substance that is already there. And for me that's a leadership challenge, not a technology or sales challenge. It's a leadership challenge to separate the hype from the substance and put the substance to good use.

> *If you want to succeed with AI, you need quality data that's specific to your business and your processes.*

Could you talk us through how you utilise data and AI at Merck?

I strongly believe in two things: one is that you can't separate data and AI. If you want to succeed with AI, you need quality data that's specific to your business and your processes. Quality is the key word here. Everyone has data, but if your data is a waste product or a side product, it won't help you with the AI. So, if you're really interested in leveraging AI to gain a competitive advantage, you must take good care of your data.

The second thing is that data and AI can't be an ivory tower capability. Obviously, you need central organisations – global organisations like mine – as a catalyst, a governance entity, strategic entity, cultural change entity. But it can't be a delegation of accountability. The accountability for data and AI needs to be embedded in everything we do as an organisation.

We want to have our understanding and capabilities as deeply embedded as possible in the organisation. But we also want economies of scale, and need to be very disciplined in our usage of technologies that can easily become very expensive.

When you approach a data product, you think in three different dimensions. Can you explain that approach?

Historically, we would have different scenarios: either a new system we're building, an application, or even just a dashboard. People would look into the source systems and identify which data they need, then they source this from different systems. Because the data doesn't directly connect to each other, people would then build mappings and clean the data. Alternatively, there would be a data set somewhere they get access to from a data warehouse, a data lake, or some other application. That's the shortcut.

But then they find out that this data is not complete, and they need additional data. So, they start to build what I like to call Frankenstein data sets, because essentially it's a second-hand data set that's then enriched by firsthand data. Nobody has a lineage of that data.

So, a data product is a more strategic approach for us. Currently we're going top-down with different data domains and what we're trying to do in the future is to push responsibility for the data as close as possible to the point of creation. We haven't implemented data contracts; we have limited data lineage for a lot of systems, but for others we do have good lineage.

So we want to establish that transparency, and most importantly, the right accountability and responsibility across the organisation. We're also aiming to support it with the platforms available to deliver good quality, to monitor the data and pipelines.

Sometimes there's the need for a quick fix. We like to call it 'stop the bleeding.' Obviously, that's a priority, and that's okay. But after you stop the bleeding, there's a tendency to stop even more bleeding, put more patches on it. But that's not a sustainable or healthy solution.

So the second part of the journey – the more strategic part – should be to heal the patient. So, how do we get to these healthy behaviours and healthy setups? Ultimately the solution lies in those who create the data, and who have awareness that this data is an asset to the organisation. That should be part of their daily job, and should ideally be incentivised as such. And they need to be transparent about what happens with their data.

This is why I'm also a big fan of data marketplace approaches, because with these you can see how that data is used. People are generating business value with it, and customers are delighted because my sales reps have that information that I made available to them.

At the moment, we don't have a real economy of data; we just have people trying to make gold out of coal, which is no easy task. But again, it's a journey and it's an organisational transformation at scale. They bring everybody on board.

We always like to talk about what I like to call the holy trinity: people, ways of working and technology. And you have to start with the people.

How about these other two dimensions that go with the data product: the platform and the analysis aspect?

To a certain degree, the data generation is the first mile of our journey. And getting there isn't typically something data folks have in mind. When we talk about data, people always start with the extraction of the data, but they don't pay any attention to the creation of the data.

So we need to introduce the quality mindset at the point where the data is created: where people put something into systems, or where interfaces are built, or where data models are designed, or when a company is acquired, for example.

That's the first mile. And when we talk about the second, third, fifth, tenth mile, that's essentially how data is made available for secondary use in the organisation or shared between systems.

This is where the platforms come in. Because I don't believe in one size fits all, and I don't believe in the Highlander principle when it comes to technology, because that would be too much of a lock-in. I'm a big fan of global standards, I'm a big fan of open source. I'm super excited about Apache Iceberg, to see that become a relevant piece of our architecture and our global standard.

The platforms for me are fundamental parts that must be effective and efficient in scale across the company. And because of that, we decided to take our key platforms and integrate them into what we call the data and AI ecosystem. We even gave it a name: UPTIMIZE.

UPTIMIZE is essentially our brand promise to our organisation. We called it UPTIMIZE because it's well integrated and it's safe, secure and compliant. And the different components work as seamlessly as possible together.

It's a journey: our ecosystem isn't yet perfect, but it is improving significantly. It's growing with every product increment, every three months. And we also have a multitude of analytics and AI products that are built on top of that ecosystem. They can be fully-fledged applications, simple dashboards, API services that we make available to our organisation. That's what our community of people, our practitioners, our sector hubs, and also our individual contributors leverage and build.

You often say: 'AI has both a first-mile and a last-mile problem.' Can you tell us what you mean by that?

So, the first-mile problem is the creation of data. This creation is governed by data specialists. Often, there's this detachment between the data creators and the other employees in the business, who say: 'that's nice, but that's not my reality. I still have to write stuff down and somebody else maybe types it up if he can read my handwriting.' Or: 'I have to put in the same information in four different systems and I don't know where they ever end up.'

That's our first-mile problem. If you start the data journey on the extraction of data from operational systems, you're already at a disadvantage. If you're serious about it as an organisation, you have to start at the creation. And that's the people on the ground and the systems we give them.

Discipline, good systems, and ideally automated data capture – those are the things that can solve our first-mile problem and could also ease a lot of pain further down in the value chain of data and analytics AI.

Currently, only a handful of companies have that strong foundation, and really think about data this way, and expose data so that every team in the organisation has to do it. Consider the API mandate by Jeff Bezos at Amazon. That's one of the very few examples of companies that have these first principles in place and can reap the benefits of data mindset literacy, and consequently have quality data available at scale.

How does a large, complex company tackle that problem?

Step by step. You have to preach, to sell the concept, and you have to collaborate. And at first you have to put some kind of Band-Aid on it. When somebody has to input data into four different systems, you can build a 'plaster' powered by AI to make sure the data is pushed into four different interfaces but is consistent. That's an emergency solution that 'stops the bleeding'.

But if you want to heal the patient, obviously you have to talk about enterprise application architecture and interfaces. You have to start talking about data languages and data models that people agree on. So, if you have three, five, ten thousand applications globally, which large corporations typically have, none of them have ever been built with a target state in mind.

These always have been built based on what the new data model is for this system. We take either what's coming out of the box or alternatively we do a project, and we start to think about how a data model should look.

But these approaches aren't strategic. There isn't a data domain, an ontology in place, and there's no mapping available. Or, you do it as acquisition, and end up with another 500 systems you're inheriting where you don't have a fighting chance to integrate them.

So, ultimately what you want to do is to define your target state, your strategy, your semantics ontologies, your data domains, and then you need to get disciplined to map to them.

But you also need to be prepared to continuously evolve them, because your business might change. You might acquire companies who have a different business or that know more about you, about the single domain, and therefore it's a journey. And the question is: do you know where you want to go, and do you have a plan of how to get there? Or are you just firefighting and putting Band-Aids on everything to stop the bleeding?

How would you describe the last-mile problem?

Well, the last mile is essentially how we translate it into the day-to-day running of the organisation. We've seen so many dashboards and reports that are fancy and look enticing, but then are not used regularly and become outdated.

So, how do we really get to scale? How do we embed this in our routines? And for that, obviously, we need to talk about user experience and user interfaces, but we also need to talk about things like responsiveness, data availability, and processes.

When you have a new fancy analytics in AI product, a new application, you need to consider: how do I need to work differently to utilise this? It boils down to what we mean by digitisation. Does digitisation just mean that we do everything the way we did previously, but with a tool that maybe gives us 5 or 10%? Or do we rethink the way we do business because of the new possibilities that we have?

This is the one thing where, because of our federated setup, I need to rely on my partners on the ground in the sector hubs, but also in the spokes. Because our remit is so broad, we can't be an expert in all fields we operate in like all fields we operate in, like material science and electronics, and also with key opinion leaders, and doctors in neurology, in healthcare. So, that's exactly where our federated approach comes in.

But ultimately, this is a team sport, so the senior leaders and the middle management must understand digitisation as an opportunity that can help us to up our game.

We can be an early follower, we can be a late follower, we can be a laggard, but ultimately things will change and it's up to us to find out how they can change. And we're coming back to the experimentation and how we really make a difference for those we serve.

Can you describe your thoughts on the term 'use case' and why it can be problematic?

The central technology where we document is still called 'the use case portal', but I'm not a big fan of this term. Because when looking at it from a business perspective, the term use case feels to me like an old man with a hammer looking for a nail. I saw an interview yesterday with Tim Cook, the Apple CEO, and he was asked: 'Apple has never used the term 'AI'. Why?' And his response was: 'As a company, we don't focus on the means to an end, we focus on the end. We don't need to tell our customer that it's AI, and that's why we're not actively using the term.'

Similarly, for me, 'use case' is about the application of technology, but I'm not interested in showcasing the technology – that's the sales mindset. I'm interested in the business impact and the value for the people we serve, for our patients; our customers.

I prefer to have the mindset of a solution builder, someone who understands a pain, an opportunity, and builds a solution to cater to it, rather than applying a technology in a use case to showcase that technology. So, we try to establish the product mindset across all categories. We talk about platform solutions, things like AWS, Snowflake and Foundry. We talk about data products to make sure that a data set is high-quality, actively managed and governed.

Then we talk about the last mile, which is the analytics and AI products, which are solutions to ensure that we're delivering the insight and the functionality needed to solve a business problem or realise an opportunity.

Do you have any final thoughts that you'd offer to current data and AI leaders, or aspiring leaders of complex organisations, who are trying to make data and AI work at scale?

My recommendation would always be: don't worry too much about scale. Focus instead on the impact and value – the difference that you can make in the lives of those that we serve. Our patients; our customers. And if you make sure you're delivering on the promise, either directly or indirectly, depending on your responsibility, then you'll be on the right trajectory.

The second thing is, as you're doing this, also try to think about the future and try to minimise the proliferation and diversity of technological depth. Try to limit the technological complexity, process complexity and process depth. Because ultimately that complexity will hinder your attempts to scale at a later point in time.

But I believe scale shouldn't be the only priority. Scale is only useful when it's built on top of the impact and value that you give to the organisation. Otherwise, scale will be unachievable.

16 BOOKS TO TRANSFORM DATA INTO WISDOM

Essential reads for data practitioners to understand the world around us on a deeper level

BY **NICOLE JANEWAY BILLS**

NICOLE JANEWAY BILLS is the Founder & CEO of Data Strategy Professionals. She has four years of experience providing training for data-related exams. She attained recognition from DAMA for a Master Level pass of the CDMP Fundamentals Exam.

You're inherently curious – you're always looking for new ways to build and apply your skills. You'd like to utilise data to make better sense of the world around you. One effective way to deepen these skills is through reading. In this article, we'll provide you with a curated list of the very best resources to provide valuable insights, improve decision-making skills, save you time in your work, and generally turn data into wisdom.

DATA REIMAGINED

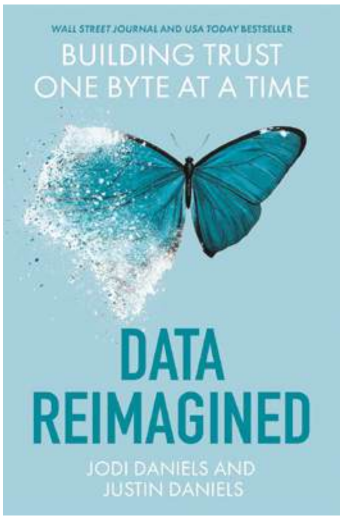
Authors: Jodi and Justin Daniels
Time to read: 6 hrs 44 min (202 pages)
Rating: 4.8/5 (28 total ratings)

SUMMARY: This is a realistic and practical view of how personal data is used in the real, commercial world. There will always be trade-offs. The data will always flow. This book sets the reader thinking about how personal data should be collected, used, stored and secured so that everyone is kept as safe as possible.

Data Reimagined will help you fast-track your approach, create trust in your data collection, and safeguard this trust with I proper data use and sharing practices.

You'll learn security measures for common vulnerabilities and how to use forward thinking as a competitive advantage to attract and build customer trust. Get ahead of the curve with this must-read for all business leaders searching for ways to build customer trust and protect their business.

TL;DR: The authors have done a great job of explaining the complex principles of data privacy and security in an engaging, easy-to-understand manner.


WALL STREET JOURNAL AND USA TODAY BESTSELLER
BUILDING TRUST ONE BYTE AT A TIME
DATA REIMAGINED
JODI DANIELS AND JUSTIN DANIELS

2

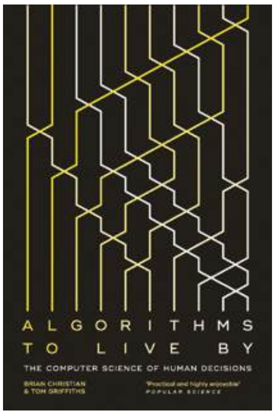
ALGORITHMS TO LIVE BY

Authors: Brian Christian and Tom Griffiths
Time to read: 12 hrs 16 mins (368 pages)
Rating: 4.5/5 (3,460 total ratings)

SUMMARY: This book describes how algorithms can be used to improve the quality of life. Co-written by a computer scientist and a psychologist, it explores themes of order versus spontaneity, finding balance in life, and how technology influences the way we think.

The authors show how algorithms can be used to untangle very human questions. For example: How to have better hunches? When to leave things to chance? How to deal with an overwhelming number of choices? How best to connect with others? Algorithms offer quick fixes that can provide a framework, helping us work through issues that could seem overwhelming at first. The concepts in this book are designed to cut down on decision-making time and free up mental resources for other important tasks.

TL;DR: Describes how algorithms can aid in decision-making, how to approach optimisation, and how to use data to guide our choices.


ALGORITHMS TO LIVE BY
THE COMPUTER SCIENCE OF HUMAN DECISIONS
BRIAN CHRISTIAN & TOM GRIFFITHS

3

THE DRUNKARD’S WALK: HOW RANDOMNESS RULES OUR LIVES

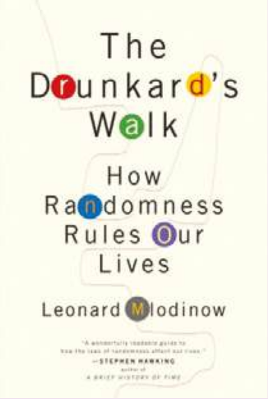
Author: Leonard Mlodinow
Time to read: 8 hrs 10 mins (252 pages)
Rating: 4.5/5 (1267 total ratings)

SUMMARY: The author’s storytelling skills and imaginative approach vividly demonstrate the impact of randomness on our lives. He demonstrates how wine ratings, school grades, political polls, and corporate success are all less reliable than we think. It also highlights how our perception of control is a seductive illusion and that we are surrounded by complexity and chance encounters.

Throughout the book, Mlodinow draws on insights from probability theory and statistical analysis to explain why seemingly random events are often not as random as they appear. He also discusses the ways in which our cognitive biases can lead us to misunderstand the role of chance in our lives. Mlodinow also explores the history of probability theory and how it has evolved over time. He discusses the contributions of early thinkers such as Blaise Pascal and Pierre-Simon Laplace, and how their work laid the groundwork for modern statistical analysis.

In revealing the fragility of control, as well as uncovering the multitude of biases and fallacies which exert their forces on our minds, Mlodinow invites us to think deeper about the decisions we make in our lives.

TL;DR: Not all events and outcomes are well thought through and calculated. Chance and randomness play more critical roles in our lives than we think.



4

EVERYBODY LIES: BIG DATA, NEW DATA, AND WHAT THE INTERNET CAN TELL US ABOUT WHO WE REALLY ARE

Author: Seth Stephens-Davidowitz
Time to read: 11 hrs 24 mins (352 pages)
Rating: 4.4/5 (2,592 total ratings)

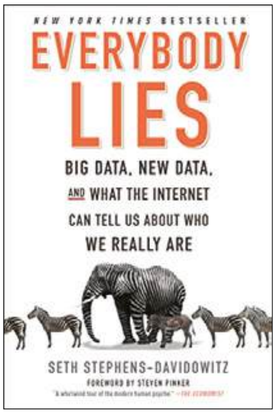
SUMMARY: Stephens-Davidowitz explores new data available to researchers – amalgamations of thousands if not millions of data points collected from social media and online search engines – ‘big data.’ By describing what big data is, how it came to be, and how we may use it to better understand ourselves, Everybody Lies will open your eyes to the unspoken truths of human nature. Stephens-Davidowitz has carefully researched examples from these data sources to support his claims that big data can revolutionise how we look at society.

Information collected from today’s largest social media platforms, as well as search engines, can teach data analysts unspoken truths about human nature which might contain hints as to how we can improve society in unexpected ways.

For data practitioners and internet users, the book’s dive into the lighter and more useful side of data collection is a welcome addition to conversations around data privacy. Your usage of the internet and consent to the collection of your information may one day, however inadvertently, help save someone’s life.

In all, *Everybody Lies* is a thought-provoking book and a good read for anyone interested in the insights that can be gained from big data.

TL;DR: Big data reveals communal secrets that individuals may be unwilling to disclose.


NEW YORK TIMES BESTSELLER
EVERYBODY LIES
BIG DATA, NEW DATA, AND WHAT THE INTERNET CAN TELL US ABOUT WHO WE REALLY ARE
SETH STEPHENS-DAVIDOWITZ
FOREWORD BY STEVEN PINKER

THE CHANGING WORLD ORDER

Author: Ray Dalio
Time to read: 19 hrs 12 mins (576 pages)
Rating: 4.6/5 (5,884 total ratings)

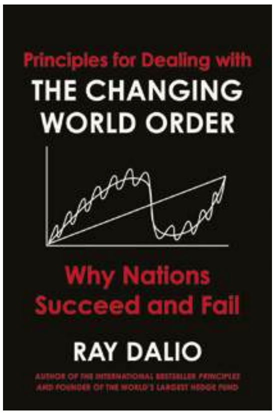
SUMMARY: In this book, legendary investor Ray Dalio examines why nations succeed and fail. He examines history's most turbulent economic and political periods to reveal why the times ahead will likely be radically different from those we've experienced in our lifetimes. He offers practical principles to guide those working with data who seek to 'predict the future' when our past data will not necessarily reflect what is to come.

In his new book, Dalio argues that the times ahead will be very different from anything that we have experienced in our lifetimes. He examines the history of the last 500 years, the rise and fall of major empires, significant economic, political, and social cycles.

5

Writing in a way that is both intuitive and enjoyable to read, Dalio focuses on fundamental ideas and recurring patterns. The book provides a fresh perspective on the current state of the world. It offers insight into how to navigate future changes. Overall, this one is a must-read for anyone thrilled by world affairs and interested in how to best navigate the changes ahead.

TL;DR: This truly fascinating book offers a broad perspective on the changing state of the world. It will teach you how to think in systems, and how to learn from the past to predict what's to come.


Principles for Dealing with
THE CHANGING WORLD ORDER
Why Nations Succeed and Fail
RAY DALIO
AUTHOR OF THE INTERNATIONAL BESTSELLER PRINCIPLES AND FOUNDER OF THE WORLD'S LARGEST HEDGE FUND

DOING GOOD BETTER

Author: William MacAskill
Time to read: 9 hrs 4 mins (272 pages)
Rating: 4.5/5 (806 total ratings)

SUMMARY: MacAskill explores the essential questions that readers should ask themselves in order to make strategic, well-informed choices about how to best invest their time and energy into making the world a better place. MacAskill provides a methodology to determine how we can do the most good in our day-to-day lives. With the critical thinking skills inspired by this book, ordinary people can leverage the incredible opportunity of improving the world we live in.

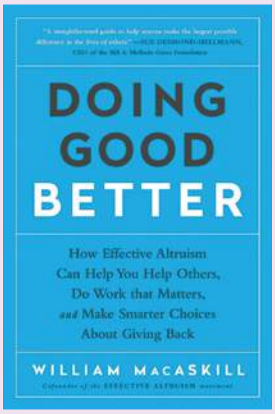
One startling statistic stands out: the most effective charities can be 10x as effective as the average charity, and 100x as effective as an underperforming organisation.

6

Effectiveness is largely about cause selection, and MacAskill argues that we should carefully consider top cause areas such as farmed animal welfare, global poverty, and catastrophic risk reduction opportunities that can have an outsized impact relative to their cost. Interested readers can check out *Give Well* and *Giving What We Can* for some ideas.

Despite the weighty message, MacAskill writes in an easy, conversational style. His excellent book will leave you feeling inspired and empowered to make change.

TL;DR: Offers a data-driven approach to maximising positive impact.


DOING GOOD BETTER
How Effective Altruism Can Help You Help Others, Do Work that Matters, and Make Smarter Choices About Giving Back
WILLIAM MACASKILL

THE PRECIPICE

Author: Toby Ord
Time to read: 16 hrs 4 mins (480 pages)
Rating: 4.4/5 (747 total ratings)

SUMMARY: In beautifully crafted prose, Ord presents a comprehensive and thought-provoking examination of the potential risks that threaten humanity's survival. And yet *The Precipice* is an imminently uplifting work. Ord emphasises that humanity's future is in our hands – we must act now to protect ourselves and generations to come.

We are positioned at a time that Ord compares to adolescence – we face great uncertainty, but we have the tools to ease escalating risks and safeguard our future.

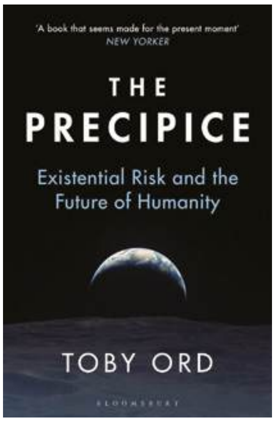
The Precipice explores a range of subjects, offering a thorough study of the potential hazards of each. Ord considers unaligned artificial intelligence, bioterrorism, nuclear war, climate change, supervolcanos, asteroids, and other potential sources of existential risk. He uses careful statistical analysis to conjecture about the harms posed by each.

7

One of Ord's many merits is his ability to communicate difficult scientific and technological concepts in a straightforward way. He offers careful analysis of the moral and ethical ramifications of potential risk mitigation steps.

Ord underlines the significance of assessing existential hazards from a broad and long-term viewpoint throughout the book. He exhorts people to think about immediate repercussions and potential effects on future generations.

TL;DR: Comprehensive analysis of existential risks the world may face in the near future.


'A book that seems made for the present moment' NEW YORKER
THE PRECIPICE
Existential Risk and the Future of Humanity
TOBY ORD
BLOOMSBURY

8

THINK STATS

Author: Allen B. Downey
Time to read: 7 hrs 32 mins (226 pages)
Rating: 4.3/5 (140 total ratings)

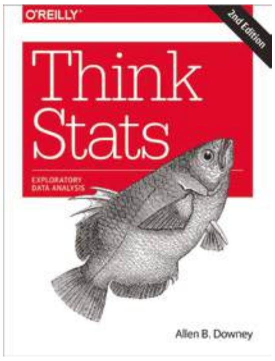
SUMMARY: Learn computational statistics (distributions, probability laws, visualisation, and more) with this book.

Downey takes the reader through various exploratory data analysis techniques such as probability mass functions, cumulative distribution functions, and the like. Additionally, he covers crucial subjects including regression analysis, hypothesis testing, and probability distributions. This book's use of actual data sets to demonstrate the ideas discussed makes it simple to comprehend and put the strategies into practice. The book also has a ton of problems and exercises that support the information presented.

This straightforward reference introduces fundamental statistical ideas alongside Python code. It's an excellent starting place for aspiring data analysts looking to streamline and advance their work.

Think Stats is a favourite of self-learners because the book teaches you statistics in practice rather than through theory, mathematical equations, or proofs.

TL;DR: Alongside simple Python code, the principles of basic statistics are laid out in a readily understandable format.



9

THE ELEPHANT IN THE BRAIN

Authors: Kevin Simler and Robin Hanson
Time to read: 13 hrs 52 mins (416 pages)
Rating: 4.4/5 (790 total ratings)

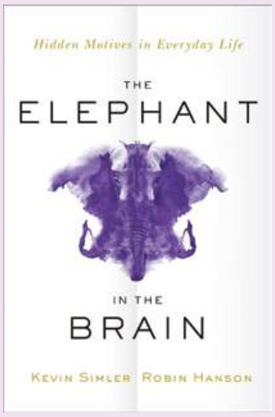
SUMMARY: Explores unconscious motives and how they drive our personal decisions, business interactions, and social institutions. The 'elephant' refers to the selfish motivations we inherited from our evolutionary ancestors.

The main goal of *The Elephant in the Brain* is to teach us to use our conscious thought processes to improve our behaviour rather than trying to justify selfish behaviours.

The acknowledgment that you are not always making rational decisions is an uncomfortable revelation, which simultaneously represents a beneficial opportunity to make more thoughtful decisions in the future.

Simler and Hanson demonstrate that while the motivating factors behind the behaviour of most animals may be opaque to us compared to human behaviour, there are likewise motivations behind human behaviour which remain unacknowledged in most social situations.

TL;DR: Recognising the egoistic 'elephant in the brain' helps a data practitioner become more aware of potential biases in their decision-making. With more clarity, they can make better decisions in conditions of ambiguity.



10

WIZARD ZINES

Author: Julia Evans

SUMMARY: These brief, inexpensive, incredibly useful guides provide clear explanations of challenging concepts in technology. *Wizard Zines* is a collection of short and succinct tech publications by Julia Evans that use comic-style graphics to explain complex concepts in a fun and easy-to-understand way.

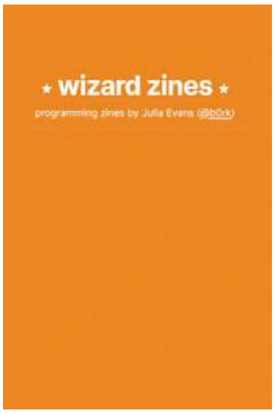
The *Wizard Zines Collection* includes 12 zines, and the *Bite Size Collection*, which includes 4 zines on topics like debugging, command line, Linux, and networking. The zines can be purchased individually for $10 to $12 each or as a whole collection for $117.

The collection is targeted towards young adult and teen creatives but can be useful for those in their 30s and data practitioners who may find Evans' approach (sticking to fundamentals, avoiding jargon, encouraging universal learning) a doctrine that can be incorporated into their workflow.

The collection covers a wide range of technical subjects such as debugging, Linux, networking, SQL, Git, containers, and more. This provides a set of complementary skills thus making them a handy tool for data practitioners.

Overall, *Wizard Zines* is a highly recommended collection. The use of enticing visuals and concise explanations make these zines a great resource for anyone looking to expand their knowledge in a variety of computer science fields.

TL;DR: Learning complex concepts is way easier when we understand the fundamentals on which they are built.



11

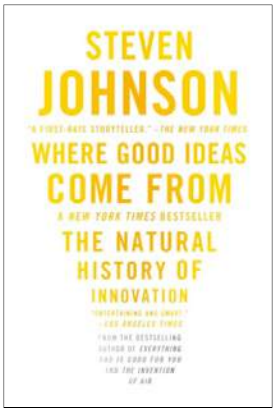
WHERE GOOD IDEAS COME FROM

Author: Steven Johnson
Time to read: 11 hrs 12 mins (336 pages)
Rating: 4.5/5 (1,116 total ratings)

SUMMARY: Johnson examines the origins of innovation in this book. From the 'slow hunch' to the 'eureka moment,' Johnson examines the numerous environments and circumstances that have contributed to some of the most important discoveries in human history. He also looks at how technology and the development of social networks have affected how we come up with and share ideas. His analytical style attempts to identify commonalities across the situations that engendered truly novel concepts, in order to provide data practitioners and general readers with practical advice on how to innovate in their own lives.

Johnson draws parallels between the innovative persistence of life in constant evolution and the need of the modern professional to innovate in order to pursue excellence in their field.

He reveals that making connections in the mind isn't necessarily about absorbing tons of new information. Rather, it is about connecting what we already know in new ways. The brain is by far the most dense, complex network on earth. In order to make the most of the information we have, we have to be able to access it. In a way that is slightly metaphysical, Johnson proposes that we already hold all the answers – they are internalised.


STEVEN
JOHNSON
"A FIRST-RATE STORYTELLER." – THE NEW YORK TIMES
WHERE GOOD IDEAS
COME FROM
A NEW YORK TIMES BESTSELLER
THE NATURAL
HISTORY OF
INNOVATION

This is a must-read for anyone looking to better understand the patterns behind genuine innovation. The diversity of your personal network is what will make your ideas more sophisticated.

TL;DR: Anyone thrilled by the history of innovation, the creative process, or who wants to learn more about the patterns that underlie true invention should read it.

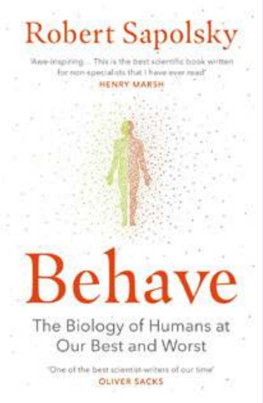
12

BEHAVE

Author: Robert Sapolsky
Time to read: 26 hrs 50 mins (800 pages)
Rating: 4.7/5 (6,492 total ratings)

SUMMARY: *Behave* is a masterpiece survey of 'the biology of humans at our best and worst,' from the endocrine system to cultural influences on our neurology. From subjects like empathy to war, language, and free will, Sapolsky builds on his experience as a university professor with his humourful and enlightening tone while incorporating experimental data and statistical examples into each subject. Sapolsky argues that understanding the complex interplay between biology and environment is essential for understanding and addressing the most pressing social and political issues of our time, from poverty and inequality to war and conflict.

Sapolsky's bestseller represents a combination of multidisciplinary material on human behaviour woven through with rigorous data analysis, hand-in-hand with anecdotal storytelling, which is sure to delight any reader. His work can help us better understand not only ourselves, but the whole human world around us.



TL;DR: This bestseller combines scientific research from every discipline with data-driven analysis to arrive at a complete picture of – not good, not bad, but fundamentally human – behaviour.

13

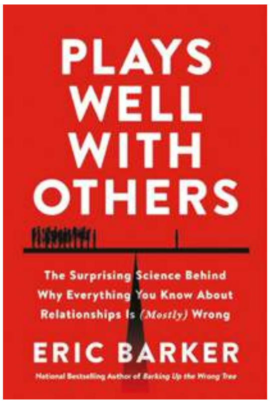
PLAYS WELL WITH OTHERS

Author: Eric Barker
Time to read: 10 hrs 8 mins (304 pages)
Rating: 4.7/5 (665 total ratings)

SUMMARY: In this self-help book that examines the science behind interpersonal interactions, Barker dives into age-old maxims, drawing on science to reveal the truth beyond the conventional wisdom about human relationships.

The book provides a thorough explanation of how people behave, particularly in forming close relationships. Barker emphasises the importance of community in determining a person's level of happiness. Additionally, it gives practical information for recognising and resolving issues in romantic partnerships as well as for enhancing interpersonal communication. No wonder the book has been described as a 'cure-all for our increasing emotional distance and loneliness.' It provides effective advice on how to improve relationships, rekindle love, and approach new people whether you're an extrovert or an introvert, socially confident or anxious.

The author uses humour and anecdotes to explain the similarities between hostage negotiation strategies and marital disputes, how a skilled con artist lied his way into a twenty-year professional soccer career, and why people with opposing viewpoints may end up being our best friends.


PLAYS
WELL
WITH
OTHERS
The Surprising Science Behind
Why Everything You Know About
Relationships Is (Mostly) Wrong
ERIC BARKER
National Bestselling Author of Barking Up the Wrong Tree

For people who work in data management, this book gives important insights on how to be the ideal colleague, how to work effectively with others, and how to resolve problems scientifically and empathically.

TL;DR: Through concise, timely, and enjoyable stories, this book vividly illustrates difficult theories and concepts, and it concludes in a memorable, impactful way.

14

ACCELERATE

Authors: Nicole Forsgren PhD, Jez Humble, and Gene Kim
Time to read: 9 hrs 36 mins (288 pages)
Rating: 4.5/5 (2295 total ratings)

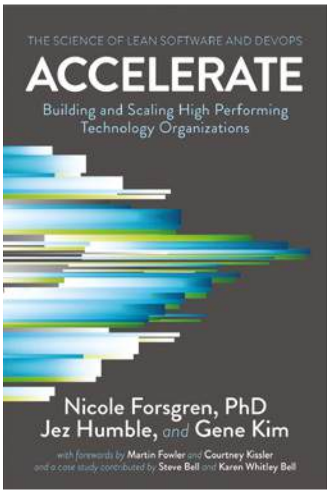
SUMMARY: Forsgren, Humble, and Kim present years of groundbreaking research conducted across more than 30,000 organisations on measuring software delivery performance and what drives it, presenting a collaborative approach to software development and operations that emphasises continuous improvement and delivery.

Uniting two thought leaders in DevOps with a world-class researcher has led to a real path to advancement in organisational design, software development culture, and systems architecture. Their research identified a set of key metrics and practices that are correlated with high-performing technology organisations, including frequent deployments, low change failure rates, and fast mean time to recovery. Using rigorous statistical methods, this book presents not only their findings, but the science behind the research, making the information accessible to readers from all backgrounds.

Similarly, the authors offer actionable practices to improve company culture, leading to high-performing teams and improved outcomes. The clarity and easy functionality of their findings make this book a friend of management at every level.

Overall, *Accelerate* is a comprehensive and practical guide to implementing a DevOps approach in technology organisations. It offers a data-driven and evidence-based perspective on how to achieve high performance in software development and operations and provides a roadmap for organisations seeking to improve their software delivery capabilities.

TL;DR: The authors present their findings on how to measure and improve the performance of software development teams, making concrete suggestions for management at every level.


THE SCIENCE OF LEAN SOFTWARE AND DEVOPS
ACCELERATE
Building and Scaling High Performing Technology Organizations
Nicole Forsgren, PhD
Jez Humble, and Gene Kim
with forewords by Martin Fowler and Courtney Kissler
and a case study contributed by Steve Bell and Karen Whitley Bell

15

THE CHECKLIST MANIFESTO

Author: Atul Gawande
Time to read: 7 hrs 28 mins (224 pages)
Rating: 4.6/5 (7,327 total ratings)

SUMMARY: Gawande presents a convincing argument in favour of the use of checklists. Any data practitioner who is battling complexity will find motivation to bring this structure into life and work.

Gawande is an accomplished medical practitioner – he practises general and endocrine surgery at Brigham and Women's Hospital in Boston, MA. He also finds time to write for *The New Yorker* and has written four best-selling books. So I would imagine he's learned a thing or two about time management and not letting things slip through the cracks.

The Checklist Manifesto offers a thorough study across a variety of disparate fields in an effort to discover how to achieve success in the face of overwhelming complexity. Particularly if you find yourself moving from task to task with little time for reflection, you will benefit from the insights in this highly practical book.

For example, you can start your checklist journey by looking at lessons learned from past projects. Working from this list of potential failure modes, pick one and create a checklist of no more than 5–9 items that aim to ensure your next project will avoid known errors. Move on to your next failure mode and repeat. Be sure to keep the set of checklists close at hand and update them frequently with new learnings and improved processes.

While a checklist isn't a magical formula, it can be a helpful communication tool to your future self, or others, about how to develop good routines. Checklists can be a great way of identifying and mitigating risk upfront.

TL;DR: Checklists bring standardisation and rigour to repeated processes. This book walks through why to use checklists, what makes for a good one, and how to use them.

DATA SCIENCE IN CONTEXT

Authors: Alfred Z. Spector, Peter Norvig, Chris Wiggins, and Jeannette M. Wing
Time to read: 11 hrs 8 mins (334 pages)
Rating: 5/5 (3 total ratings)

SUMMARY: Data science underlies applications used by billions of people every day. This book addresses this important field and raises fundamental questions about data quality, fairness, and privacy.

The book offers frameworks for critically evaluating the ethical considerations needed to apply data science productively and conscientiously, which makes it a must-read for data practitioners. It also explores the fundamental ideas of data science and its effects on contemporary society, including the development of new tools, forms of entertainment, economic development, and potential answers to challenging, complex issues.

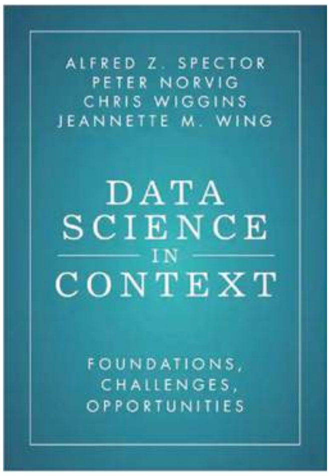
16

All teams working with data and data scientists will need to read this book. To apply data science more effectively and morally, students and professionals should thoroughly understand this perspective.

The book is a collaboration between Alfred Spector, Peter Norvig, Chris Wiggins, and Jeannette M. Wing, all experts in their fields, and their combined knowledge and experience has resulted in well-written, informative and insightful resources for data stakeholders and practitioners.

Read it for free at datascienceincontext.com, and check out author Peter Norvig's class *Practical Deep Learning for Coders*.

TL;DR: This book offers a highly comprehensive, compelling, and readable overview of data science with a range of real-life examples and a keen awareness of the ethical challenges involved.



CONCLUSION

If you're excited by the prospect of learning different ways to think about data, you should consider signing up for Data Strategy News. Each month we do a deep dive into one recent story related to data management and provide one productivity or health recommendation that's especially relevant to data practitioners.

UNDERSTANDING AND MITIGATING
AI SECURITY THREATS IN THE ENTERPRISE

STEVE ORRIN is the Federal Chief Technology Officer and Senior PE for the Intel Corporation. In his role at Intel, Steve orchestrates and executes customer engagements in the federal space, overseeing the development of solution architectures to address challenges in government enterprise, national security, and other federal areas of focus.

Steve has successfully positioned Intel as a leading expert in the application of technology in government. He has orchestrated projects for federal government customers on security, AI and inferencing, and Edge ISR sensing. He's also played a lead role in the funding and launch of the Intel Trusted Compute Pools architecture, a first-of-its-kind integrated security solution for trustworthy virtualisation and cloud stacks.

In his 19 years at Intel, Steve has developed a reputation for successfully building strong teams and playing a key role in leading all facets of technology and strategy.

What are the main types of AI security threats businesses should be aware of?

This was discussed at last year's DEFCON conference in Las Vegas, where there was an entire track dedicated to AI hacking. It included demos of the kind of things you could do to both large language models and generative AI as well as to some of the more classic CNNs and DNNs: the object recognition and natural language processing kind of AI tools, we saw poisoning and prompt injection as a common attack vector. There are some more esoteric or well-designed attacks going after the model weights: first understanding the model to be able to then create images or inputs that will fool the model.

One of the classic examples – I believe it was done by Purdue University – showed that if the attacker understands how your model is set up, and what are the weights it's applied to, and the metrics it's using, then if I had glasses with, say, stripes on them, I could fool the AI and it would no longer recognise me, even to the point where sometimes it would no longer recognise me as a human. So not only would it miss the facial recognition, but you could actually trick it into not knowing what a human was.

Another example from the early days was being able to put a sticker onto a stop sign and suddenly it couldn't detect it was a stop sign.

With a deeper attack, a lot goes into understanding the model. It involves stealing the model, analysing it, looking at the data that's generated or what its weighting bias is, and then being able to use that against the model later. And we're seeing a lot more of that research, where the goal is to bypass or skew an AI for a variety of reasons.

There's another area to think about when we talk about malicious use. Obviously, the AIs themselves are being targeted, whether it's to embarrass or to do malicious activity. But we're also seeing the adversaries, the criminals, the cyber gangs, use AI to conduct their malicious scams. We're seeing much better-crafted phishing that's being generated by AI. We're seeing deepfakes being used to prompt people to send money, or to respond to a support call that's not really from support. And we're seeing AI being used to do information gathering to understand what all the services are and what are all the ports to really speed the process of malware development.

And so, as the adversaries are using these AI tools, we as defenders have to do a better job of adopting and defending against these kinds of attacks ourselves.

What role does hardware play in securing AI systems?

At the end of the day, all of this runs on hardware. There are a couple of key properties that hardware provides. One is that hardware can accelerate the things you want to do to protect AI. So, being able to encrypt your models and data feeds using hardware acceleration to accelerate their cryptography; the key management and the protocols allow you to turn on all those security bells and whistles without impacting performance.

One of the baselines is leveraging the hardware acceleration that's been built in, sometimes as long as 20 or even 30 years ago. Crypto acceleration has been in your commercial off-the-shelf hardware platforms, your Xeons and your PC clients, and it's been available since 2010. A lot of those features are already baked in and much of the software stack takes advantage of it; you just need to turn it on.

But the other area is understanding that hardware is physical, it's not virtual. There are technologies like confidential computing where I can use the hardware to lock down access control to memory and be able to use the hardware-based encryption of memory.

So, if you think about the data security model we've talked about for years: data at rest, security and data in transport. Confidential computing is that last mile of data and in use protection. And for AI, that is really where all the fun happens – in the actual inferencing, in the training and execution of the AI algorithms, and being able to put that into an encrypted memory container where the memory itself is encrypted, and the access to that memory is locked down by the CPU. That's a capability that allows you to protect your AI even if you have malware resident on the platform.

Actually, it even protects you, if someone physically walks up and tries to put a probe onto the platform and tries to read the memory in real-time, it's all encrypted. And so, one of the key roles that hardware will play in securing AI is providing that safe place to stand; what we call a trusted execution environment so that your execution of your AI engine is protected, its algorithms are protected, the models, the weights, all of that can be protected no matter whether it's in the cloud or deployed at the edge where you don't have guards with guns protecting it.

You can use that hardware to protect your AI in its execution, and ultimately you also want to protect the responses. You want to ensure no one's trying to change the response. So, it can give you that end-to-end protection that we've all been looking for.

What are some of the challenges to be aware of when integrating AI security features directly into hardware?

The same challenges you'll always find when it comes to integrating security: one is you have the tradeoff of security versus performance. That's always a classic problem when adding security. One of the things that Intel has spent 30 years working on is how to introduce security features that don't impact performance. And part of that is building those features into the hardware itself.

But the biggest challenge is twofold: first is preparing the software stack; making it easy to integrate into the products and applications you already use. A key thing we've learned (and the industry has learned) is that the more stuff you have to do in order to take advantage of security, the less likely it is that the end user or customer is going to be able to want to do it, or be able to successfully do it. So, removing friction is a key thing.

By performing that integration early, by having the operating system, the security vendors themselves have that hardware integration before it gets deployed. So, when you buy a software product, it's already taking advantage of it. Or when you go onto the cloud, it's just a click of a button and you get confidential computing.

> *One challenge that all security professionals have to deal with is that we must be right 100% of the time. The hacker has to be right just once.*

It's those kinds of integrations that Intel and its ecosystem are doing both in open source and in commercial software to ease that friction between adopting security and deploying it.

The second issue is that often there's a lack of understanding or apathy. A lot of people just don't realise they have a lot of security at their fingertips if they just turned it on. In the security industry, often the reason we still have security problems is you have the right security capabilities, but you haven't configured them, you haven't flipped the switch to turn them on, you haven't deployed it to all the different parts.

One challenge that all security professionals have to deal with is that we must be right 100% of the time. The hacker has to be right just once.

What do you think will be the biggest issues in AI security over the next 5 to 10 years?

I think as AI becomes more pervasive, some of the bigger challenges are going to be around data privacy. AI is a data engine. It's hungry for data, it consumes data, it lives on data. And the more data you put in, the more opportunity there is for exposure of that data. And once data gets learned, it's very hard to unlearn that data. So, I think data privacy and security is a key part of that, but it's bigger than just a question of whether I can protect the data throughout it going in and coming out.

And even in a geopolitical environment, different countries have different determinations of what's considered to be privacy, what's considered to be sufficient security. Different industries from their regulations are going to have different standards. So one of the key challenges of securing and trusting AI is trying to rationalise across all the different domains of trust and regulation to provide a solution that can service all those different environments.

The other thing to keep in mind is that AI is a tool. It's a tool for organisations, industry, and governments to use for the betterment of their customers, their business, their citizens. It's also a very powerful tool for the adversaries. I think we're going to continually see this cat and mouse as they adopt technology much quicker than the legitimate industries do. So how do we keep that balance and how to basically not have implicit trust? This is why the term 'zero trust' is so important today. It's changing that risk dynamic from trust and then verify, to don't trust, verify twice and then maybe trust.

'I'm not going to trust you until proven that you can be trusted.' I think that shift is one of the big challenges of our time. How do I get that dynamic trust built into the use of AI?

The adversaries always take advantage of the fact that these are just implicitly trusted things, whether it be identities, or credentials, or users that are trusted. Zero trust flips that on its head.

It's about doing the right things at the time of the transaction to figure out whether I can trust just this transaction. That shift in the model may help us get a better handle. As AI's constantly changing and evolving, I think zero trust will play an even more important role. We'll consider how to leverage that AI and how do we give it access to the things we want, but ultimately how do we trust it? In some cases we'll decide not to.

These days everyone goes to Chat GPT, and if they get something really weird they know that's inaccurate. But there's a lot of in-between: stuff that comes out of these AI systems that's actually not true, but it looks good enough that we'll think, oh, that must be true. So we need to say: 'I'm not going to trust you until proven that you can be trusted.' And I think that shift is one of the big challenges of our time. How do I get that dynamic trust built into the use of AI?

What should companies do now to prepare for future AI security problems?

There's a key aspect to consider, and it's crucial to consider the aspect itself, and also the point in time that you implement it. That aspect is data governance.

Data governance is critical, but what often happens is that you start building your AI, you're doing your dataset, and only further down the road do you realise you need to add some data governance. That's a huge mistake. Data governance must happen at the very beginning, and at the same time you should start to do your problem definition, because that problem definition will actually inform the data governance and vice versa. It will form how you craft the discovery phase.

What data governance does – besides giving you a framework for applying controls – is it can inform you if you're in a regulated industry or getting data from a regulated industry, or your marketing people think someday you may want to go into a regulated industry. Having that data governance framework built into your model will allow you to apply controls early, so that you can be more agile downstream. I use the word framework, because it's not just a tool that you use once; it's an overlay on the whole process that constantly needs to be informed and integrated in.

Another key point is that it's not just about technology, it's about the organisational dynamics. As you're building these AI solutions, I often talk about having a diverse team building it, not just the data scientists and the developers, but the business unit owners who are actually going to generate the revenue or the benefit.

Having legal and compliance involved from the get-go is a critical step to making sure you're both informed on the potential challenges and planning for them, but also what's coming out the other end, so that when you come out with your AI, they're not like a deer in headlights, panicking and saying: 'Oh God, we're going to shut you down. We haven't looked at this from a compliance perspective.' Having the key stakeholders involved from the beginning is actually a recipe for success time and time again.

What are some common misconceptions about AI security?

One is that you can use AI to solve any security problem. A lot of people think: 'AI is this powerful tool. I'm going to use it to detect the next advanced persistent threat that no one has ever seen before.' But because AI is built on data, it needs to train on a lot of data about how things happen, to be able to make a prediction about how things will happen.

So if there's only ever one of these attacks, that's not enough data to really train a good AI to detect the next one of them.

And that's been one of the fundamental challenges around using AI in cybersecurity. Everyone's looking to use it to catch that one-time, really well-crafted, nation-state advanced threat. And the reality is that's not a good use of AI. That's one of the main misconceptions.

A second common misconception is the opposite of the first: 'I can't use AI for security because I can't trust it,' or

'I've got smart cyber hunters. They're going to do that.'

But there is a place where AI is absolutely going to show real value. Think about a day in the life of a cybersecurity professional inside an organisation; 90% of their time is spent firefighting the hack du jour, a blip on the firewall, the ransomware phishing campaign that's coming in. It's the mundane, everyday issues that happen all the time. They're constantly figuring out which applications are affected, and then patching those applications. It's like a whack-a-mole kind of approach to security, and they get no time to deal with the really important next-generation attack because they're always consumed by the day-to-day.

That's an area where AI can actually shine: in the automation and the application against these mundane, repetitive daily processes. We have a lot of data of what it means to do searching vulnerability databases. That's something that can absolutely be automated – and very effectively, using AI machine learning.

And if you can use machine learning for 80% of what I call the stupid stuff, then your underfunded, overworked and sleep-deprived team of security professionals can focus on the 20% of really hard problems.

That leads us to the third myth: that AI is going to replace my job, whether it be in security or any other field. The reality is AI is a tool and it's something we should harness, and it's not going to replace your job, it's going to augment you. It will enable you to focus on the more challenging and interesting problems – the ones you really want to get up and go tackle. Because AI can take care of the 80% of the mundane, the repetitive, the manual processes.

That separation is actually one of the ways I think that organisations will get the largest ROI of the application of AI for cybersecurity – it's not in trying to find that one-off attack, but in automating and getting more efficient at dealing with the regular, day-to-day kind of vulnerabilities.

[...] every industry, even meatpacking, is dependent on technology for operating and so is vulnerable to ransomware.

What industries are most at risk from AI security threats?

I think we've seen that almost any industry that's adopting AI can be at risk of the threats. Certainly, the regulated industries; the ones that have value. Consider the question that was asked of a bank robber back in the 1800s: 'Why do you rob banks?' And his answer was, 'That's where the money is.' The same is true today.

So industries with money or assets, or are critical infrastructure regulated are going to be targeted. But we've also seen that the motives behind the attacker can vary. It can be financial gain, or it can be sowing chaos and disruption. It could be influence, it could be revenge, it could be geopolitical. The motives are across the board. And when you have this diversity of motives, it means that the target space is much richer.

If it's financial gain, they're going to go after financial assets or the large user bases. But if the motives are otherwise, critical infrastructure, supply chains, logistics – those can also be ripe targets.

And so, it actually doesn't serve us to assume one industry is more at risk than another. It's really about understanding that every industry could be a target, and it's about your organisational risk appetite and risk posture to determine how much security you need to deploy to meet your risk bar.

I believe it's an absolute mistake to think: 'Well, I'm not important. I'm not doing something critical, so I'm not at risk.' A really good example of this came during Covid with one of the big ransomware attacks.

We all know about the attack on the capital pipeline on the US East Coast, but one of the ones that got a lot of news for a short time, but was really informative, was the JBS attack; the one on the meat processor in Australia. They got a ransomware attack, it shut down meat production, and consequently there was a global shortage of meat supply.

And I learned two important things about that. Number one: no one is immune from attack. You couldn't get a less sexy business than meatpacking. There's no money; it's not critical infrastructure, it's not energy. It's the least likely industry you'd expect to be a target.

The second thing was just how dependent we are on technology. Because again, you would think of meatpacking as not being technology-dependent, yet ransomware took down the line. It shows that every industry, even meatpacking, is dependent on technology for operating and so is vulnerable to ransomware.

So, going back to your point, I'd argue that every industry is a target for different reasons. The crucial question for an organisation is: 'Where do I fit on that risk profile?' They should consider whether they have assets that are of value from a cybercriminal perspective; whether they're servicing a critical infrastructure or a critical constituency that potentially makes them an activist target; whether they have an important part of the global supply chain. So, understanding where you fit in that risk profile and then applying the right controls to map to your risk.

And that's why things like the cybersecurity framework from NIST and the risk management frameworks are really important. Because it's not a one-size-fits-all solution. It's understanding what the right controls for my risk are, for my environment, at this given time. And what zero trust adds on to that is the idea that it's not a once-and-done action, it's a continuous process of reassessing risk.

EMBRACING THE FUTURE OF DESIGN:

AI'S TRANSFORMATIVE ROLE IN CREATIVE PROCESSES. AI-CENTRIC DESIGN™ METHODOLOGY

DANIEL MARTINEZ is the founder of Designing Tomorrow®, a startup specialising in integrating AI into the design process. With nearly 30 years of experience at the intersection of tech and design, Daniel has developed AI-Centric Design™, a methodology that profoundly integrates AI to innovate problem-solving and creativity. This approach complements traditional methods, expanding the toolkit for designers and innovators.

With senior roles at global firms like Publicis and Atos, he has successfully integrated AI to optimise design workflows and drive business success.

A recognised thought leader, public speaker and mentor, Daniel regularly shares his insights and vision on ResearchGate and Medium. He holds a specialisation in Creativity & AI from Parsons School of Design and is pursuing a degree in Computer Science for Artificial Intelligence from Harvard. Daniel is passionate about shaping a future where design and AI work hand in hand to create meaningful and impactful solutions.

Undoubtedly, many industries are exploring how artificial intelligence (AI) can enhance operations and drive innovation. The design industry is a leading example of how AI can transform traditional workflows, enhance creativity, and deliver tangible value. This article explores the profound impact of AI on the creative process, highlighting how it necessitates adopting entirely new ways of working to leverage its potential fully.

THE CHALLENGES OF TRADITIONAL DESIGN

Throughout my career in the design industry, I have always found it amusing and challenging to justify the ROI of design. The results and benefits of design are often intangible, making it difficult to quantify its actual value. This has been a constant puzzle, as the subjective nature of creativity often needs to match with the objective metrics businesses seek. This challenge, however, is not unique to design. Many industries grapple with similar issues, balancing qualitative insights with quantitative demands.

In recent years, my curiosity led me to research how new technologies like AI can improve the performance of our design processes and help quantify their actual value. This journey has been enlightening, revealing how AI can be a powerful ally in enhancing creativity and providing tangible metrics to demonstrate the value of design. The lessons learned here are applicable far beyond design, offering insights into how AI can streamline workflows and foster innovation in various fields.

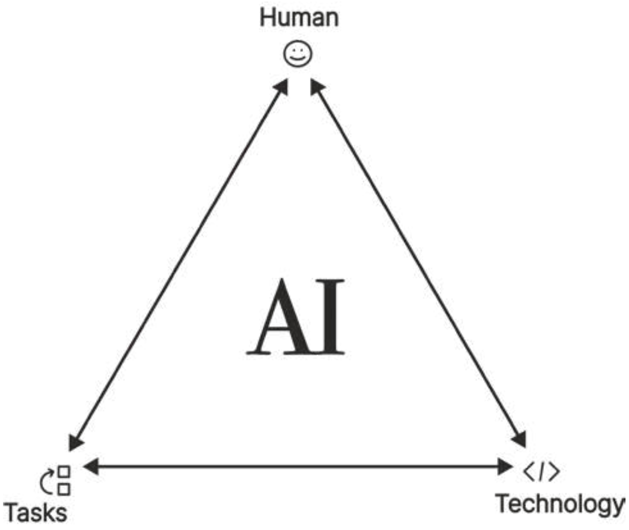
THE TRANSFORMATION BROUGHT BY AI

The advent of AI introduces a transformative shift in how we approach design. AI is not here to replace designers but to augment their creativity and streamline their workflows. By adopting specific AI tools, we can address the inefficiencies of traditional Design Thinking and open up new avenues for innovation. We utilised various AI tools, including OpenAI custom GPTs for natural language processing and ideation, Midjourney for rapid image generation, Uizard and Figma for developing and maintaining design systems, Colormind for generating colour palettes, and Tokens Studio for managing design tokens. These tools helped automate repetitive tasks, synthesise diverse insights, and provide real-time assistance, allowing professionals to focus on more strategic and creative aspects of their work. For instance, AI can help generate ideas, optimise designs based on data, and even anticipate user needs. These capabilities can be applied to marketing, product development, customer service, and more.

Through concepts like anticipatory design, for example, a travel app powered by AI can anticipate a user's preferences based on past behaviour and suggest personalised travel itineraries. We could go further by using GenUI components to create tailored and unique user interfaces for this specific user on demand. This level of personalisation enhances user experience and inspires us to think differently about how we approach our design challenges. It's no longer about pushing pixels on our screens; it's more about strategic design and hyper-personalisation.

THE NEW ERA OF AI+DESIGN

Based on my extensive research and experimentation with streamlining design processes, I have concluded that we are on the brink of a new design methodology, which I have named AI-Centric Design™. I spearheaded this initiative due to the lack of innovation in traditional design processes and the rapid development of AI tools oriented towards creative outputs. While Design Thinking gained formal recognition and application in the 1980s and 1990s through the work of IDEO and its founder, David Kelley, and further traction in the early 2000s with institutions like Stanford University's d.school, it has been around for several decades. This longevity highlights the need for a fresh approach that integrates the advanced capabilities of AI.



AI-Centric Design™ integrates AI into every stage of the design process, leveraging advanced tools and algorithms to augment human creativity, streamline workflows, and optimise outcomes. By examining AI's transformative effects on design, other industries can glean valuable insights into how to integrate AI into their processes.

AI USAGE AT EACH STEP OF THE DESIGN PROCESS

Ideation and Content Generation:
Utilise OpenAI custom GPTs to brainstorm and generate diverse ideas and concepts based on data-driven insights. These custom models help synthesise information from various sources to inspire new design directions. Constructing prompts meticulously is critical to ensure outputs align with brand guidelines and maintain tone of voice consistency. OpenAI's custom GPT models can analyse vast datasets, generate creative content, and provide nuanced suggestions that align with project goals. Additionally, MidJourney is employed for rapid image generation, creating visual elements and concepts that can be iterated quickly. This tool facilitates extensive experimentation with different design styles, enabling designers to visualise and refine ideas swiftly. MidJourney's AI-driven approach ensures that visual content is generated efficiently, supporting a dynamic and flexible design process.

- OpenAI custom GPTs can be tailored to specific projects to brainstorm ideas and generate content based on extensive data analysis. These models can synthesise insights from various sources, offering creative and innovative suggestions.

For instance, they can generate detailed textual content, propose new design concepts, or even simulate user interactions to inspire fresh ideas. By leveraging data-driven insights, these models ensure that the ideation process is grounded in relevant information, enhancing the quality and relevance of generated concepts. Constructing prompts with careful consideration of brand guidelines and tone of voice is essential to maintain consistency and ensure that the generated content aligns with the brand's identity.

- MidJourney specialises in rapid image generation, making it an invaluable tool for swiftly creating visual elements and concepts. It allows designers to experiment with various styles and ideas without requiring extensive manual work. By generating high-quality images based on text prompts, MidJourney accelerates the visualisation process, enabling quick iterations and refinements. This capability supports extensive experimentation, helping designers efficiently explore and validate different design directions. The careful construction of prompts ensures that the visual outputs adhere to brand guidelines and maintain consistency in tone and style.

These tools combine to provide a robust ideation and content generation framework, leveraging AI to enhance creativity and streamline the design process.

Design Systems:
Implement Uizard and Figma to develop and maintain design systems. These platforms integrate seamlessly with AI-generated outputs. OpenAI custom GPTs can generate JSON files for design tokens based on brand guidelines and style kit requirements, ensuring alignment with the overall design strategy. Additionally, leverage Colormind to generate colour palettes that align with the project's aesthetic goals. Use Tokens Studio to manage design tokens, ensuring consistency and efficiency in the design process. Combining these tools allows you to create a cohesive and scalable design system that supports continuous improvement and innovation.

- Uizard is an AI-powered design tool that simplifies the creation of wireframes and prototypes. It can quickly turn sketches into digital designs, making it an excellent choice for developing and maintaining design systems. Figma is a powerful design platform known for its real-time collaboration features and extensive plugin ecosystem, making it ideal for creating and managing comprehensive design systems. These platforms can seamlessly integrate AI-generated outputs, enhancing the efficiency and effectiveness of the design process.

- OpenAI custom GPTs can generate JSON files for design tokens, ensuring all design elements adhere to brand guidelines and style kit requirements. These models can automate the creation of design tokens, reducing manual effort and ensuring consistency. Custom GPTs provide valuable insights that help maintain alignment with the brand's design strategy by analysing and synthesising information from various sources.

- Colormind is an AI-powered colour palette generator that creates aesthetically pleasing colour schemes based on deep learning. It analyses millions of images to generate colour palettes that align with the project's aesthetic goals. By integrating Colormind into your design system workflow, you can ensure that colour choices are visually appealing and consistent with the overall design vision.

- Tokens Studio for Figma (formerly known as Figma Tokens) is a powerful plugin for managing design tokens. It allows you to define, use, and manage design tokens directly within Figma, ensuring consistency across your design system. Tokens Studio supports importing and exporting JSON files, making it easy to integrate with AI-generated outputs from OpenAI custom GPTs. This tool helps maintain efficiency and consistency in the design process by linking tokens to Figma styles and components.

By utilising Uizard, Figma, OpenAI custom GPTs, Colormind, and Tokens Studio, you can create a robust design system that is both scalable and consistent. These tools work together to streamline the design process, reduce manual effort, and ensure all design elements align with the project's aesthetic and functional goals.

Testing and Iteration:
Conduct AI-driven usability tests using machine learning models to analyse user interactions. Integrate AI-generated feedback into the design iterations, using insights from these models to identify areas for improvement and optimise user experience. Tools like Maze and UsabilityHub offer robust features for analysing user behaviour. Maze utilises AI to generate dynamic follow-up questions, identify common themes, and produce automated reports, helping teams make data-driven decisions quickly and efficiently. UsabilityHub provides AI-enhanced usability tests that categorise user responses and offer actionable insights, ensuring continuous feedback and iterative

improvements. Continuous feedback from AI models ensures that designs remain aligned with user needs and project objectives, allowing for iterative improvements and adjustments.

- Maze leverages AI to enhance usability testing by analysing user interactions and generating actionable insights. It can identify common themes in user feedback, provide dynamic follow-up questions, and produce automated reports. This allows teams to quickly pinpoint areas for improvement and optimise the user experience based on real-time data.

- UsabilityHub offers AI-driven usability tests that help categorise user responses and generate actionable insights. It enables designers to better understand user behaviour and make informed decisions to refine and improve their designs continuously.

These tools facilitate efficient, data-driven iterations and ensure user feedback is systematically integrated into the design process, ultimately enhancing the overall user experience.

The Experiment. How AI-Centric Design™ Outpaces Traditional Processes:

In an experiment comparing Design Thinking with AI-Centric Design™, I embarked on two comparable projects regarding complexity and design requirements. The first project involved designing a mobile financial trading application for one of the largest financial services firms in the world. This project included core functionalities such as account opening, login, profile management, transactions, and portfolio management.

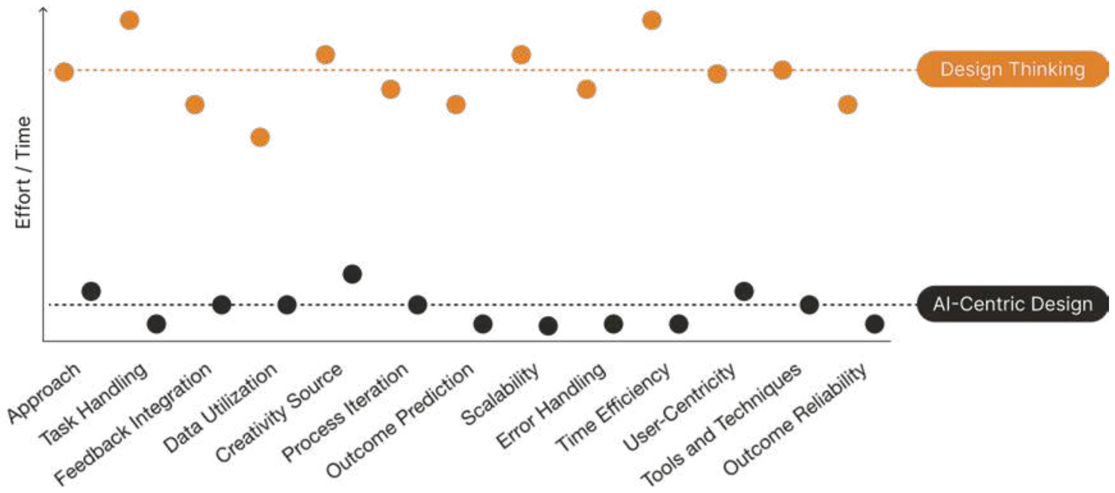
The second project was the design of a DeFi application developed on Web3, enabling users to trade and earn cryptocurrency. This project included similar core features: account opening via MetaMask wallet integration, login, profile management, transactions, and portfolio management. Unlike the first project (which was part of an established financial institution), the second project was set in a start-up environment, requiring us to build a brand-new digital experience from scratch.

Both projects were evaluated for their complexity from a design standpoint, considering aspects such as digital brand experience, features, user flows, and design systems. This comparison ensured that both projects required a comparable level of effort and design innovation, providing a fair basis for assessing the impact of AI-Centric Design™ versus Traditional Design Thinking.

In the first case, following a Design Thinking process, it took six months from kickoff to implementation and launch of an MVP with a team of ten designers across different disciplines (user research, UX, UI, experience technologist, etc.).

In the second case, the team achieved even better results in a shorter time – twelve weeks – with fewer resources (two designers and an army of AI co-workers).

To assess the performance of Traditional Design Thinking vs. AI-Centric Design™, I conducted a heuristic evaluation – measuring the total time taken, human effort required, and quality of outcomes across different areas of the design process – and then performed a comparative analysis to benchmark these metrics, calculating efficiency gains and measuring improvements in design quality.



The AI-Centric Design™ process was by far the winner, being:

- 36 times faster,
- required eight times fewer resources,
- 20% more reliable (five significant errors in project one vs one minor error in project two post-MVP launch).

To understand the practical implications of this shift, let's examine the detailed outcomes of the heuristic evaluation conducted during this project. The heuristic evaluation criteria focused on three main aspects: time taken, human effort required, and the quality of outcomes. For this project, I tailored the heuristics to measure the impact of AI on the design process precisely. Here are the parameters considered:

Results:

Time Efficiency: The AI-Centric Design™ approach significantly reduced the project timeline from six months to twelve weeks. This 36-fold increase in speed was achieved through AI's ability to automate repetitive tasks, quickly generate design iterations, and integrate feedback in real-time. Traditional methods required extensive manual effort and multiple iteration cycles, extending the timeline considerably.

Resource Optimisation: With AI tools assisting in repetitive and data-intensive tasks, the AI-Centric Design™ process needed only two designers compared to the ten required in the traditional approach. This 8-fold reduction in human resources was possible because AI-streamlined tasks like initial ideation, prototyping, and feedback integration allowed the small team to focus on high-level creative and strategic work.

Quality and Reliability: The final designs produced through the AI-Centric approach were 20% more reliable with fewer post-launch issues. Specifically, the heuristic evaluation identified five significant errors in the traditional design process compared to just one minor error in the AI-Centric Design™. This improvement is attributed to AI's ability to process and analyse large amounts of data, providing insights that might be overlooked by human designers and ensuring a more thorough vetting of design elements before implementation.

Innovative Outcomes: The AI-Centric process fostered more significant innovation by freeing designers from tedious tasks, allowing them to focus on creative exploration. AI tools suggested design variations and enhancements, leading to more innovative and user-centric solutions. The AI-driven approach encouraged designers to experiment with new ideas and refine their work based on data-driven insights, resulting in a more dynamic and creative process.

These results underscore AI's transformative potential in the design process. By embracing AI-Centric Design™, we can achieve remarkable gains in efficiency, resource management, and design quality. AI accelerates the design process and enhances designers' creative capabilities, enabling them to produce more innovative and reliable outcomes.

Final Thoughts

Integrating AI into the design process represents a significant evolution in how we approach creativity and problem-solving. AI-Centric Design™ offers a robust framework that enhances efficiency and innovation and aligns design outcomes more closely with user needs and expectations. By systematically evaluating key performance indicators such as task handling, feedback integration, data utilisation, creativity, process iteration, outcome prediction, scalability, error handling, time efficiency, user-centricity, tools and techniques, and outcome reliability, we can gain a comprehensive understanding of the strengths and areas for improvement within AI-Centric Design™ projects.

This detailed evaluation framework empowers design teams to harness AI's full potential, ensuring its implementation is strategic and effective. The insights from this assessment process will guide continuous enhancement, fostering an environment where AI and human creativity work harmoniously to produce exceptional design solutions.

As we move forward, adopting AI-Centric Design™ will transform individual projects and set new standards for the industry. By embracing this methodology, design professionals can stay ahead of the curve, delivering innovative and reliable outcomes that meet the ever-evolving demands of users and stakeholders. This document provides the foundation for achieving these goals, offering a clear path to leveraging AI in ways that maximise its benefits and drive the future of design.

AI KINGDOM IN THE DESERT: MY JOURNEY TO UAE AND SAUDI ARABIA

RUZE LIANG is the AI architect and AI Product Tech Lead for Tonomus, NEOM. He has 10 years of experience in big data, artificial intelligence, machine learning and deep learning, together with advanced knowledge and hands-on practice in large language models. He's an expert in end-to-end solutions and architecture design, successfully delivering several megaprojects and PoC across the Middle East region for public and private sectors. He's garnered approximately 1,500 citations and holds an h-index of 21 following his PhD, and in 2022 he was honoured as an outstanding alumnus by the National Taiwan University of Science and Technology.

As highly internationalised countries, the UAE and KSA offer tax-free packages and a top-notch environment to attract a significant influx of international talent, forming a cross-culture atmosphere and work style.

The artificial intelligence (AI) revolution began in October 2012, when neural networks returned to the academic forefront. Following this pivotal year, advancements in AI rapidly accelerated, garnering increasing attention from the US, China, Europe, and the MENA region, including the United Arab Emirates (UAE) and the Kingdom of Saudi Arabia (KSA).

After obtaining my master's degree in Taiwan, I headed to Saudi Arabia for my PhD. With the world-class supercomputing facility and strong academic resources, I've learnt many machine learning and deep learning knowledge and skills, and have been trained to be a problem solver. With the AI wave starting to surge, my AI journey evolved with the Middle East region.

The decision to work in UAE and KSA was driven by the region's rapid technological advancements and the growing demand for data science expertise. UAE is on the track of being a preferred global AI hub thanks to the concerted efforts of the government in partnership with the private sectors, making the country an indispensable partner of choice for AI developers, startups, and companies from around the world. By 2023, more than 100k workers were employed for AI-related jobs. Meanwhile, KSA was determined to become the centre of AI innovation and the largest investor for the 2030 vision, creating > 40 billion USD investments to foster AI and other key technologies. All these recent actions indicate that UAE and KSA set goals for creating AI-incubating settings, attracting global talent to move to the Middle East.

Immersed in such an attractive and competitive working environment, my roles have been diverse, impactful, and challenging. In my early career as a data scientist, I have dedicated myself to several projects that heavily utilised natural language processing (NLP), computer visions (CV), and audio recognition to solve a variety of problems in different scenarios, ranging from social media, marketing, healthcare, education, and the oil & gas field. In order to achieve a minimum viable product (MVP), I developed a data pipeline and ML pipeline starting from data acquisition, preprocess, storage, data upload, model building, training, validation, and deployment, eventually automating the entire pipelines on the cloud or on-premises to deliver. With my extensive professional experience and expanded network, I started leading various teams, including AI, delivery, and engineering teams, offering direction and guidance to each. Concurrently, I designed comprehensive solutions and architectures by integrating different products, assessed network latency, calculated computational costs, developed solution matrices, and planned project delivery phases. I believe I'm on the right track to keep sharpening my AI skills, and to absorb these in-domain and out-domain knowledge and experiences.

As highly internationalised countries, the UAE and KSA offer tax-free packages and a top-notch environment to attract a significant influx of international talent, forming a cross-culture atmosphere and work style. Unlike the work culture in Asia, which often involves long hours, the work environment in the Middle East is more aligned with European and American standards. Some companies provide hybrid working styles, allowing employees to decide to work from home or office. The public sectors (together with a few private sectors) further introduce a four-day, work-week program, meaning employees would get three days off – Friday through Sunday. Besides this, the emphasis on continuous learning and opportunities for professional growth is obvious across several firms, encouraging people to participate in various AI events or conferences organised by the government, including Kaggle Days Meetup Dubai, AI Everything Global, LEAP Riyadh, and so on. This demonstrates that the UAE and KSA provide excellent working conditions with professional events, conferences, and meetups offering the opportunity to connect with fellow AI specialists and to stay updated with industry trends, fostering the development of leading AI technology giants.

Living & social life in the UAE and Saudi Arabia has been an enriching and unique experience, composed of a blend of Arabic tradition and modernity. In the UAE, cities like Dubai and Abu Dhabi offer a cosmopolitan lifestyle with world-class amenities, providing residents with a variety of choices to enjoy. During the weekday, I will work out or exercise in the gym, or night-walk to refresh after work. On the weekend, I prefer going shopping with my local Emirati friend in the world's biggest Dubai mall, visiting the Louvre Abu Dhabi for a tour of stunning modern architecture and Arabic history, or enjoying a social drink alone. The same life is also aligned with KSA, except alcoholic beverages aren't allowed due to the law. Riyadh, the capital of KSA, offers a rich cultural life in such a great cosmopolitan city. My first site-tour was to visit Diriyah, guided by a Saudi colleague, enjoying the traditional ancient Saudi scene and participating in traditional Saudi festivals. The blend of historical charm and modern advancements makes Riyadh a fascinating place to live and work, reflecting the kingdom's commitment to progress while preserving its cultural heritage. In brief, it's a memorable life in the Middle East, with its astonishing city culture and the Arabic people's warm hospitality.

Currently the KSA and UAE offer immense growth opportunities in data science and AI, focusing on smart city planning, AI-driven solutions, digital transformation projects, and several Giga-scale construction projects. Each of the projects required several experts and professionals in the domain field to bring up key AI solutions, project management, integration capability, contractor coordination, practical implementation plans, and timely project delivery. Meanwhile, the local government also puts in place various investments and vision funding to support AI and other advanced technology development, encouraging more and more people with ambitions and skills to start their businesses here. So far, there are nearly 2000 AI start-ups established in the Middle East, and various headhunters and talent acquisition teams are eagerly looking for candidates to relocate. All of which suggests that the opportunities for AI talents and for firms in UAE and KSA are tremendously high.

> *I believe: Overcome the challenges, and the rewards will follow.*

CONCLUSION

Reflecting on these years in the Middle East, I've invested my best time and finest years to these countries and contributed myself to their AI development and expansion. Combined with making good friends with several Emiratis and Saudis, the blend of professional opportunities and cultural enrichment makes my journey here full of meaning, and very precious. I feel very proud and honoured to be part of them and to offer my AI expertise and project experiences during their crucial time. We have not yet reached the final destination, and there will be more challenges to come. But I believe: Overcome the challenges, and the rewards will follow.

THE GROWING SIGNIFICANCE OF KNOWLEDGE GRAPHS AND GRAPHRAG IN IMPROVING MODEL PERFORMANCE

KIRK MARPLE is based in Seattle and has spent the last 30 years in software development and data leadership roles. His early career included working at Microsoft and General Motors. Later on he successfully exited from his first startup RadiantGrid, which was acquired by Wohler Technologies.

He's currently the Founder and CEO of Graphlit, which streamlines the development of vertical AI apps with their end-to-end cloud based offering that ingests unstructured data, and leverages retrieval augmented generation to improve accuracy, domain specificity, adaptability, and context understanding – all whilst expediting development.

Can you give us a definition of knowledge graphs, and explain what they are?

Knowledge graphs deal with extracting relationships between bits of knowledge. Classically, we talk about people, places and things. So that could be a company with knowledge on where they're located, what their revenue is, and the number of employees. We would call that the metadata on that entity. And that company may have relationships with, say, Microsoft in Seattle. So there's an edge that you create in the knowledge graph between those entities. The knowledge graph is really just that blown up. It's all those different interactions, and inner relationships between those little bits of knowledge, and the value becomes information retrieval.

It's a great way to represent the knowledge, in a way that you can then retrieve it and walk the graph from one place to the other, to be able to learn more from the knowledge that's embedded in it.

Why do you think data scientists and data engineers working in industry should be paying attention to knowledge graphs, in the current AI era?

I think people are familiar with RAG – that the R in RAG is information retrieval. The data presents a search problem or a filtering problem via metadata, and knowledge graphs give another axis, a resolution to the problem of how to retrieve data to feed into large language models and large multimodal models.

Up until recently, knowledge graphs have been a bit of a sidecar in very specialised parts of the industry. But now they're giving you another view on the data. I think vector search has been a big thing over the last couple of years. It's been around for a while before that, but graphs are another facet of information retrieval that complement that as well.

How does the knowledge graph differ from other commonly used data structures, such as a JSON format?

So a JSON structure is basically a kind of linked list. You can follow the links, or operate it like a DAG workflow. The problem is you might get recursion; you might be coming back around to the same element.

For example, when I worked at Microsoft, there might be another person that worked at Microsoft, and then Microsoft bought their company. And then that person links back to me, and so you could get cycles in that graph. That's an area which would be hard to represent, from a serialisation standpoint, in a JSON structure. But there are ways to work around that. You only walk the graph so far, where you collapse links together.

How do knowledge graphs contribute to data integration and the semantic understanding of such datasets?

Say you have a couple of pages of text, and you're mentioning companies, people and places. You can use a vector embedding to find Adobe, or Amazon or someone in that text. It would use a keyword search or even a vector embedding. In the RAG process, you want to pull back the relevant text, and then provide that to the LLM.

But what if there's data about the entity that's not in the text? That's really where knowledge graphs shine: where it's the current year's revenue, or how many employees they have, or other things you can enrich around the entity, that you can then pull back from anything that you're finding, via the vector betting.

So to me, it's an enrichment step where you could just retrieve on the graph itself, but then it's more of a global set of data. You're unlikely to get anything relevant to a specific question. Maybe the question starts with a text kind of vector search, and then you can do another set that pulls data in from the graph. But there may be use cases where you just want to talk to the graph itself, if you have enough information in there. So basically, pull from the metadata around the nodes, from more than the node itself, which might just be like a word.

What kind of methods or techniques are you using to integrate knowledge graphs with LLMs?

There are two sides to it: one is the data ingestion path – the question of how to get data in, and how to extract entities, and how to store it. That's the first step, pulling in data from Slack, email, documents and podcasts. Then you have to do named entity extraction and named entity recognition on that; building up a knowledge graph from the data. So we had a good data structure to pull from.

And now we've moved into the second issue: the RAG graph concept. Now I can ingest all this data, all this representation in a graph, how can I use it?

That's where we've done some experiments starting with vector search, figuring out those similar text chunks that I found, and those text chunks have entities that we've extracted from them. Then you can essentially use this as a way of expanding your retrieval, and pulling it into more content that also observes those same entities. And that's a way to use these graphs for expanded retrieval.

Another way is using graphs and faceted queries to get data analytics. We've actually got a demo on our website of how to ingest a website and use any art to build a graph, and then essentially get a histogram of all the topics, people, places and companies that were mentioned in the data. So you're summarising the graph into a chart form. It's a really easy way to get a different view on what's inside your data.

How can knowledge graphs help with the interpretability and explainability of LLMs?

One way to look at it is: if you have citations, so you have the LLM, it responds. It gives you back a list of citations of your sources that you used. You can then visualise the citations in a graph form and see the topics in addition to the text that it found – the topics and entities that were essentially cited.

You can then check the commonality between these citations. Are they similar? Are they different? I was actually planning to build a demo app that does just that from the citations: do a graph search, get that data and be able to visualise it. That approach is really useful.

What are the signals that a knowledge graph could be the right tool for a particular problem a data scientist is working on?

The thing that comes to mind is the interrelationships between the data. A typical data set might be row-based and there's not as much interrelationships between the rows, or we're seeing different data sets like that. And this is an instance where there's an implicit grouping or classification, and you can bucket data into different classes of nodes.

We were based on schema.org, and the JSON-LD kind of classification structure, the taxonomy – that's really where there's already standardisation around what is a person, what is the organisation, and defining the metadata.

So I think you can look at it one way, where you're kind of mapping data to an existing graph structure. Or you can take a different approach, where you're mapping it, and inventing your own data structure in your own relationships.

But to me, it's really about that classification metaphor of: 'This is a something, and you can assign that, and then that becomes a relationship between other data in the graph.'

Could you explain the concept of graph retrieval augmented generation (GraphRAG) and its importance?

It's something that's come up over the last year or so. It's actually an area that was really core to how we were thinking about our implementation of RAG and using our knowledge graph. Microsoft released a paper on it several months ago.

The concept is basically about leveraging a knowledge graph as extra context for the generation side, which is the prompt you provide to the LLM. So where typically you're using a vector search to find bits of text that are relevant to the question or to the prompt and GraphRAG, you can augment that information you get back with the relationships of that text to the graph.

So, to use our earlier example, with the extraction of Microsoft and OpenAI and Seattle in the sighted text, you can expand your footprint of information. And this is where you have to guide it a bit. Because you don't want to pull in everything – all the last 10 years of Microsoft revenue or something like that. So you have to guide the graph retrieval, and say: 'I'm asking you a question, it seems financially related, go import all the revenue of the last several years from the entities that we've identified in the query.' And that's where you can then use the graph as a secondary search mechanism and retrieval mechanism. It's still early days. There have been some prototypes out there, some papers, but I think it's still an evolving area.

What are the differences between a traditional RAG system and GraphRAG?

So with a typical RAG, you're starting with the text that you've ingested. Text is extracted from documents, transcripts, from audio files, and you're chopping that up into chunks and that's what's provided to the LLM.

By contrast, with GraphRAG you can actually get data that wasn't in the cited text. So it could be data that was extracted from other documents; maybe it found the revenue of a company in a different document, or it pulled it from an API service somewhere else, to enrich the knowledge graph. So it's really a way that it can see outside the domain of the cited sources and provide more context, to answer the question. That's really how we see GraphRAG.

What are the key components or modules within the GraphRAG framework?

A lot of it is really ingestion. So you have to create the knowledge graph. That's probably where the vast majority of the work is. You have to do named entity extraction, or you have to use LLMs, to do the identification of the entities or the nodes in the knowledge graph, and have a really rich ETL pipeline for creating the knowledge graph. So that's the majority of the work; having a pipeline that can deal with changes in data. You edit a document, you remove a reference to a company – how do you remove that from your knowledge graph? Those kinds of things.

The other side of it is during the retrieval stage of the RAG workflow; having a way that you're not just going to look at the text in the cited text, but you're also going to widen your vision and start walking the graph to pull in extra data.

The thing we've seen, at least right now, is it has to be guided. I haven't seen a way that you can make it dynamic and be like: 'Hey, I'm going to use GraphRAG if I see this scenario, and I'm going to use normal RAG if not.' Because there's extra work that has to be done during retrieval, and you had to create the knowledge graph in the first place.

But the question of whether we make it dynamic, and pull from the graph as needed without being guided – that's something we're looking at.

How does GraphRAG handle scalability or efficiency concerns with large-scale graph data?

I think a lot of that is in the architecture. In ours it's more of an index. So we're not storing the text, or we're not storing a lot of data in the graph database itself. So the walking, the graph queries end up being pretty fast.

And then we're able to pull in the metadata from a faster layer, a storage layer. So I think scalability is key. When you're going to have millions or billions of entities, you need a graph database that can handle that.

And I would say most, if not all, of the existing graph database solutions can handle that scale. So that's typically not a problem. The queries are usually fast, they've been tuned enough. That's why I think a lot of the hard work, at least what we've seen, is on the creation side: just making sure you can actually get the right data, keep it fresh, and that kind of thing.

How does GraphRAG address challenges related to building a graph presentation, or to graph representation learning, like node embedding or graph classification?

What hopefully we'll get to soon is the ability to – once you have the graph created – create embeddings of the relationships. So you have a subgraph of information, say, by page. I've observed these companies, these people, these places, and their relationships on a page of text, and I'll be able to run higher-level algorithms like graph embeddings, and say: 'What is similar to this page of text with its graph relationships?' Because today, you can run a vector embedding on the raw text. I think having a graph embedding that takes both of them, or a multi-vector embedding, could be really interesting. And I think I could see that evolving over the next year or so as this becomes more commonplace.

What are some current limitations or potential drawbacks of GraphRAG?

I would say the biggest is that it does take more work. You could go back and backfill the graph from the text, if you've ingested like 1,000 documents, and you want to create the graph, that may be a little more trouble.

We make you opt-in so you can set up a workflow and say: 'I want to build an entity graph from this data,' there's a little bit of extra cost. You're doing some analysis on that text, or you're running LLMs, you're using up tokens. I would say that's probably the biggest limitation: managing cost and scale. Because if you have a tonne of data, maybe you want to pick and choose the data that you want to create a knowledge graph from.

But there are other solutions. If you're not assuming cloud hosted, like an open API, you can build a lot of this more in your data centre with local models and things like that, and keep the cost down that way.

What are the problems that the development community behind GraphRAG are currently working on?

The biggest thing I've seen is that there's not really any standardisation on what GraphRAG means. I think everybody has their own interpretation of it. There was a Microsoft paper that drew a line in the sand about it, and a lot of what they described we were already doing. We previously hadn't really talked about it in that way, about putting a definition to it.

But I think we'll start to see a bit more coalescing. Some people don't agree with this, but I think RAG has standardised a lot. There are still knobs you can turn, like the usury ranking or not, those kinds of things. But the concept of RAG I think is stabilised now. We're not there yet with GraphRAG, and so I think we'll have to go through a wave of trying things, seeing what works and what doesn't, before that settles into a pattern.

How effective are Knowledge Graphs in reducing hallucinations in LLMs?

The grounding concept of providing sources for the LLM to pull from is really what's at the heart of creating an accurate RAG algorithm. And the ability to have graphs and pull in more context, that feeds into providing that extra accuracy. It's something speculative and prototyped at this point, but what we've seen is that you can wind down your amount of hallucinations by proper grounding, by giving good content sources, even by some prompt engineering, to really have the LLM focus on the content you're giving it, not what it's been trained on.

So I think the GraphRAG really feeds into this idea that you're providing it more context, giving it more data to chew on that isn't in its training set. And that should minimise the amount of hallucinations because you're giving it a wider set of context.

What are the developments or advancements that you see coming with knowledge graphs, or GraphRAG that you're most excited about?

We've seen a swell of people talking about it on Twitter, on Reddit, and more papers coming out. So you can see the momentum around awareness and implementation, though a lot of what's out there is still in the demo phase. I think we've been really focused on the ingestion path and constructing the knowledge graph. We'll be releasing something more formal from a GraphRAG feature very shortly.

But what we're seeing in the market is that as people's awareness grows, we'll start to see more people talking about it and implementing it. But it brings up a whole other question of evals: how do you even evaluate that this is better? Some companies have been doing RAG evals. I don't know how well they'll apply to GraphRAG evals and those kinds of things, that can be tricky. But that could open up a whole new market for products or companies to help with that area. So evals aren't going away, they're only going to become more important.

And getting into the future of where the RAG pattern can be useful, we really see two areas: repurposing your content that you've ingested is one of them. So, using the retrieval part of this, to find interesting information in a large dataset might be good for marketing materials, for technical reports. Being able to give a rough outline of what you want to pull from, or create, like a blog post or report, and use the graph to fill in the details. And that graph could be constructed from other structured data, or unstructured data. It's a really interesting area, where we can use this pattern with a really rich data set to create really high quality content.

Then the other is the agent concept that's starting to be talked about. I think there are some open source projects that are working on this. There's also a lot of past work around actor models and distributed architectures and things like that. That's a pattern that can apply, where the RAG concept is really just a set of functionality that gets called from the agent and can feed the output of the agent into the input of another agent.

But from our perspective, we see them as two different layers, where the RAG is more analogous to a database query, where there's some input or some output, and then you have a programming system on top of it for asynchronous agents, and those kinds of things.

So we'll see how it evolves. I think other people have different perspectives; maybe they're more integrated. But I see it more as a kind of workflow layer: a graph workflow, and the RAG is this functionality that fits in underneath it.

ANTHONY ALCARAZ is the Chief AI Officer and Partner at Fribl. His work at Fribl is at the cutting edge of integrating advanced AI solutions into HR, streamlining the final stages of candidate evaluation, and ensuring optimal matches between job seekers and available positions. Beyond his role at Fribl, Anthony is a consultant for startups, where his expertise in decision science, particularly at the intersection of large language models, natural language processing, knowledge graphs, and graph theory is applied to foster innovation and strategic development.

Anthony's specialisations have positioned him as a leading voice in the construction of retrieval-augmented generation (RAG) and reasoning engines, regarded by many as the state-of-the-art approach in our field. He's an avid writer, sharing daily insights on AI applications in business and decision-making with his 30,000+ followers on Medium.

Anthony recently lectured at Oxford on the integration of artificial intelligence, generative AI, cloud and MLOps into contemporary business practices.

PLANNING AS THE CORE CHALLENGE IN AGENTIC AI:

SOLVING IT WITH REINFORCEMENT LEARNING

Picture a team of AI agents working together seamlessly to tackle a complex business strategy problem – one agent researching market trends, another analysing financial data, and a third crafting recommendations, all coordinating their efforts towards a common goal.

This logic of collaborative artificial intelligence, known as agentic AI, represents the next frontier in automation and problem-solving. As AI systems become more sophisticated, there is growing interest in moving beyond rigid, predefined processes to embrace flexibility, adaptation, and teamwork among AI agents.

Agentic AI holds immense promise for automating intricate, open-ended tasks that have long resisted traditional automation techniques. By breaking down complex problems into specialised roles and leveraging the unique capabilities of individual AI agents, multi-agent systems can orchestrate intelligent automation in ways that were previously unimaginable. Pioneering frameworks like crewAI, LangGraph, and AutoGen are paving the way for this new paradigm, enabling developers to design and deploy crews of AI agents that can autonomously navigate and execute complex workflows.

> *Agentic AI holds immense promise for automating intricate, open-ended tasks that have long resisted traditional automation techniques.*

However, as we venture into this new territory of collaborative AI, we encounter a fundamental challenge that lies at the heart of agentic systems: planning.

How do we enable AI agents to effectively plan their actions, coordinate with each other, and adapt their strategies in dynamic, open-ended environments?

This article argues that planning is the core challenge in agentic AI, and that reinforcement learning (RL) offers a promising solution to this critical problem. In the following sections, we will explore the rise of agentic AI and its key principles, explain why planning poses such a significant challenge for these systems, and examine how reinforcement learning techniques can address these difficulties.

By understanding the interplay between planning and reinforcement learning in agentic AI, we can gain crucial insights into the future of intelligent automation and collaborative artificial intelligence.

The Rise of Agentic AI

Agentic AI represents a paradigm shift in how we conceptualise and implement artificial intelligence systems. At its core, agentic AI envisions autonomous AI agents working together in teams, or 'crews,' to tackle complex, open-ended tasks. This approach moves beyond the limitations of single-model AI systems, leveraging the power of specialisation and collaboration to achieve more sophisticated and flexible problem-solving capabilities.

Several key frameworks have emerged at the forefront of this agentic AI revolution, each offering unique approaches to multi-agent collaboration:

1. **crewAI:** This framework enables developers to design AI teams with specialised roles, equipping them with curated sets of research and analytic tools based on their specific tasks.
2. **LangGraph:** LangGraph takes a more structured approach, using explicit directed graphs to define workflows between agents. This gives developers fine-grained control over agent coordination and task allocation.
3. **AutoGen:** This platform relies on emergent workflows arising from multi-turn conversations between agents, allowing for more dynamic and adaptive collaboration patterns.

While these frameworks differ in their specific implementations, they all share core principles that define the agentic AI approach:

Specialisation and Collaboration: One of the most striking commonalities across these systems is how they leverage multiple specialised agents that work together. Rather than relying on a single monolithic model, agentic AI decomposes tasks into subtasks that are delegated to agents with different roles and skills. This specialisation allows each agent to focus on its area of expertise, while collaboration enables the team to tackle problems that would be challenging for any individual agent.

For example, in a job-seeking scenario, a crew might comprise agents specialising in tech job research, personal profile engineering, resume strategy, and interview preparation. By working together, these specialised agents can guide an individual through the entire employment journey more effectively than a single generalist AI.

Leveraging Language Models and External Tools: Another key pattern in agentic AI systems is the use of large language models (LLMs) as the 'brains' underpinning each agent. These pretrained models allow agents to engage in open-ended language interactions, interpreting natural queries, generating fluent replies, and making judgement calls.

However, agentic AI doesn't rely on language models alone. To ground agent knowledge and extend their capabilities, these systems also connect agents to external tools and data sources. Whether retrieving passages from the web, querying structured databases, or calling third-party APIs, agents use real-world information to inform their decisions and actions.

This combination of linguistic flexibility and external grounding allows agentic AI systems to maintain coherent dialogues while drawing insight from the broader world – a key step towards replicating how humans use language as a gateway to knowledge and action.

Managing Agent State and Workflows: Perhaps the most varied aspect of agentic AI design is how platforms handle the state and workflow orchestration of their agent teams. Since agentic tasks often involve many steps and dependencies between agent outputs, maintaining a coherent global state and control flow is crucial.

Approaches to this challenge vary across platforms. LangGraph uses an explicit directed graph to define workflows, giving developers fine-grained control. AutoGen relies more on emergent workflows arising from multi-turn conversations between agents. crewAI falls somewhere in between, with high-level task flows that guide agent interactions but flexibility for agents to autonomously delegate and respond to subtasks.

Despite these differences, some consistent priorities emerge for agentic state and workflow management:

- Providing a mechanism for agents to build on each other's work and decisions over time
- Enabling flexible definition of task division and agent coordination patterns
- Allowing task-specific customisation of agent roles, tools, and delegated authority
- Gracefully handling exceptions and nonlinear dependency graphs between agent outputs

As we can see, the rise of agentic AI brings with it tremendous potential for flexible, intelligent automation. By leveraging specialisation, collaboration, and the power of language models grounded in external data, these systems can tackle complex, open-ended tasks in ways that were previously out of reach for traditional AI approaches.

However, this potential also comes with significant challenges. Chief among these is the problem of planning: how do we enable these diverse teams of AI agents to effectively coordinate their actions, make decisions under uncertainty, and adapt their strategies in dynamic environments? This brings us to the core challenge that lies at the heart of agentic AI systems.

Planning as the Core Challenge

As agentic AI systems grow in complexity and capability, the need for effective planning becomes increasingly critical. Planning in this context refers to the process by which AI agents determine sequences of actions to achieve their goals, coordinate with other agents, and adapt to changing circumstances. While planning is a fundamental aspect of intelligent behaviour, it poses particularly difficult challenges in the domain of agentic AI.

Why is planning so challenging for AI systems, especially in multi-agent scenarios? There are several key factors that contribute to this difficulty:

1. High-dimensional state and action spaces: In agentic AI, the state space (all possible configurations of the environment and agents) and action space (all possible actions agents can take) are extremely large and complex. This is due to the combinatorial explosion that occurs when multiple agents, each with their own capabilities and potential actions, interact in open-ended environments.
2. Partial observability: Agents often have incomplete information about the state of the environment and the actions of other agents. This uncertainty makes it difficult to predict the outcomes of actions and plan effectively.
3. Non-stationary environments: In multi-agent systems, the environment is constantly changing as agents take actions and interact with each other. This non-stationarity means that the effects of actions can be inconsistent over time, complicating the planning process.
4. Long-term dependencies: Many tasks in agentic AI require long sequences of actions with dependencies between steps. Planning over these extended time horizons is computationally challenging and requires balancing immediate rewards with long-term goals.
5. Coordination and communication overhead: Effective planning in multi-agent systems requires coordination between agents, which introduces additional complexity and potential bottlenecks in the decision-making process.

To address these challenges, researchers have turned to formulating the planning problem in **agentic AI as a Markov decision process (MDP) (arxiv.org/abs/2406.14283)**. An MDP provides a mathematical framework for modelling decision-making in situations where outcomes are partly random and partly under the control of a decision-maker.

In the context of agentic AI, we can define the components of the MDP as follows:

- **State space (S):** The space of all possible thought processes and environmental configurations
- **Action space (A):** All possible combinations of thoughts or document retrievals
- **Transition dynamics (P):** How new thoughts are generated based on previous thoughts and actions
- **Reward function (R):** Evaluation of the quality of answers or progress towards the goal
- **Discount factor (Y):** Prioritisation of short-term vs. long-term rewards
- **Problem horizon (T):** Maximum number of reasoning steps allowed

By framing the planning problem as an MDP, we can leverage a wide range of techniques from the field of reinforcement learning to address the challenges of planning in agentic AI. However, this formulation also highlights a fundamental tension in the planning process: the exploration-exploitation dilemma.

The exploration-exploitation dilemma refers to the trade-off between exploring new, potentially better solutions and exploiting known good solutions. In the context of agentic AI planning, this manifests as a balance between:

- **Exploration:** Trying out new combinations of thoughts, retrieving diverse documents, or pursuing novel lines of reasoning that might lead to breakthrough solutions.
- **Exploitation:** Focusing on known effective strategies, building upon successful thought processes, or refining existing solutions to maximise immediate rewards.

Finding the right balance between exploration and exploitation is crucial for effective planning in agentic AI systems. Too much exploration can lead to wasted computational resources and inconsistent performance, while too much exploitation can result in suboptimal solutions and an inability to adapt to new situations.

Traditional planning approaches, such as those based on symbolic AI or exhaustive search, often struggle to address these challenges in the context of agentic AI. These methods typically rely on complete knowledge of the environment, deterministic action outcomes, and clearly defined goal states – assumptions that rarely hold in the complex, uncertain, and open-ended domains where agentic AI operates.

What's needed instead is a flexible, adaptive approach to planning that can:

1. Handle high-dimensional state and action spaces efficiently
2. Deal with partial observability and uncertainty
3. Adapt to non-stationary environments
4. Plan over long time horizons with complex dependencies
5. Balance exploration and exploitation dynamically
6. Coordinate actions between multiple specialised agents

This is where reinforcement learning enters the scene, offering a powerful set of techniques that are well-suited to addressing the unique planning challenges posed by agentic AI systems.

Reinforcement Learning and Advanced Techniques as Solutions

Reinforcement learning (RL) has emerged as a promising approach to tackle the complex planning challenges in agentic AI. RL is a type of machine learning where an agent learns to make decisions by interacting with an environment, receiving feedback in the form of rewards or penalties.

This learning paradigm is particularly well-suited to the planning problems in agentic AI for several reasons:

1. **Learning from experience:** RL agents can learn optimal strategies through trial and error, without requiring a complete model of the environment. This is crucial in the complex, partially observable domains where agentic AI operates.
2. **Balancing exploration and exploitation:** RL algorithms have built-in mechanisms for managing the exploration-exploitation trade-off, allowing agents to discover new strategies while also leveraging known good solutions.
3. **Handling uncertainty:** RL methods are designed to work in stochastic environments, making them robust to the uncertainties inherent in multi-agent systems.
4. **Long-term planning:** Many RL algorithms are explicitly designed to optimise for long-term rewards, allowing them to plan over extended time horizons and capture complex dependencies between actions.
5. **Adaptability:** RL agents can continuously update their strategies based on new experiences, making them well-suited to non-stationary environments.

One particularly powerful RL technique that has shown promise in addressing planning challenges is Monte Carlo tree search (MCTS). MCTS is a heuristic search algorithm that combines tree search with random sampling to make decisions in complex spaces. It has been successfully applied in various domains, including game-playing AI like AlphaGo.

In the context of agentic AI planning, MCTS can be used to efficiently explore the vast space of possible thought processes and action sequences. The key steps in MCTS are:

1. **Selection:** Traverse the tree from the root using a tree policy (e.g., Upper Confidence Bound) to balance exploration and exploitation.
2. **Expansion:** Add a new child node to expand the tree.
3. **Simulation**: Run a random simulation from the new node to estimate its value.
4. **Backpropagation:** Update node statistics along the path back to the root.

By iteratively applying these steps, MCTS can focus computational resources on the most promising regions of the search space, making it well-suited to the high-dimensional state and action spaces encountered in agentic AI planning.

Another key RL concept that can be applied to agentic AI planning is Q-learning. Q-learning is a model-free RL algorithm that learns to estimate the expected cumulative reward (Q-value) for taking a specific action in a given state. In the context of agentic AI, we can use Q-learning to estimate the value of different thought processes or document retrievals.

Recent advancements in the field have led to the development of several innovative approaches that build upon these foundational RL concepts to address the specific challenges of planning and reasoning in agentic AI systems.

Let's explore three cutting-edge techniques that show particular promise:

Q*: Improving Multi-step Reasoning with Deliberative Planning (arxiv.org/abs/2406.14283)

The Q* framework, introduced by Wang et al. (2024), represents a significant leap forward in improving the multi-step reasoning capabilities of large language models (LLMs). Q* combines the power of A* search with learned Q-value models to guide LLMs in selecting the most promising next steps during complex reasoning tasks.

Key features of Q* include:

1. Modelling the reasoning process as a graph, with each node representing a partial solution to the given problem.
2. Using a learned Q-value model as the heuristic function for A* search, estimating how promising each potential next step is for solving the overall problem.
3. Employing Monte Carlo tree search (MCTS) to efficiently explore the vast space of possible reasoning paths.
4. Incorporating a self-evaluation mechanism where the LLM scores its own refined answers, allowing for continuous improvement of the reasoning process.

The Q* framework addresses several critical challenges in agentic AI planning:

- **Handling long contexts:** Q* can process large batches of documents from knowledge sources, overcoming the limitations of fixed context windows in traditional LLMs.
- **Robustness to irrelevant information:** By exploring multiple branches of reasoning, Q* is resilient against unsuccessful information retrieval and misleading documents.
- **Adaptability:** The framework can be applied to a wide range of reasoning tasks without task-specific fine-tuning of the underlying LLM.

Experimental results have shown that Q* significantly outperforms baseline methods on various mathematical reasoning and code generation tasks, demonstrating its potential to enhance the planning and reasoning capabilities of agentic AI systems.

LLM Compiler for Parallel Function Calling (arxiv.org/abs/2312.04511)

While Q* focuses on improving the reasoning process itself, the LLM Compiler approach tackles another crucial aspect of agentic AI planning: efficient orchestration of parallel function calls. This technique, inspired by classical compiler design, aims to optimise the execution of multiple function calls in LLMs.

Key aspects of the LLM Compiler approach include:

1. Automatic decomposition of user inputs into a series of tasks with their inter-dependencies.
2. Parallel execution of independent tasks, significantly reducing latency in complex workflows.
3. A planning phase that creates a directed acyclic graph (DAG) of tasks, allowing for efficient scheduling and execution.
4. Integration with external tools and APIs, extending the capabilities of LLMs beyond pure language processing.

The LLM Compiler addresses several important challenges in agentic AI planning:

- **Efficiency:** By identifying parallelizable patterns and managing function call dependencies, the compiler can significantly reduce the latency of complex tasks.
- **Scalability:** The approach is designed to handle large-scale and complex tasks that involve multiple function calls and data dependencies.
- **Flexibility:** The compiler can adapt to different types of LLMs and workloads, making it a versatile tool for various agentic AI applications.

Early results have shown that the LLM Compiler can achieve substantial speedups compared to sequential execution methods, with latency improvements of up to 3.7x and cost savings of up to 6.7x on certain tasks.

Monte Carlo Tree Self-refine for Mathematical Olympiad Solutions (arxiv.org/abs/2406.07394)

Building on the success of MCTS in other domains, researchers have developed a Monte Carlo tree Self-refine (MCTSr) algorithm specifically tailored for tackling complex mathematical reasoning tasks, such as those encountered in mathematical Olympiads.

Key features of MCTSr include:

1. Integration of large language models with Monte Carlo tree search to enhance problem-solving capabilities.
2. An iterative process involving selection, self-refine, self-evaluation, and backpropagation steps.
3. A feedback-guided refinement process that allows the model to iteratively enhance its solutions.
4. A strict and critical scoring mechanism to ensure only genuinely improved solutions receive high scores.

MCTSr addresses several challenges in mathematical reasoning and planning:

- **Handling complex, multi-step problems:** The algorithm is designed to tackle intricate mathematical tasks that require multiple reasoning steps and strategic thinking.
- **Continuous improvement:** Through its self-refine and self-evaluation mechanisms, MCTSr can progressively enhance the quality of its solutions.
- **Adaptability to different problem types:** The framework has shown success across various mathematical domains, from grade school arithmetic to Olympiad-level challenges.

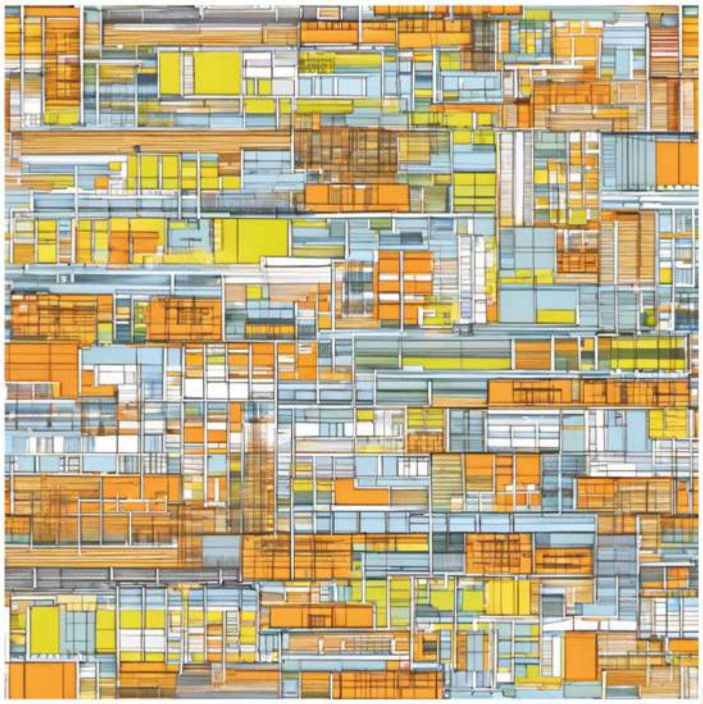
Experimental results have demonstrated that MCTSr can achieve GPT-4 level performance on mathematical Olympiad problems using much smaller models like LLaMA-3 8B, showcasing its potential to dramatically improve the reasoning capabilities of AI systems.

These three approaches – Q*, LLM Compiler, and MCTSr – represent the cutting-edge of planning and reasoning techniques in agentic AI. By combining reinforcement learning principles with innovative search and optimisation strategies, these methods are pushing the boundaries of what's possible in AI-driven problem-solving.

However, it's important to note that the application of these advanced techniques to agentic AI planning is not without challenges:

1. **Computational complexity:** These methods often involve intensive computational processes, which can be resource-demanding for large-scale applications.
2. **Balancing exploration and exploitation:** Finding the right balance between exploring new solutions and exploiting known good strategies remains a delicate task.
3. **Interpretability:** As these systems become more complex, ensuring transparency and interpretability in their decision-making processes becomes increasingly challenging.
4. **Generalisation:** While these approaches have shown impressive results in specific domains, further research is needed to assess their generalisation capabilities across diverse task types.

Art by AI



HOW SMEs CAN NAVIGATE THE DATA & AI LANDSCAPE

FRANK STADLER is a senior data leader, mentor and advisor based in Augsburg, Germany. He's spent around 15 years in various software engineering roles, consulting for BMW, developing RFID-software for industrial gas companies and leading a team working on a medical product software, decreasing the mortality for thousands of chronically ill people. About 6 years ago Frank shifted his focus to data science. Ever since, he's been delivering custom data science solutions in many projects.

What are some common misconceptions small and medium enterprises have about AI and data science in general?

There are two common misconceptions: one is that AI and data science are optional and don't really generate advantages for the business. Many SMEs haven't realised that AI is going to be crucial for success.

The second is that many companies, especially SMEs, think you need lots of data to be able to generate value from AI and data science projects. And that's not accurate: the data needs to be good, but you don't need to have big data.

That's an interesting point, because the conventional wisdom is that machine learning models, for example, require lots of data to train them. Can you elaborate?

It depends on the amount of data, and also the kind of data that you have. If you're doing forecasting, for example, you need historic data, but that doesn't necessarily need to be big data. You could just have a couple of million rows and you'll be able to achieve very good results with classical machine learning.

> *Many SMEs haven't realised that AI is going to be crucial for success.*

It's true that if you're using deep learning models, then you'll probably need more data. But machine learning models can work with smaller data sets and generate good value from them.

Are there any other parameters, in terms of size, that are the minimum guidelines for ML models to work effectively?

It really depends on the data set. I would say that the quality and the completeness of the data is more relevant than the size and the amount of data. You need to have the right data and the right data points, and the right features and parameters.

If you don't have the relevant data – for example, if you're trying to do daily forecasts, but you only have monthly sales numbers – then that's not going to work out.

What are the key challenges SMEs typically face when implementing AI and data science projects?

Usually, it's a lack of technical expertise and subject-matter expertise. SMEs often don't have many people with a statistical and data science background on their payroll. And they also often don't have people knowledgeable in the technologies commonly used.

The second challenge many SMEs face is whether they have the support from senior management that's required to make sure implementing the AI aligns with company goals.

Thinking about where they should start, how can smaller companies identify opportunities where AI and data initiatives can make an impact?
I would advise that they start looking at their core business processes and their strategic goal – what they're trying to accomplish. Next, they should hold brainstorming sessions with the relevant stakeholders. Think of ways they can use their data to improve upon those processes, and maybe even create new processes that support their core operations and their core strategy.

Are there any use cases that you see come up more regularly, that SMEs should think about first?
It depends on the business case. If you're selling things, then of course there's always an AI use case for forecasting and optimisation of logistics and storage. If you're working with services and working with customers directly, then areas like customer segmentation, customer churn and customer lifetime value calculations are where AI and data science can be of most benefit.

What are the risks associated with AI and data science for SMEs, and how can they mitigate them?
One of the biggest risks, especially for smaller companies, is the legal side; the issues around data privacy and data security. Once you start working with data sets, you need to make sure that you're allowed to use them for those kinds of use cases. You also need to make sure that there are no data breaches or confidentiality issues.

> *One of the biggest risks, especially for smaller companies, is the legal side; the issues around data privacy and data security.*

Can you give us an example of where an SME might get tripped up by this legal aspect of data security?
One issue many companies might face is if they collect technical data from their products for more generic purposes, but they then try to use that data on an individual basis, to create personal recommendations for that customer.

That's a common use case where you're not really allowed to use the data where individual customers could be identified. If the product is an entertainment product that sends usage data back to the company, that could also be critical or medical data, such as data from health devices.

What are some of the common pitfalls SMEs should avoid when taking on AI and data science projects?
A common issue is overspending on a technical solution that's needlessly big and complex. Small businesses should start small initially, but also have a growth plan in mind so they can expand on their current projects.

Another potential pitfall is that businesses spend too much money on projects that don't have ROI. So the project might be very modern and much-hyped, but in the end, if there's no return that helps support your business processes then it's money wasted.

How do businesses go about assessing whether a potential project is likely to have a positive ROI?
They should start by carefully considering what KPIs they should measure before they implement the solution and once it's been implemented. Also, think in advance about how AI and data science could improve upon current processes. There needs to be a theoretical way that the impact will be quantified. And it needs to be measurable afterwards.

Ideally, the new project shouldn't be too big or complex, because that increases the risk it will fail. It should be somewhat basic and isolated, but not too isolated that it's not relevant for other processes or other use cases later on.

I would say a mid-sized topic that's not too complex is probably a good way to start. It's also a good way to gauge whether the company is ready for these kinds of projects. How is the team structure? How is the technical infrastructure? How is the data quality? If your data infrastructure and the data quality aren't up to the job, then you've wasted a lot of time and money.

Where should an SME that's new to data science and AI start, in terms of their data infrastructure?
It's a good start to identify the data sources you have in the company. Usually they're multiple and heterogeneous, and not accessible through one single interface or technology. Also consider what you want to achieve. It's usually better to do data analysis first and check what kind of insights and benefits you can generate out of that before you start doing more advanced stuff like data science or AI.

To understand what kind of data is available and what kind of insights can be generated from it, the data should be stored in a unified platform like a data warehouse or a data lake. One single point where they can access the data. It's easier to build the technology stack around that.

And in terms of a solid data infrastructure, what's required in addition to this central repository?

The main points to consider are: where is the data coming from and where do we want to store it? So in the cloud, you could use Snowflake. Otherwise, you'd need some kind of automatic data pipeline to make sure that the data is transferred from the source system to the warehouse or the data lake, where it can then be aggregated, quality-checked and improved for the use cases of data analysis and data science.

What size and structure should a data team be in an SME? What's the minimum number of employees a smaller business can work with?

That depends a lot on the skill set and the requirements you have. I myself have started out as a one-person data team in the past, but I also had the necessary software engineering background to make sure that I could deal with all the topics that came up. So it's possible to start with one or two people. But you need to have at least one person with an engineering skill set; a data engineer or software engineer who can make sure you build up the infrastructure.

The next person I would probably look out for is a data analyst or some similar role, so you're able to analyse and generate insights from your data.

And only after you have those specialists in place would you consider looking at data scientists for more complex use cases and projects.

Who should the data leader ideally report to in an SME?

Ideally the CEO. It should be a core topic for the business. If that's not viable, another option would be the COO or the CPO if they have one; whoever is closest to the real business operations and the product or the services the company provides.

I don't think the CTO is necessarily the best role to report to, because the deciding factor for the impact and the success of data science projects and data analysis is not the technology that's being used, but rather the impact it has on the business. And the CTO is commonly very focused on the technology and not the business processes.

How important is it that the first hire has domain knowledge of that area?

People might disagree with me here, but I would argue it's not too important because they can learn the necessary domain knowledge on the job. Technical knowledge is more important, especially for the first person hired. That's provided there are people available who can assist with the necessary domain knowledge, and can teach the first hire as well.

Thinking about SMEs that are too small to hire a dedicated technical employee, how can they get started? Is it possible for them to build a data infrastructure?

It's certainly possible to allocate a part-time role to an existing technical employee. I have experienced that myself in the past. Otherwise, it could also be possible to work with external resources that support the company for at least a couple of days or weeks to give them a head start to continue working on this on their own.

How can SMEs leverage AI and data to improve customer experience and retention?

Customer experience is very closely related to individualisation. So if the SME is able to use the data they have for each individual customer and use that to improve upon the services or products provided, then that will help keep the retention high.

You mentioned earlier that data quality is more important than the size of the data set. How do SMEs ensure they have high-quality data?

SMEs need to check their data quality from the outset. And you have to work closely with the people who are actually generating the data and inputting the data into the systems. These people need to have good data literacy and be made aware of the impact that the data quality will have later on. If the people generating or inputting the data realise the value that good data quality has, then you can ensure that the quality stays high.

And data quality should always be as high as possible as upstream in the process as possible. There may be some quality issues that can be fixed later on, but it will always be better to have initial high-quality raw data.

What's your take on SMEs leveraging GenAI?

A lot of companies, especially in the SME arena, try GenAI because of the hype and the promise that it's easy to implement (because it's pre-trained). They go to Hugging Face and download an open source model, and then they try some RAG implementation with their own data.

The probability of this type of initiative delivering real value for SMEs is usually low, unless they already have very good data infrastructure and data management in place. That's because the results are very dependent on the data quality that you use as input. So if the SME is already very advanced in their data architecture and data quality, they might have good results. Otherwise, the results will probably be subpar.

Do SMEs need a separate database for GenAI? Are there additional and infrastructure requirements for GenAI projects?

Most often, yes, but it also depends on the infrastructure you have in place already. For example, if you want to do a RAG use case, then of course you need a vector database. But there are also vector plugins for existing databases, and additional tools to circumvent the need for a completely new database system.

How much data is needed for a successful RAG implementation in a small company?

That very much depends on the use case. But it doesn't have to be much. It's more a case of the right vectorisation and the right chunking of the data. And it needs to be relevant to the use cases and the questions you want to answer with the system. So more data doesn't necessarily perform better or create better results. It needs to be the right data and it needs to be prepared well.

Would you argue that most of these small-scale RAG implementations are just a glorified enterprise search?

There are certainly many projects where this is the case. And in many instances, when you try to implement a RAG, you will have to combine it with more classical methods of searching as well. So if you only have a limited number of documents, having a regular search or some kind of software-based search tool might be more efficient.

SMEs need to check their data quality from the outset.

What advice can you give about maintaining these models? SMEs sometimes assume they're just like standard software that can be left alone once built.

So there are multiple issues that could be problematic. First of all, if you use apps to access public LLMs and OpenAI, those will change without notice and can create different results than you would previously have expected. And you also need to make sure that the vectorisation of your documents is up-to-date. So if new documents get added, then they will have to be added to the vectorisation database as well. And old documents might need to be removed. You need to have processes in place to take care of that.

ZIAD AL-ZIADI is the Co-founder and CPO at Banqora, an applied-AI fintech automating post-trade processing banks, asset managers, and hedge funds. Previously, he led data science product development at Channel 4, implementing recommender systems and ad-targeting models for over 29 million users.

As Product Lead at Raft, he managed a team of over 10 machine learning engineers deploying NLP and computer vision models for large-scale document processing. As a researcher at the buy-now-pay-later unicorn Zilch, he developed their first machine learning model for credit risk. He also held various product roles at ClearScore after exiting his startup, Paperclip.

LLMS AREN'T PRODUCTS:
THE CHALLENGE OF PRODUCTIONISATION AND DELIVERING VALUE

Let's face it: the technology scene is drunk on AI right now, and LLMs are the drink of choice. Organisations are scrambling to integrate shiny new models into their workflows, but here's a sobering thought: LLMs aren't magic bullets, they're just sophisticated hammers. And not every problem is a nail.

LLMs themselves aren't products. They're foundational technologies that need sincere and serious development before they can deliver any real value. It's like giving an intern a supercomputer – there's potential, but also a lot of hand-holding required. The challenge isn't just technical implementation, although that's a hurdle in itself. It's about creating a user experience that doesn't frustrate people or put them at risk. Building a functional AI system is one thing, but making it actually useful? That's where the real work begins. Unlike most conventional software, LLMs require a tailored approach rather than a simple plug-and-play solution.

Building a system that can generate human-like text is impressive, but it's only the first step. The key is to harness that capability in a way that solves real problems and enhances existing processes. This requires a deep understanding of user needs, industry-specific challenges, and the limitations of the technology itself.

Let's consider legacy sectors such as finance and government that aren't known for their agility. They have many tangled workflows that capture complex nuances and processes, and for good reason too. These aren't just arbitrary processes – they've been developed over decades, often in response to specific regulatory requirements or hard-learned lessons. You can't simply overlay an AI solution on top of these complex systems and expect seamless integration.

The challenge here is twofold. First, there's the technical aspect of integrating AI capabilities into existing systems. But perhaps more importantly, there's the human element. How do you convince professionals who've spent their careers mastering these workflows to trust and adopt AI-driven solutions? This transition requires a delicate balance of innovation and respect for established practices. It's not about replacing human expertise but augmenting it in meaningful ways.

One of the most intriguing challenges in developing LLM-based products is dealing with their non-deterministic nature. In other words, these models can be unpredictable. This presents a unique hurdle for user experience design. How do you create an interface that's intuitive and reliable when the underlying technology might produce different results each time? It's like designing a car where the steering wheel might turn into a joystick – the user experience becomes jarring and dangerous.

Product designers need to carefully consider where human intervention is necessary. This might involve creating checkpoints for user validation, designing interfaces that provide context for AI-generated content, or developing systems that allow for easy human override when needed. It's a delicate dance.

The next wave of innovation in LLMs and broader GenAI may not be centred solely around datasets, models, and computing. While the technical advancements in AI are undoubtedly impressive, I believe the next wave of true innovation in this field may come from unexpected places. While the tech world is obsessed with bigger models and faster processors, I think the real revolution might be brewing in boardrooms and courthouses.

It's not enough to have a powerful AI model – you need to understand how it fits into your users' lives and workflows.

Innovation could be a lawyer figuring out how to make AI-generated text legally binding, or a business strategist developing a new pricing model for AI. The technology is great, but the real game-changers might come from people in suits. While significant progress has been made on the performance front through the development of larger models and robust technical infrastructure, we should begin to focus on creating core value through innovative procedures. These are the kinds of factors that will shape the future of AI implementation. And they'll require input from a diverse range of professionals – not just engineers and data scientists, but lawyers, ethicists, business strategists, and more.

For most organisations, those vaunted foundation models are just raw ingredients. Unless you're in the rarefied air of OpenAI or DeepMind, your job isn't to create the AI – it's to cook with it. The real value comes from specialisation. It's the difference between having a Swiss Army knife and a surgeon's scalpel. Both are useful, but when you need to perform open-heart surgery, you know which one you'd prefer.

This process of specialisation is where the true potential of LLMs can be realised. It's the difference between a general-purpose tool and a precision instrument designed for a specific task. It's the responsibility of product teams to post-train these models, moving them from being generally capable to specialising in specific product offerings. This might involve additional training on industry-specific data, fine-tuning for particular use cases, or integrating them with proprietary systems and workflows.

At its core, building a successful product – AI-powered or not – comes down to solving real problems for real users. We need to resist the temptation to start with the technology and work backwards. Instead, the focus should be on identifying genuine user needs and pain points, and then exploring how AI can address these issues.

This approach requires empathy, user research, and a willingness to iterate based on feedback. It's not enough to have a powerful AI model – you need to understand how it fits into your users' lives and workflows. Put differently, the hardest problems that an organisation can tackle cannot solely be solved with today's foundation models. While LLMs offer powerful capabilities, they are tools to be wielded in service of solving real user problems.

As with any technology product, experimentation is crucial for the development of LLM-based solutions. The relative ease of spinning up and testing LLM systems compared to traditional software development offers an opportunity for rapid iteration and learning. However, this ease of experimentation should not lead to a lack of rigour in evaluation.

Running evaluations should be considered a core component of any LLM-based product's specification. These evaluations need to be rooted in the domain or problem space the product is addressing. It's not sufficient to rely solely on unit tests or measures of data drift, for example; instead, product teams must test whether LLM outputs meet acceptable end-user criteria. This might involve domain experts reviewing AI-generated content, measuring improvements in workflow efficiency, or assessing the accuracy and relevance of AI-assisted decision-making.

Building effective LLM products requires a diverse team of professionals. It's not enough to hire machine learning engineers with the hope of shipping a functional product. You need professionals who understand the underlying infrastructure required to support AI systems. This includes expertise in data collection and storage, backend and frontend development, and the ability to design systems that can effectively consume and render model outputs.

Moreover, product teams should include members with expertise in UX design, domain knowledge, ethics, and legal compliance to ensure that LLM-based products are not only technically sound but also user-friendly, trustworthy, and aligned with regulatory requirements.

Many organisations have genuine use cases and opportunities where workflows can be improved by LLMs. However, this isn't an excuse to retrofit organisation-wide workflows with vanity AI features that please executives. Not every process needs to be AI-powered, and not every AI implementation will deliver meaningful value.

Organisations should be strategic in how they approach LLM integration. This means carefully evaluating potential use cases and prioritising those that align with core business objectives and user needs. The process should involve:

1. Conducting thorough user research to identify pain points and opportunities
2. Mapping existing workflows and identifying areas where AI can add significant value
3. Prioritising LLM implementations based on their potential impact and feasibility
4. Developing clear success metrics that tie back to organisational goals and user satisfaction

It's crucial not to lose sight of the human element. The most successful AI implementations will be those that enhance human capabilities rather than trying to replace them entirely.

Product designers need to think carefully about how AI and human expertise can work together synergistically. This might involve using AI to handle routine tasks, freeing up humans to focus on more complex, nuanced work. Or it could mean using AI as a tool to augment human decision-making, providing additional insights and analysis. The key is to design systems that leverage the strengths of both AI and human intelligence, creating workflows that are more efficient and effective than either could achieve alone.

As we navigate this new landscape, let's not lose sight of what truly matters: creating products and solutions that make people's lives better.

As we stand on the brink of this AI revolution, it's clear that we're entering uncharted territory. The potential of LLMs is enormous, but so are the challenges we face in harnessing this technology effectively.

We need to approach this new frontier with a combination of excitement and caution. Yes, we should be exploring the possibilities of AI and pushing the boundaries of what's possible. But we also need to be mindful of the potential pitfalls and unintended consequences.

Success in the AI era will require more than just technical prowess. It will demand creativity, empathy, and a deep understanding of human needs. It will require us to rethink our approaches to product design, team building, and problem-solving. As we navigate this new landscape, let's not lose sight of what truly matters: creating products and solutions that make people's lives better. It's up to us to use these technologies responsibly and shape a future where AI serves humanity, not the other way around.

EDITORIAL INTELLIGENCE:

COMPREHENSIVE APPROACHES AND METHODOLOGIES FOR MODERN AND STRATEGIC CONTENT PUBLISHING

DR ANA MOYA is a data scientist and analytics expert. Her career is based on a solid foundation in statistics and data science, culminating in a doctorate awarded by the Technical University of Dortmund.

During her subsequent 15+ year tenure in the data field at the FUNKE Media Group and Handelsblatt Media Group, she worked on a variety of projects, including data integrations, the development of data and text mining algorithms, user research, and the application of statistical and advanced AI models in order to derive strategic recommendations.

Ana is also a sought-after academic lecturer, having conducted research on the world of data at TU Dortmund for nearly 20 years. Since 2018, she has been teaching 'Data Science and Business Intelligence' at the International School of Management in Germany. Additionally, she is an author of books and articles, and addresses the application of statistical methods in data journalism newspaper articles.

Beyond data journalism, statistics derived from news offerings drive the core work of editorial teams: selecting topics, shaping coverage, and publishing content. This is particularly evident in online journalism, where search engines and social media exert increasing influence. The growing use of mobile devices introduces diverse usage scenarios and reader needs. News media and publishers aim to reduce dependence on the advertising market, increasing the need to better understand readers and subscribers.

With changing media consumption patterns, editorial teams must adjust their techniques to stay relevant and achieve desired outcomes. This involves navigating a dynamic environment while maintaining content excellence and honesty, using contemporary techniques to meet audience expectations and maintain competitiveness.

THE ROLE OF DATA IN EDITORIAL BUSINESS

Data touches every part of a publishing business, from content creation to customer engagement, helping to understand and optimise editorial processes. Historically, the tangible nature of printed newspapers made the business more concrete. Now, with digital content, new forms of presenting information, such as apps, websites, and newsletters, are essential.

In subscription-based businesses, ERP (enterprise resource planning) systems track subscription details, such as billing cycles, renewal dates, and account statuses. CRM (customer relationship management) systems focus on managing customer communications and tracking interactions.

Beyond these core systems, additional data sources provide valuable insights into reader behaviour and engagement. Analytics tools monitor how users interact with content on websites, apps and so on, capturing metrics such as page views or time spent on articles.

Additionally, content data from CMS (content management systems) play a pivotal role by storing and organising article metadata, such as publication dates, author information, categories, and tags.

Integrating and analysing diverse data sources gives publishers a comprehensive understanding of audience preferences and behaviours. A data-driven approach in journalism involves making decisions based on quantitative data and analysis. This method, encompassing algorithms, models, and historical performance metrics, places the success or failure of published content at the forefront of decision-making. While statistical evaluations identify trend topics and audience preferences, they are considered alongside journalistic criteria such as objectivity, balance, and timeliness, making the strategy data-informed rather than purely data-driven.

FROM DATA TO WISDOM: INTELLIGENCE FROM DATA

The process of transforming raw data into valuable knowledge involves stages illustrated by the DIKW (data, information, knowledge, wisdom) pyramid. This progression helps understand past events, predict future trends, and derive actionable insights for intelligent decisions.

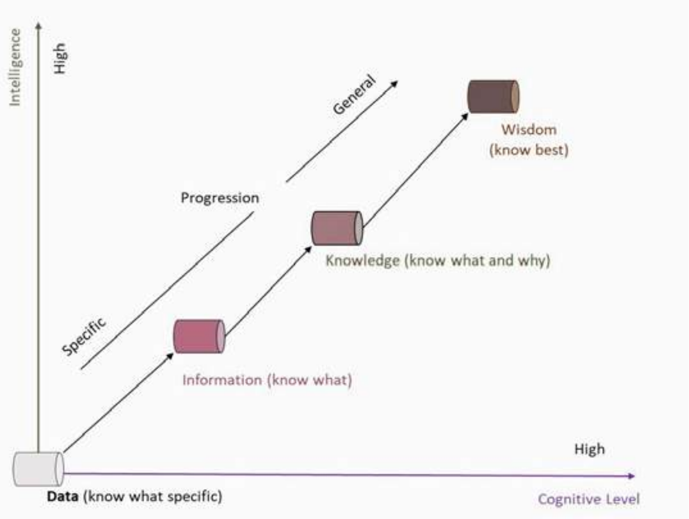
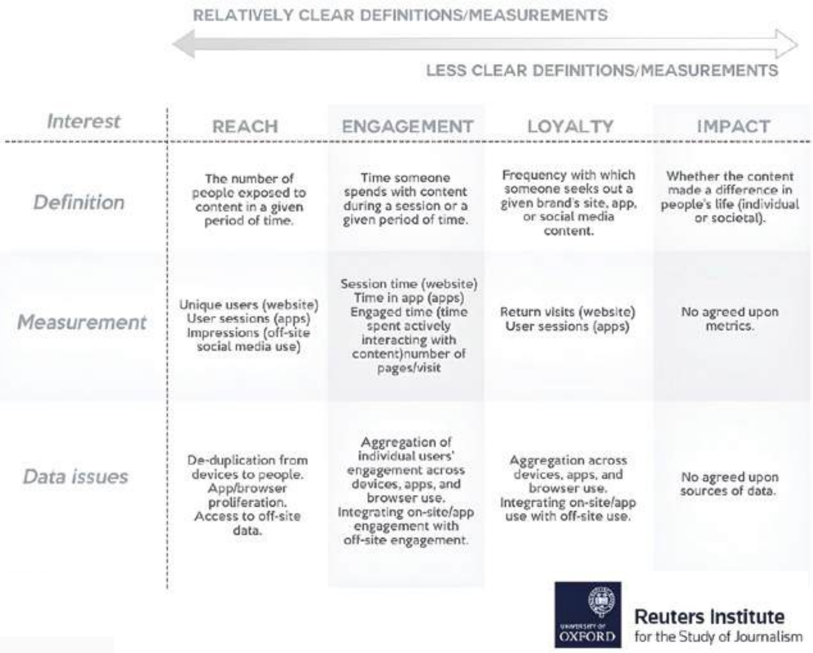


FIGURE 1: Cognitive Progress: Based on specific data, general knowledge is eventually acquired (according to Cao 2018)

The development of data into wisdom occurs through application of progressive analyses, predominantly predictive analytics, not only to understand the past, but also to make forecasts for the future, to generate meaningful, action-oriented insights to derive 'intelligent' decisions.

Measuring success in data-informed strategies requires establishing key performance indicators (KPIs). These KPIs define essential metrics for evaluating performance and ensuring alignment with strategic goals.

Different metrics can thus serve different purposes, but some things are currently harder to measure than others. Some will likely always resist quantification. Some sources of data, like sessions, are used in attempts to understand quite different things, like reach versus engagement. These aspects are clearly visualised in the following figure:



RELATIVELY CLEAR DEFINITIONS/MEASUREMENTS

LESS CLEAR DEFINITIONS/MEASUREMENTS

Interest	REACH	ENGAGEMENT	LOYALTY	IMPACT
Definition	The number of people exposed to content in a given period of time.	Time someone spends with content during a session or a given period of time.	Frequency with which someone seeks out a given brand's site, app, or social media content.	Whether the content made a difference in people's life (individual or societal).
Measurement	Unique users (website) User sessions (apps) Impressions (off-site social media use)	Session time (website) Time in app (apps) Engaged time (time spent actively interacting with content)number of pages/visit	Return visits (website) User sessions (apps)	No agreed upon metrics.
Data issues	De-duplication from devices to people. App/browser proliferation. Access to off-site data.	Aggregation of individual users' engagement across devices, apps, and browser use. Integrating on-site/app engagement with off-site engagement.	Aggregation across devices, apps, and browser use. Integrating on-site/app use with off-site use.	No agreed upon sources of data.

UNIVERSITY OF OXFORD Reuters Institute for the Study of Journalism

FIGURE 2: A range of metrics mapped in terms of relative clarity of definitions and measures (according to Cherubini, 2020)

The effectiveness of data analysis hinges on how results (based on KPIs) are processed and communicated. Dashboards offer a graphical, interactive user interface that displays live data, while reports summarise the current status at specific times and are actively distributed. Both forms of communication should have clear goals and be tailored to recipients' needs. Consistent dashboards and comprehensive reporting ensure relevant teams and stakeholders are informed and can act on the insights provided.

EDITORIAL INTELLIGENCE

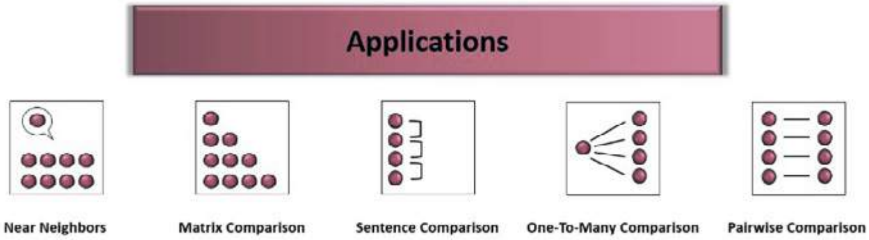
Editorial intelligence refers to a media organisation's ability to analyse, interpret, and generate predictive insights from collected data. This process supports data-informed editorial strategies, enhancing content quality and relevance. By integrating editorial intelligence, media professionals can make informed decisions that align with audience needs, continuously improving the editorial process. The ability to understand and use data competently is crucial for effective editorial intelligence.

Advanced analytics, including predictive and prescriptive methods, play a crucial role in informed decision-making. Predictive analytics uses historical data to forecast future events, such as user conversion rates or cancellation likelihoods, involving techniques like statistical modelling, machine learning, and text mining.

For instance, a project aimed at understanding subscriber reading habits might use clustering algorithms to segment users into groups based on their content preferences and engagement levels. Classification algorithms could then be employed to categorise new content into these segments, while regression analysis might predict how different types of content will perform. Text mining techniques can analyse and extract characteristics from articles like sentiment and identify trending topics.

A powerful text mining technique is latent semantic analysis (LSA). It evaluates documents and looks for the underlying meaning or concept of the documents. LSA would have an easy job if each word only had one meaning, but oftentimes, words are not only ambiguous but also are synonyms or have multiple meanings. One example could be the word 'may', which could be a verb, a noun for a month, or a name. To overcome this problem, LSA fundamentally compares how frequently the words appear together in one document and afterwards compares it across all other documents. By grouping words with other words, it tries to identify those words which are semantically related to each other and eventually to get the true meaning of ambiguous words. Near neighbours, comparison matrix, one-to-many comparison and Pairwise comparison are LSA methods, as stated by Rozeva and Zerkova, 2017.

FIGURE 3
Latent semantic analysis applications;
Source: Rozeva/Zerkova 2017:



Prescriptive analytics goes beyond prediction by providing actionable recommendations based on data insights. This involves using optimisation, AI, and simulation techniques to suggest specific actions that can enhance business processes and content strategies.

GENERATIVE AI IN EDITORIAL INTELLIGENCE

While generative AI is not yet a major component of Editorial Intelligence, traditional statistical models and algorithms remain essential tools for data analytics. Methods like robust text mining, clustering, and classification continue to offer reliable and efficient benefits.

Text mining, for instance, remains a highly reliable technique for extracting insights from unstructured data. Its robustness comes from its ability to handle large volumes of text efficiently, uncovering valuable patterns, trends, and sentiments that inform content strategies and audience engagement. This long-established method has proven its effectiveness over time and continues to deliver reliable results.

Additionally, traditional methods are resource-efficient, requiring less infrastructure and energy compared to newer technologies, making them environmentally sustainable.

CHALLENGES AND SUCCESS FACTORS

Ensuring high data quality is critical for implementing data-informed strategies. Poor data quality can lead to incorrect decisions. Practical steps to address this issue include implementing alerts for data anomalies and fostering responsibility among departments to maintain data quality.

Promoting data culture and literacy within departments is also essential, enabling individuals to understand and utilise data effectively, transforming insights into actionable strategies. A collaborative environment enhances internal processes. It's important to involve stakeholders in understanding the complexity and value of different analytical approaches. Transparency about the time and resources required for various types of analyses helps set realistic expectations and achieve better outcomes.

REFERENCES AND RECOMMENDED LITERATURE

Cao, Longbing: Data Science Thinking. Cham [Springer] 2018.

Cherubini, F.: Editorial Analytics: How News Media Are Developing and Using Audience Data and Metrics - Reuters Institute Digital News Report. 2020.

Rozeva, A.; Zerkova, S.: Assessing semantic similarity of texts – Methods and algorithms. In: Pasheva, V.; Popivanov, N.; Venkov, G. (eds.), 2017.

Shi-Nash, Amy; Hardoon, David R.: Data Analytics and Predictive Analytics in the Era of Big Data. In: Geng, Hwaiyu (Hrsg.): Internet of Things and Data Analytics Handbook. Hoboken [John Wiley & Sons], 2017.

FRANCESCO GADALETA

MIXING METHODOLOGIES: EXPLORING THE POWER OF PHYSICS-INFORMED NEURAL NETWORKS

FRANCESCO GADALETA is a seasoned professional in the field of technology, AI and data science. He is the founder of Amethix Technologies and Intrepid AI, building an AI-powered platform for autonomous robots. Francesco also shares his insights and knowledge as the host of the podcast *Data Science at Home*.

His illustrious career includes a significant tenure as the Chief Data Officer at Abe AI, which was later acquired by Envestnet Yodlee Inc. Francesco was a pivotal member of the Advanced Analytics Team at Johnson & Johnson. His professional interests are diverse, spanning applied mathematics, advanced machine learning, computer programming, robotics, and the study of decentralised and distributed systems. Francesco's expertise spans domains including healthcare, defence, pharma, energy, and finance.

We all recognise the precision of physics and the accuracy of neural networks. But have you ever considered that these two fields could be combined? This concept is now becoming a reality with the development of physics-informed neural networks (PINNs).

At first glance, merging these two disciplines might seem challenging to visualise. However, many recent physics discoveries, particularly those from the last decade, have been enabled by advancements in both hardware and software. Neural networks, on the other hand, have the ability to learn about the world, including the principles of physics, directly from data and observations. So, how exactly do physics and neural networks complement each other in this innovative technology? Let's explore this in more detail.

OPENING THE DOOR FOR CLOSED-FORM SOLUTIONS

Traditional physics models are typically developed by expert physicists who create representations of the world or the natural phenomena they wish to study. For instance, modelling aircraft dynamics involves equations that account for drag, gravity, thrust, acceleration, and orientation. In robotics, the complexity of the robots and their operating environments results in increasingly intricate sets of equations.

On the other hand, neural networks adopt a purely data-driven approach. They can model nonlinearities and do not require closed-form solutions. In other words, they don't need the phenomena they predict to be solvable through explicit mathematical expressions.

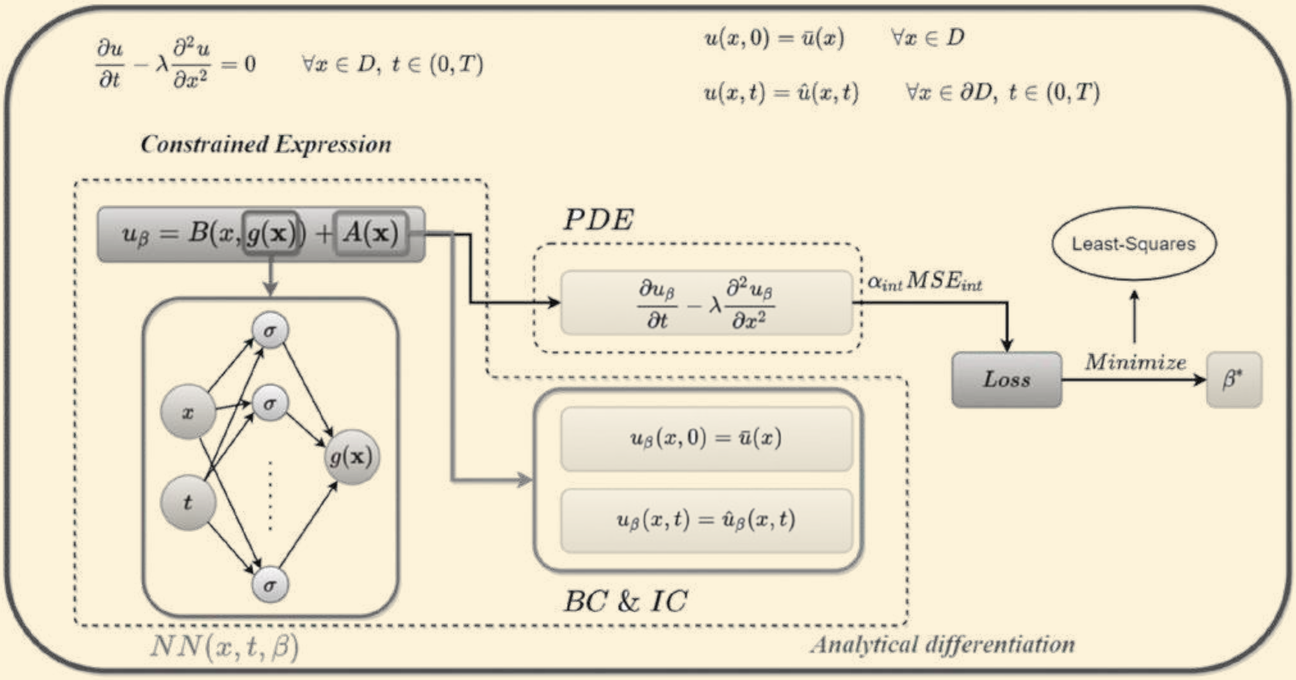
Closed-form solutions are highly desirable for mathematicians and physicists because they provide explicit solutions to problems. A closed-form expression or formula directly offers a solution using a finite number of standard mathematical operations, which may include logarithms, trigonometric functions, and more.

Conversely, non-closed-form solutions require approximations or numerical methods to find solutions that approximate the real answer. This is common in physics, where many problems do not have closed-form solutions. In such cases, function approximators like neural networks can be invaluable.

PINNS COMBINE OPPOSITES TO DELIVER GREATER POWER

Physics-informed neural networks (PINNs) are innovative models that integrate physics and neural networks. They exist at the intersection of these two fields, combining the data-driven prediction capabilities of neural networks with the model-driven precision of physics.

For physicists, accuracy hinges on the quality of the model and the equations used to represent real-world phenomena. In contrast, neural networks achieve accuracy by extrapolating from observations of the world and the natural phenomena they are predicting. Neural networks rely on data during both training and inference, making them data-oriented rather than model-oriented.


$\frac{\partial u}{\partial t} - \lambda \frac{\partial^2 u}{\partial x^2} = 0 \quad \forall x \in D,\ t \in (0,T)$
$u(x,0) = \bar{u}(x) \quad \forall x \in D$
$u(x,t) = \hat{u}(x,t) \quad \forall x \in \partial D,\ t \in (0,T)$
Constrained Expression
$u_\beta = B(x, g(\mathbf{x})) + A(\mathbf{x})$
PDE
$\frac{\partial u_\beta}{\partial t} - \lambda \frac{\partial^2 u_\beta}{\partial x^2}$
$\alpha_{int} MSE_{int}$
Least-Squares
Loss
Minimize
β^*
σ
x
t
$g(\mathbf{x})$
$u_\beta(x,0) = \bar{u}(x)$
$u_\beta(x,t) = \hat{u}_\beta(x,t)$
BC & IC
$NN(x,t,\beta)$
Analytical differentiation

By Mariodeff – Own work, CC BY-SA 4.0, https://commons.wikimedia.org/w/index.php?curid=117757726

PINNs blend these two fundamentally different approaches, creating a methodology that leverages the strengths of both. This hybrid approach can yield more powerful results than either method alone, offering the best of both worlds.

This principle is not new to machine learning. For instance, ensemble methods combine different models to learn various aspects of the same data, with each model contributing its unique insights. When combined, these models produce a more powerful and comprehensive solution than any single model could achieve on its own.

PINNS IN PRACTICE: PROJECTILE MOTION

To illustrate the concept of physics-informed neural networks (PINNs), let's consider the classic physics example of projectile motion. When you launch a projectile, such as a missile or rocket, what happens to its trajectory? How does its velocity, position, and acceleration change over time?

First, consider the displacement vector, which represents the position of the projectile at time T, typically as a function of time. The first derivative of this position gives the velocity vector, and the second derivative provides the acceleration. Simply put, the first derivative of velocity is acceleration.

For a projectile, the simplest formula for acceleration must account for drag – the resistance of air – and gravity. The formula for the acceleration of a projectile considering both drag and gravity is:

$$\vec{a} = -k\vec{v} - \vec{g}$$

Where:

- $\vec{a}$ is the acceleration vector of the projectile.
- k is the coefficient of drag.
- $\vec{v}$ is the velocity vector of the projectile.
- $\vec{g}$ is the acceleration due to gravity vector.

In this context:

- The term $-k\vec{v}$ represents the drag force per unit mass, which acts opposite to the direction of the velocity.
- The term $-\vec{g}$ represents the gravitational force per unit mass, which acts downward.

From physics we also know that:

- The drag force is proportional to the velocity and acts in the opposite direction.
- The gravitational force is constant and acts downward.

This equation explains the projectile's motion in terms of acceleration, velocity, and position over time. By integrating acceleration, you obtain velocity, and by integrating velocity, you return to position.

Using numerical integration methods such as Runge-Kutta[1], you can describe the trajectory of your projectile according to the physics model in use. This is, of course, a simplification, as all physics models are approximations and may not account for all forces at play. However, you can still estimate the landing point of the projectile with reasonable accuracy.

Now, what happens when you use a neural network to predict this trajectory? The neural network will attempt to predict future positions by analysing past data. As long as the neural network has sufficient data, its predictions for the next position or acceleration would be accurate.

The challenge arises when the neural network lacks data, causing its predictions to deviate from the physics model's trajectory. Missing data results in inaccurate or missed predictions, a very well-known problem with non-linear function approximators like neural networks. Usually, this issue can be mitigated with techniques like regularisation.

For example, L2 regularisation helps the network or model generalise by fitting through the data rather than overfitting it. Regularisation prevents the neural network from excessively conforming to the training data. Consequently, even with varying amounts of data, the neural network learns effectively and maintains accurate predictions.

THE BENEFITS OF PINNS

A physics-informed neural network (PINN) leverages known physics principles to enhance neural network predictions. It's akin to instructing the neural network: 'Learn from the data, but be cautious. If your predictions contradict the physics model, those predictions shouldn't be considered valid.' This approach adds additional constraints to the neural network, particularly to its loss function, enabling the network to find solutions that align with both the data and established physical laws.

Additionally, expert physicists provide valuable insights, such as expected trajectories based on factors like drag coefficient and gravity. By integrating this expert knowledge with observational data, the model is better constrained and capable of making more accurate predictions. This is the core function of PINNs.

During inference, there may be instances where data is plentiful and easy to collect, followed by periods where data is sparse. Data availability can fluctuate, so combining the data-driven approach of neural networks with physics-based constraints ensures more reliable predictions. PINNs use the governing physics equations, such as partial differential equations, to guide their predictions when data is insufficient.

Physics serves as the foundational knowledge that PINNs must always consider. If a neural network's predictions suggest, for example, the absence of gravity in a scenario where gravity should be present, it indicates a problem with the model. Physics provides the essential prior knowledge that the neural network must consistently incorporate.

THE MATHEMATICS BEHIND PINNS

The mathematics underpinning physics-informed neural networks (PINNs) is intricate, involving calculus, partial differential equations (PDEs), ordinary differential equations (ODEs), and derivatives. These complex mathematical concepts are ubiquitous across various fields where predictions are crucial.

> *Data availability can fluctuate, so combining the data-driven approach of neural networks with physics-based constraints ensures more reliable predictions.*

[1] Runge-Kutta Methods en.wikipedia.org/wiki/Runge–Kutta_methods

In a purely data-driven approach, the loss minimisation process typically involves mean squared error (MSE). During backpropagation, the network's weights are adjusted to minimise the error between the network's predictions and the ground truth. The goal is to reduce the distance between the predicted and actual values, optimising the network's performance. This involves defining the MSE loss function and minimising it using techniques like stochastic gradient descent, a common practice in traditional neural networks.

With PINNs, an additional loss component, termed the 'physics loss,' is introduced. This component incorporates the physics model into the loss function, which now not only minimises data errors but also adheres to the constraints imposed by the physics model. The physics model often includes derivatives, which can be calculated using numerical tools available in frameworks like PyTorch or TensorFlow.

In PINNs, the loss function is augmented to account for the physics model. This expanded loss function aims to minimise discrepancies in both data-driven predictions and physics-based expectations. By combining and weighting these two loss components, the network can learn more effectively. The weighting factor or parameter is also learned by the network, ensuring an optimal balance between data and physics constraints.

One remarkable aspect of PINNs is their ability to learn unknown coefficients. For instance, in the case of the drag coefficient (often denoted as K), which acts in the opposite direction of the velocity, can vary based on atmospheric conditions like air density and the presence of clouds. Typically, this coefficient must be approximated or measured. However, PINNs can infer this value as a trainable parameter by analysing the data and the physics model, leading to an accurate approximation of the drag coefficient.

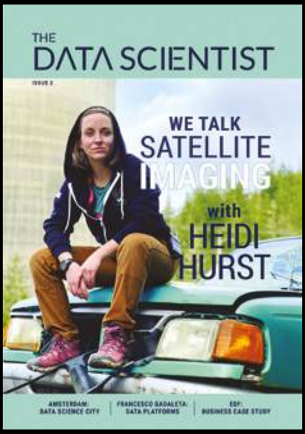
There may be other unknown coefficients within the physics model that PINNs can estimate during experiments. This capability of PINNs to infer and estimate unknown parameters is another instance where they can significantly enhance predictive accuracy and model reliability.

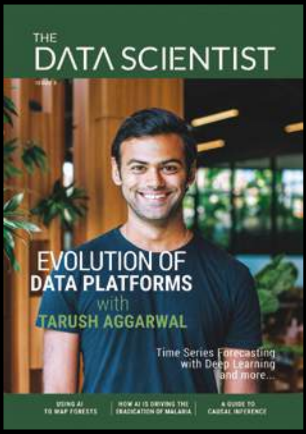
FINAL THOUGHTS

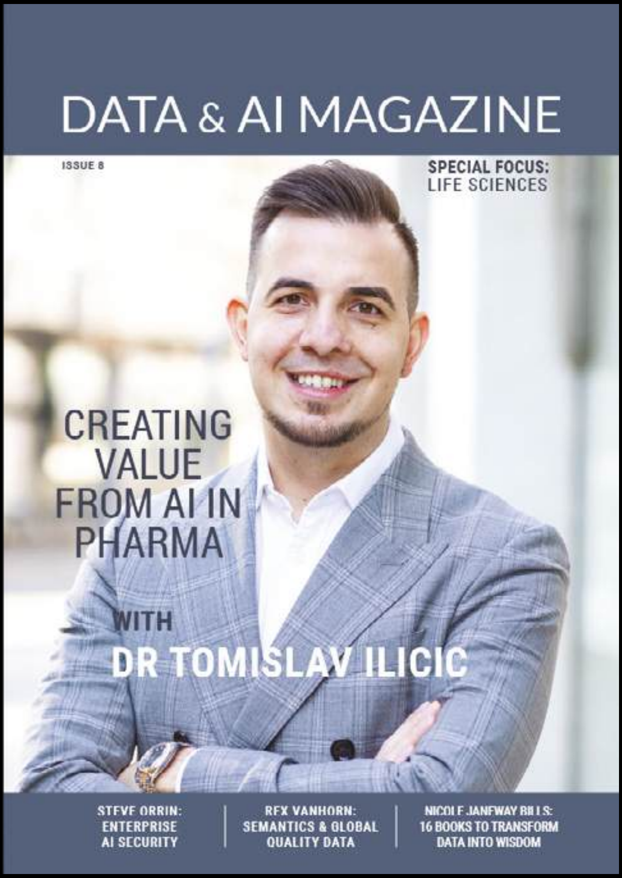
This is clearly a brief exploration into the world of Physics-Informed Neural Networks (PINNs). Amid the boundless hype and speculation surrounding neural networks, PINNs stand out for their potential to deliver real-world impact. Every mathematical method should be evaluated based on its true capabilities and potential. In the case of PINNs, we recognise neural networks as valuable and effective function approximators, highlighting their significance and utility.

DATA & AI MAGAZINE

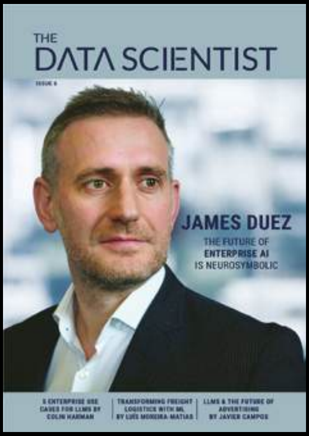
WOULD YOU LIKE TO BE FEATURED IN **DATA & AI MAGAZINE?**


THE DATA SCIENTIST
WE TALK SATELLITE IMAGING with HEIDI HURST
AMSTERDAM: DATA SCIENCE CITY
FRANCESCO GADALETA: DATA PLATFORMS
ESP: BUSINESS CASE STUDY

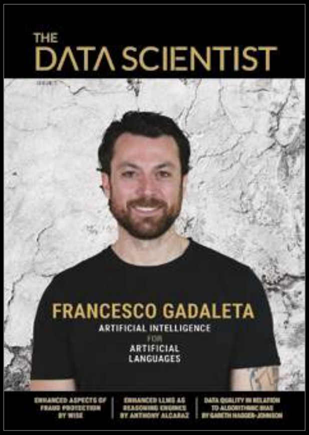

THE DATA SCIENTIST
EVOLUTION OF DATA PLATFORMS with TARUSH AGGARWAL
Time Series Forecasting with Deep Learning and more...
USING AI TO MAP FORESTS
HOW AI IS DRIVING THE ERADICATION OF MALARIA
A GUIDE TO CAUSAL INFERENCE


DATA & AI MAGAZINE
ISSUE 8
SPECIAL FOCUS: LIFE SCIENCES
CREATING VALUE FROM AI IN PHARMA
WITH DR TOMISLAV ILICIC
STEVE ORRIN: ENTERPRISE AI SECURITY
REX VANHORN: SEMANTICS & GLOBAL QUALITY DATA
NICOLE JANEWAY BILLS: 16 BOOKS TO TRANSFORM DATA INTO WISDOM

THE DATA SCIENTIST
ISSUE 5
ALEC SPROTEN & IRIS HOLLICK
ADVANCED DEMAND FORECASTING AT BREUNINGER
THE PATH TO RESPONSIBLE AI
ENTERPRISE DATA & LLM'S
HOW ML IS DRIVING PATIENT-CENTRED DRUG DISCOVERY


THE DATA SCIENTIST
JAMES DUEZ
THE FUTURE OF ENTERPRISE AI IS NEUROSYMBOLIC
5 ENTERPRISE USE CASES FOR LLMS BY COLIN HARMAN
TRANSFORMING FREIGHT LOGISTICS WITH ML BY LUIS MOREIRA-MATIAS
LLMS & THE FUTURE OF ADVERTISING BY JAVIER CAMPOS

AI SPECIAL ISSUE
THE DATA SCIENTIST
WALID MEHANNA
TRANSFORMING DATA CULTURE AT MERCK
FROM HUMANS TO HYBRIDS: PREPARING YOUR WORKFORCE FOR AI and more...
HARNESSING GEN AI TO SUPPORT NEURODIVERSITY
THE COST EFFECTIVE METHOD FOR RETRAINING LLM'S
CONTINENTAL'S APPROACH TO GENERATIVE AI


THE DATA SCIENTIST
FRANCESCO GADALETA
ARTIFICIAL INTELLIGENCE FOR ARTIFICIAL LANGUAGES
ENHANCED ASPECTS OF FRAUD PROTECTION BY WISE
ENHANCED LLMS AS REASONING ENGINES BY ANTHONY ALCARAZ
DATA QUALITY IN RELATION TO ALGORITHMIC BIAS BY GARETH HAGGER-JOHNSON

NEXT ISSUE

28th January 2025

To set up a 30-minute initial chat with our editor to talk about contributing a magazine article, please email imtiaz@datasciencetalent.co.uk