

Artificial Intelligence in Drug Development and the Impact on the Future Workforce

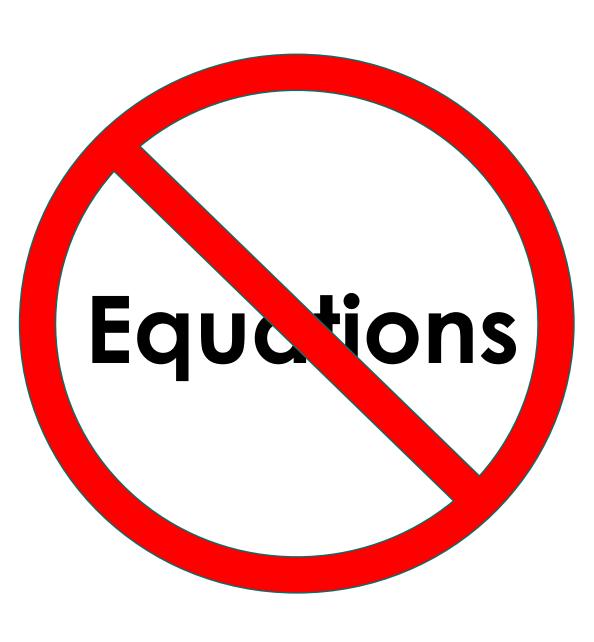
October 25, 2023

Brandon Higgs, PhD VP, Head of Translational Data Science



Background on AI

A brief primer





Introduction to AI

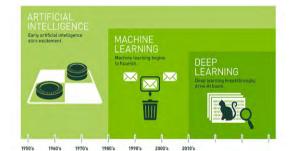
A brief background

- Definition of Artificial Intelligence (AI):
 - The simulation of human intelligence in machines that are programmed to think, learn, and perform tasks independently
 - Strive towards consciousness and sentience
- Origins of AI:
 - Dartmouth Conference in 1956, AI pioneers Alan Turing and John McCarthy

• Key Milestone:

Genmab

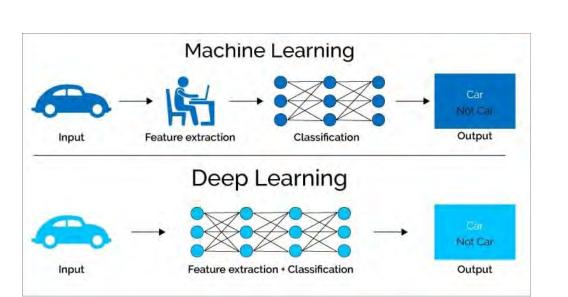
• IBM's Deep Blue defeating Garry Kasparov in chess, 1996-97







Machine learning vs. Deep learning



Related, but distinct disciplines

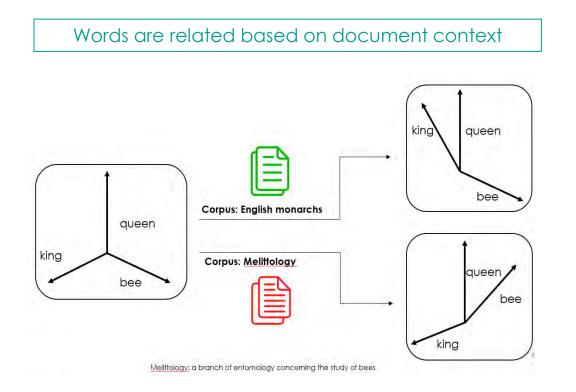




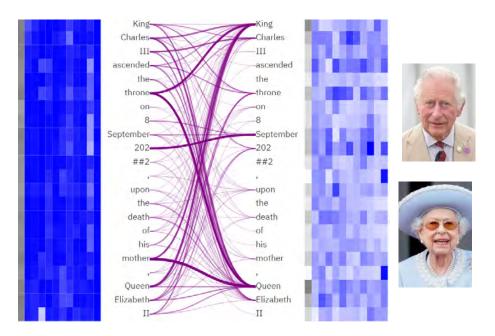
Large Language Models (LLMs): A very quick introduction



Sentence: "King Charles III ascended the throne on 8 September 2022, upon the death of his mother, Queen Elizabeth II"



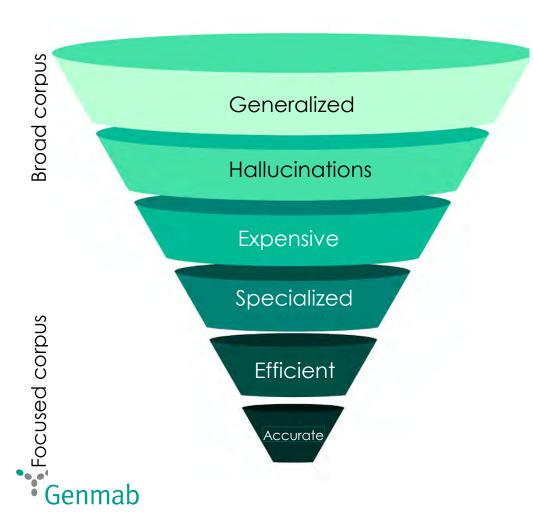
Words are related based on sentence context



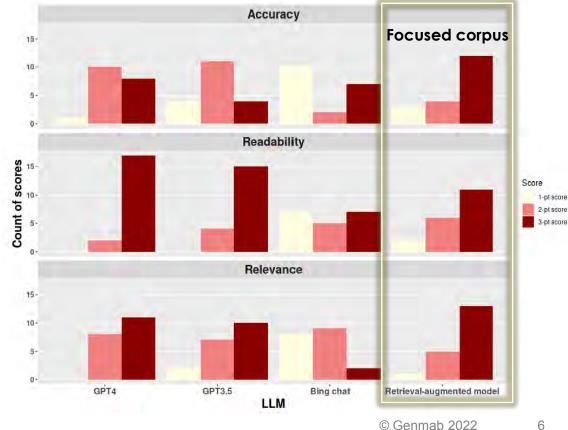
Genmab

Context matters for training or fine tuning LLMs

A focused corpus versus a broad corpus



Comparison between GPT3.5, GPT4, Bing, and a focused corpus



Soong et al, arXiv 2023, 2305.17116





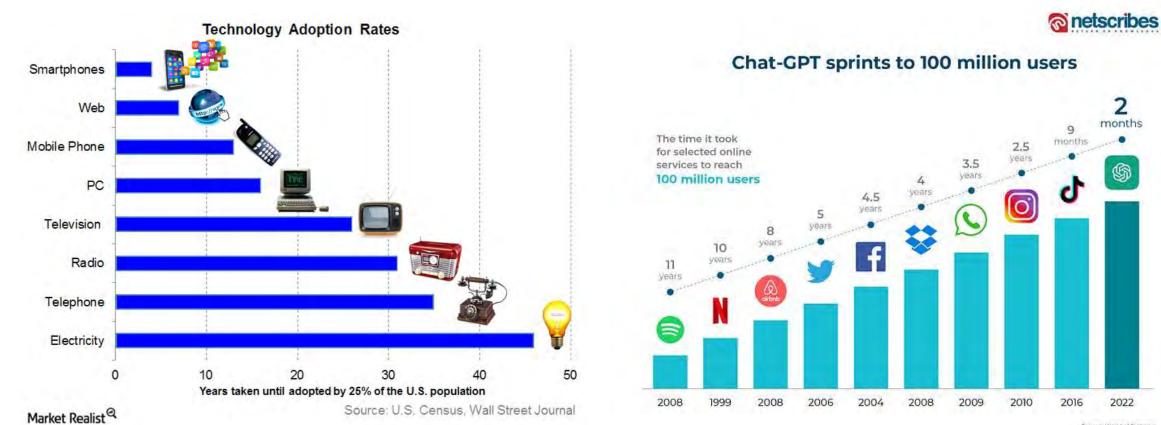
7

airquaity Ipioad a file or use demo nalyses in English or oth	data. Then jus	Browse No file	selected			i dozens of human languages. I am still in college learning new things. But I try to be helpful. I did finish thousands of books, millions of code repositories, and billions of web pages.
nalyses in English or oth	data Then Jus					
Apart of the output design (support bove, if unaccessful, iny) code works correctly som Submit penAl charges 2¢ per 10 lease use your own accord clava Ukraini!	cially the relev again with the e of the times. Settings 00 tokens/word	For general questions rant columns. See ex- same request or ask To use volce input, cl D Pythe Pythe	, briefly amples , anterently, lick Settings, 	Releases as a prototyp It can take a few tries tr seeking alternative solu Prepare and clean your Once uploaded, your di especially when numbe Before sending your rei afterward, If you are no Your data is not sent to emailing a statistician w	e, it is still being developed. Please a get it correct. If it still does not work ditions. It data in Excel first! RTutor can only ata is automatically loaded into RTut ers are used to code for categories. Quest to OpenAI, we add "Generate it using any data (plot a function or s the AI. To ask generic questions with who knows nothing about your data.	ests such as cistributions, basic plots, or simple models. Then customize it or add variables, send us your valuable feedback (lower left). k, rephrase your request. Also, increase the "Temperature" setting will make the Al more aggressive in analyze traditional statistics data, where rows are observations and columns are variables. tor as a data frame called df. Check if the data types of the columns are correct, Change if needed, Data types make a big difference in analysis and plots! R code" before it, and append something like "Use teff data frame. Note that hwy is numeric" aimulations), choose "No data" from the Data dropdown, thout mentioning column names, briefly describe your data, especially the relevant columns, just like ted data. If you want to build upon the current code, select the "Continue from this chunk" checkbox.
) Comments & questions					II be inserted before your next chunk an go back to any previous chunks a	k and get executed. This is especially important for data wrangling when you remove rows, add columns and continue from there.
				Data Types	Description	
				\$ Solar.R: int 190 118 \$ Wind : num 7.4 8 1 \$ Tomp : int 67 72 7	2 18 NA 28 23 19 8 NA 149 313 NA NA 299 99 19 194 2.6 11.5 14.3 14.9 8.6 13.8 20.1 4 62 56 66 65 59 64 69 levels "5","6","7","8",: 1 1	1 8.6





Rates of technology adoption



Source: World of Stability

Benefits of AI in the workforce



- Increased Productivity and Efficiency
 - Automation of routine tasks, optimization of processes, and handling vast amounts of data, leading to enhanced productivity and efficiency

- $\langle \gamma \rangle$
- Enhanced Decision-making
 - Analyze complex data, identify patterns, and generate insights, enabling better decisionmaking and strategic planning



enmah

- Automation of Repetitive Tasks
 - Automate mundane and repetitive tasks, freeing up human resources for more creative and strategic endeavors



Al is reshaping specific industries

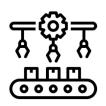
Not exhaustive



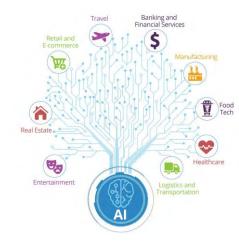
- Finance and Banking
 - Fraud detection, algorithmic trading, and personalized financial recommendations



- Healthcare
 - Medical imaging analysis, drug discovery, and patient risk prediction for better healthcare outcomes



- Manufacturing
 - Predictive maintenance, quality control, and autonomous robotics to optimize production processes
- Transportation
 - Autonomous vehicles, traffic management systems, and route optimization to improve safety and efficiency



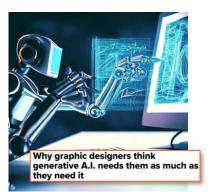
Professions impacted by advances in generative AI

Automating processes and collaborating with humans to solve problems



Copy editing

Generate draft content, allowing copy editors to focus on polishing, fact-checking, and editing instead of writing from scratch



Graphic design

Design mockups, logos, and marketing materials, enhancing designers' workflows



Customer service

Chatbots for simple customer inquiries, allowing agents to handle more complex issues



Medical imaging

Automate annotation of scans to spot anomalies, acting as a diagnostic aid for radiologists



Software engineer

Prompting to accelerate coding, allowing engineers to focus on more complex problems

https://scieditor.ca/2023/01/edito vs-aiis-editing-a-dead-career/ https://www.cnbc.com/2023/08/05/w hy-graphic-designers-think-gen-aineeds-them-as-much-as-they-needit.html

https://www.bcg.com/publications/2 023/how-generative-ai-transformscustomer-service https://healthitanalytics.com/news/researchers -leverage-aito-detect-causes-of-alzheimersdisease#:~:text=September%2021%2C%202022 %20%20Published%20in, early%20indicators% 200f%20Alzheimer's%20Disease.

https://www.forbes.com/sites/jodiecook/2023/07/1 2/ai-prompt-engineers-earn-300k-salaries-hereshow-to-learn-the-skill-for-free/?sh=65b40129d4a1

The promise of AI in drug discovery

Speed and efficiency

Accelerate with more efficient ways of identifying the right targets. Improved decision making

2

Enabling researchers to make better and more informed decisions at every stage of the process. Cost savings

3

Reduce timeconsuming experiments and high failure rates and increase the likelihood of success

Transformative potential

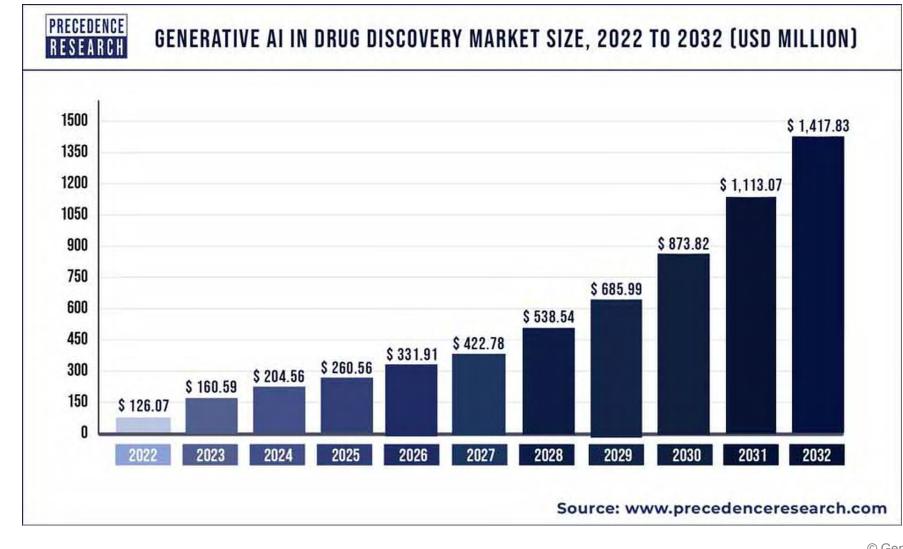
4

Investors, scientists, and regulators are interested in AI, as it is poised to revolutionize the future of drug development





Al in drug discovery has enormous profit potential

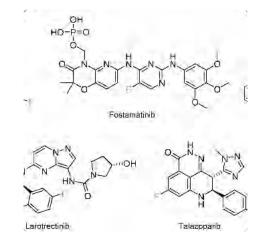


••••

Genmab

Some key components of drug discovery





Understanding Diseases and Selecting Therapies

Al is helping scientists analyze vast amounts of data to better understand diseases and the best therapies for them. Identifying and Validating Targets

ML methods can learn from existing data and predict which protein targets might be most effective in treating a given condition. Designing Therapeutic Modalities

Al algorithms can help in designing therapeutic modalities, such as small molecules and biologics, that have the desired biological properties.

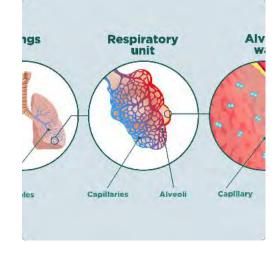


Al-designed compounds in clinical trials



Exscientia

The first AI-designed drug for the treatment of obsessive-compulsive disorder (OCD); currently in Phase I clinical trials.



Insilico Medicine

Al-designed drug for idiopathic pulmonary fibrosis (IPF), in Phase II clinical trials.



Relay Therapeutics

Drug candidates for the treatment of cancer and other diseases, including inhibitors for SHP2 and FGFR protein; advancing in clinical trials.



HR: Unlocking the potential of AI





Job descriptions (JDs): Provide a previous JD or a set of requirements and let generative AI compose the rest

Resumes: Summarize qualifications in few words, prioritize with weighting metrics for experience and skillsets across candidates



8 8 8 Performance

Chatbots can provide 24/7 support, answering common questions and guiding employees through the onboarding process

Al-based recommendation systems can offer personalized training programs based on an employee's role, skills, and career goals Al-driven performance metrics, feedback, and employee data to identify patterns and trends

Al mapping tools to identify skill gaps and recommend targeted training programs to enhance employee performance



Engagement & Retention

Al-powered sentiment analysis can analyze employee feedback, surveys, and social media data to measure employee satisfaction



What are some of the ways we are leveraging AI at Genmab?





Al applications at Genmab



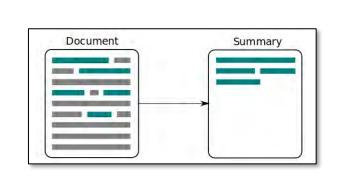
Virtual control arms

Use real-world data (RWD) to simulate clinical outcomes in comparator arm; ultimately de-risking pivotal clinical trials

Patient biopsy Features extracted from pathology image Features extracted from Feat

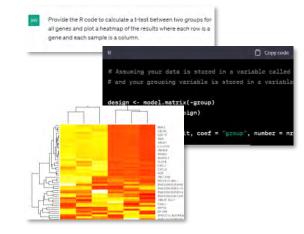
Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



Summarization

Automated summarization and simplification of scientific conference content



Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics

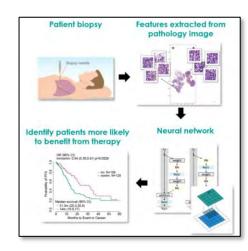


Al applications at Genmab



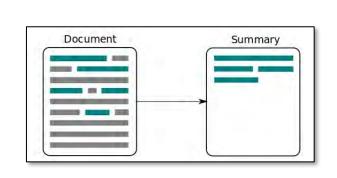
Virtual control arms

Use real-world data (RWD) to simulate clinical outcomes in comparator arm; ultimately de-risking pivotal clinical trials



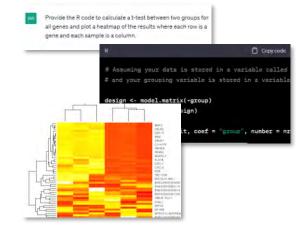
Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



Summarization

Automated summarization and simplification of scientific conference content



Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics



The balance of virtual control arms



- Large real-world data sets provide broader, more generalizable patient populations
 - Eliminates cost of patient recruitment
 - Reduces duration of trial

Genmab

- Contemporary data reflects current standards of care
- Matching methods balance baseline characteristics and reduce bias
- Can answer research questions not feasible
 in RCTs

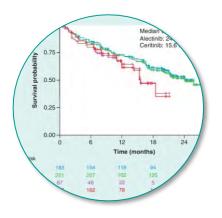


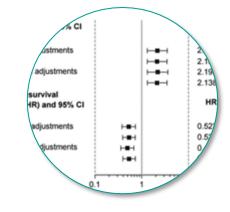
- Residual confounding from unmeasured factors
- Data quality limitations (e.g. missing data, variability in collection methods)
- Endpoint differences (RECIST vs clinical documentation of progression)
- Generalizability to clinical trial population
- Acceptance of methodology by stakeholders



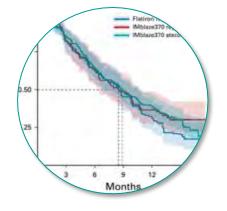
Examples in oncology trials

Demonstrate improved efficacy or noninferiority over standard of care

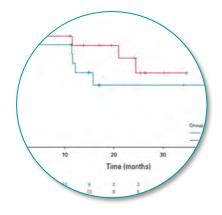




- Alectinib in ALK+ NSCLC (n=138) vs RWD control (n=306): OS HR 0.65
- Blinatumomab (CD19xCD3) in relapsed/refractory ALL (n=134) vs RWD control (n=478): CR 40% vs 25%



 Atezolizumab + Cobimetinib (MEKi) in mCRC (n=183) vs RWD controls (n=173-179): OS HR 0.67-0.86



 GEN-1 (IL-12) + chemo in ovarian cancer (n=18/15) vs RWD control (n=41/15): PFS HR 0.53

Davies et al., 2018

Gökbuget et al., Blood Cancer Journal, 2016 Schroder et al., JCO Clinical Cancer Informatics, 2021

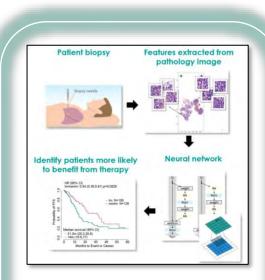
Yin et al., JCO Clinical Cancer Informatics, 2023

Al applications at Genmab



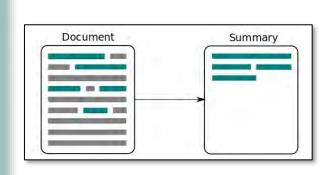
Virtual control arms

Use real-world data (RWD) to simulate clinical outcomes in comparator arm; ultimately de-risking pivotal clinical trials



Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



Summarization

Automated summarization and simplification of scientific conference content



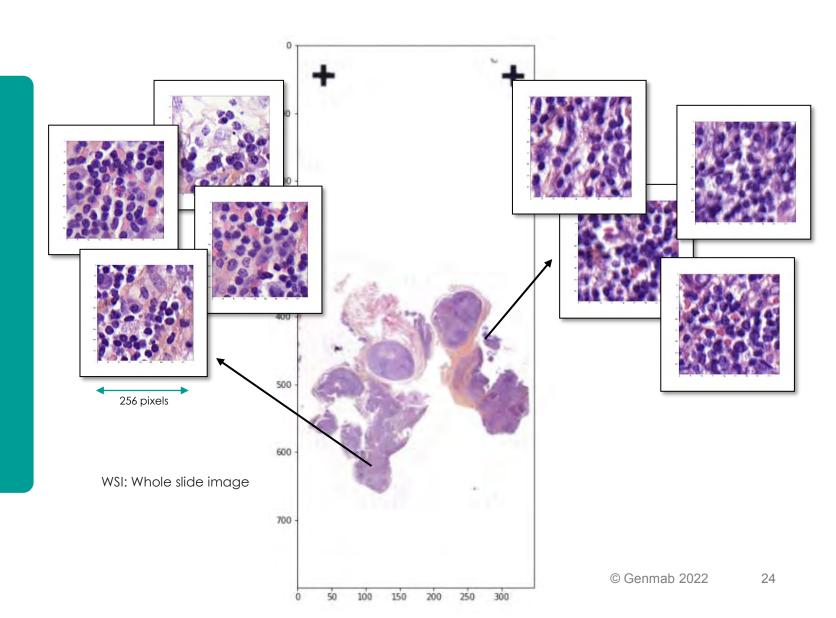
Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics



Tile/Patch extraction from whole slide image

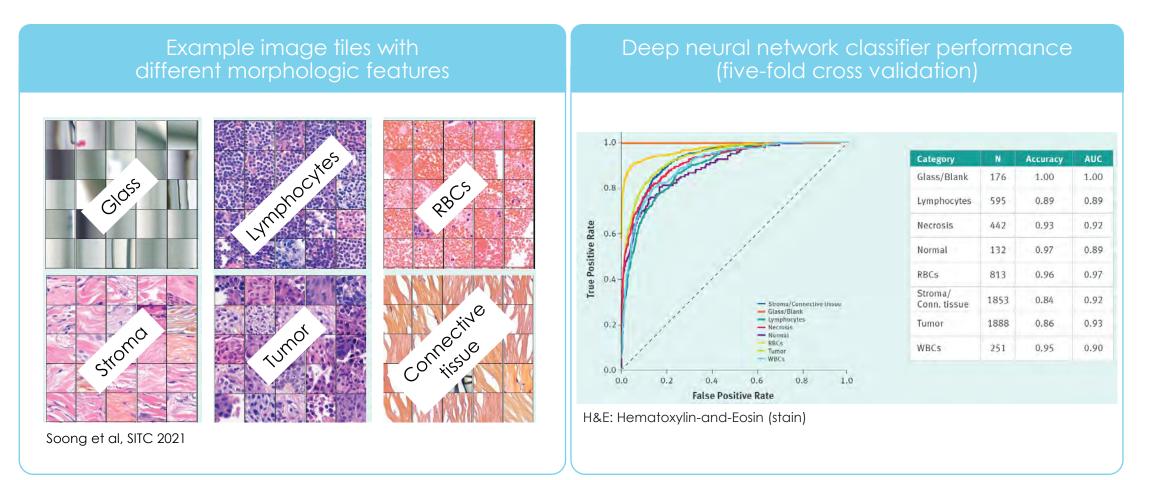
- Tissue contexture is heterogeneous
 and varies by region
 - WSI is prohibitively large search space
- Can divide the WSI into thousands of small image tile representations
- Use these tiles as microrepresentations of the tissue
- Size of WSI versus tile
 - Single WSI: 1-3 GB
 - Tile size:
 [(256x256 pixels) x 24 bits]/8 =
 197KB





Automated histologic, morphologic, and artifact detection

• ResNet-50 CNN used to learn pathologists' annotations



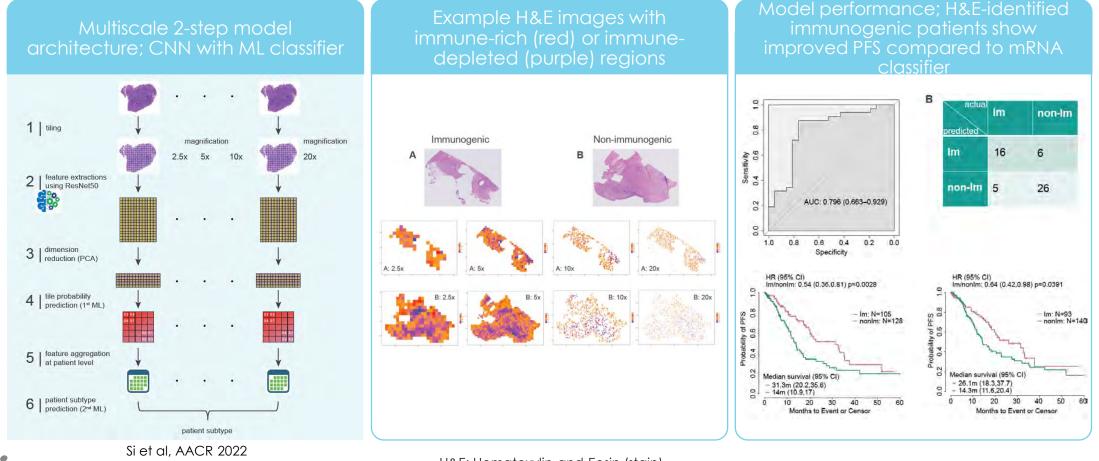
Genmab

Image-based prediction of immunogenic subtype in pancreatic cancer

- Identifying patients with immunogenic phenotype using whole slide H&E images
- ResNet-50 CNN architecture implemented

Genmab

ŏ.

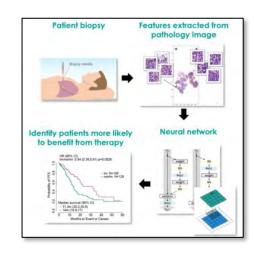


Al applications at Genmab



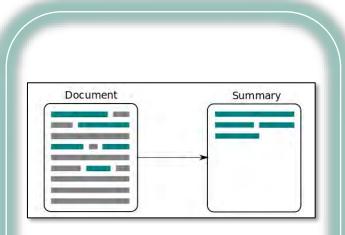
Virtual control arms

Use real-world data (RWD) to simulate clinical outcomes in comparator arm; ultimately de-risking pivotal clinical trials



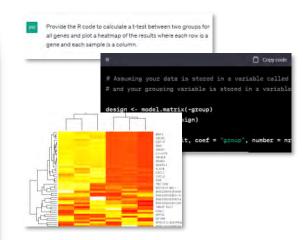
Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



Summarization

Automated summarization and simplification of scientific conference content

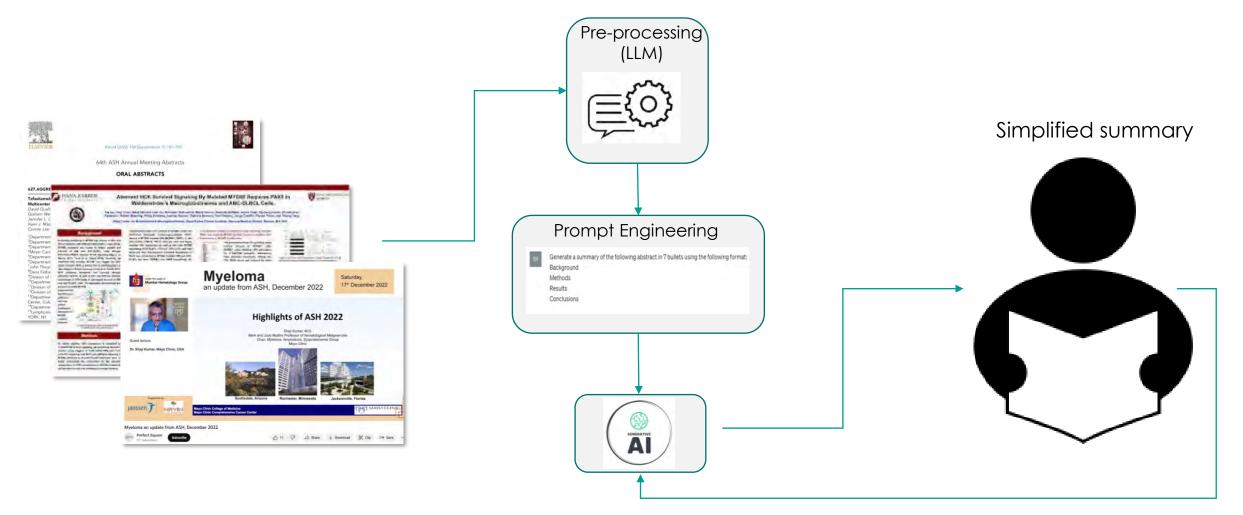


Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics



Al-generated conference content summaries



Fine tuning



Technical challenges of AI summarization

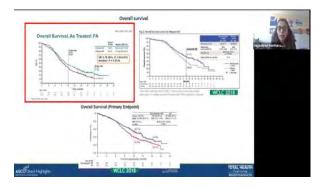


Image quality & resolution

OCR may struggle with low resolution, poor image quality, or too many small images on a slide, making it difficult to recognize characters and numbers

accurately

Genmab

Ò



Graph interpretation

Interpreting and extracting statistics or graphical elements from KM graphs, bar graphs, or pie charts can require more advanced capabilities

	N	A	В	P-value
		N = 81	N = 169	
age	250	60 65 70	61 67 73	0.1041
Sex	250			0.453^{2}
male		78% (63)	73% (124)	0) ++
Race	250			0.189^{2}
Black		0% (0)	3% (5)	045
Caucasian		98% (79)	96% (163)	
Oriental		1% (1)	0% (0)	5
Other		1% (1)	1% (1)	¥
height	250	164 169 175	164 168 173	0.185^{1}
weight	250	65 75 89	63 74 85	0.516^{1}
bmi	250	23.3 25.4 29.6	22.6 25.8 29.4	0.9961
smoking	250	37% (30)	35% (59)	0.743^{2}
pack_yrs	250	30 48 60	30 42 60	0.862^{1}

a b c represent the lower quartile a, the median b, and the upper quartile c for continuous variables. N is the number of non-missing values. Numbers after percents are frequencies. Tests used: ¹Wilcoxon test; ²Pearson test

Complex table structures

Merged cells, missing values, indentation, footnotes, nested sub-tables, or varying row or column sizes, make the OCR extraction process more difficult

Al applications at Genmab



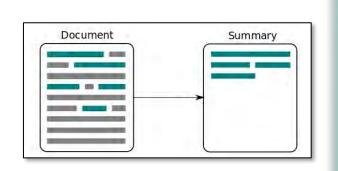
Virtual control arms

Use real-world data (RWD) to simulate clinical outcomes in comparator arm; ultimately de-risking pivotal clinical trials

Patient biopsy Features extracted from pathology image Features extracted from Features extracted

Digital pathology

From patient biopsy to histology slide, to predicting likelihood of responding to a therapy



Summarization

Automated summarization and simplification of scientific conference content



Prompt-driven analysis

Biomarker and clinical correlative analysis using generative AI to derive code and graphics



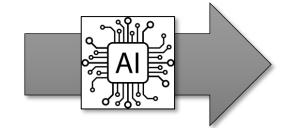
Writing prompts to create programming code and graphics

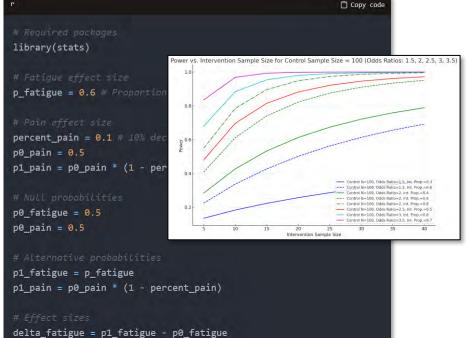
BH

can you provide R code to calculate this using the method in the uploaded paper?

Sure, here is some R code to calculate the sample size for the example I walked through, using the approach from the paper you provided:

Article	SMMR
Decision-making with multiple correlated binary outcomes in clinical trials	Standical Hetheds in Phridal Newards 2020 Vol 39(11) 326.3277 The Advisuity 3020 COCC Anada estate gatellanes tagged convolution and accession tagged convolution and accession portfol seguida convolutioned sensi- port as seguida convolutioned sensi- port as seguida convolutioned sensi- port as seguida convolutioned sensi- SISAGE
Xynthia Kavelaars ¹ , Joris Mulder ^{1,2} and Maurits Ka	ptein ² 🤒
Abstract Clinical trials often evaluate multiple outcome variables to form a con- treatment. The resulting multidimensional insight contributes to chincall treatment superiority. Common statistical procedures to make these sup- tive important shortcomings, llowever: (1) Outcome variables are often consider the relation between outcomes; and (2) superiority is often defi- or on all outcome(s), and tacks a compensatory mechanism that allows tar- (s) to outweegl small registive effects on other outcomes. To address t- Bayesian model for the analysis of correlated binary outcomes based on to lexible decision criterion with a compensatory mechanism that capture simulation study demonstrates that efficient and onbiased decisions can- controlled. The performance of the framework is illustrated for (1) fixed, ion-informative and informative prior distributions.	y relevant and efficient decision-making about eriority decisions with multiple outcomes have innedled individually, and consequently fail too ge positive effects on one or multiple outcome have shortcomings, this paper proposes (1) a the multivariate Bernoulli distribution; and (2) a es the relative importance of the outcomes. A be made while Type Lerror rates are properly
eywords ultude outcomes, compensatory decision rules, multivariate Bernoulli	model, efficiency, Bayesian analysis





delta_fatigue = p1_fatigue - p0_fatigu delta_pain = p1_pain - p0_pain

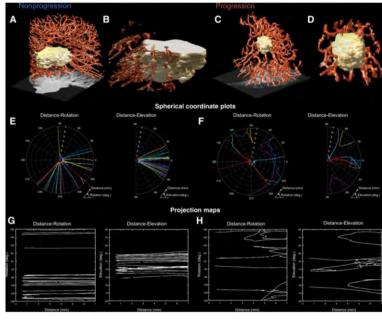


Radiomics as a noninvasive biomarker of target expression, safety, and clinical outcome

Vasculature morphology from imaging as a biomarker of 1L chemo response

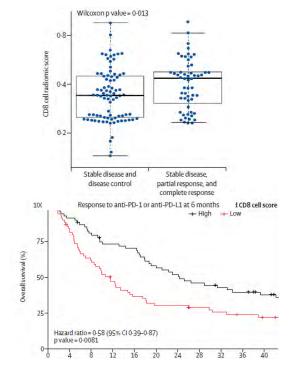
Radiomic signature correlates with CD8 infiltrate and response to aPD(L)1

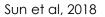
Integrated molecular data with radiomic features predicts survival in NSCLC

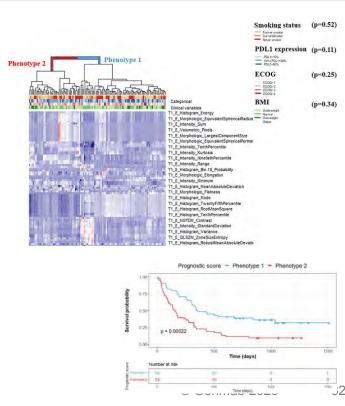




Braman et al, 2022



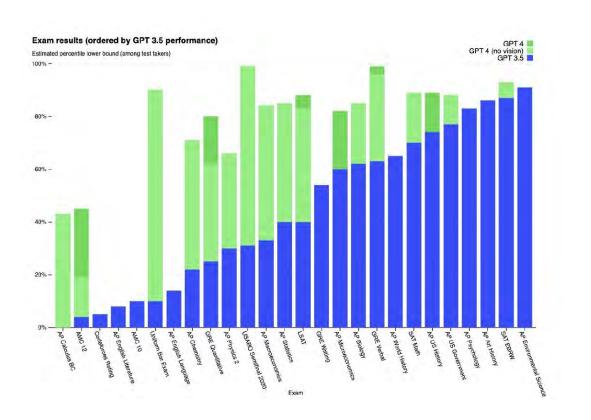




Singh et al, 2022

What skills are most affected or unaffected by generative AI

Exam performances by GPT3.5/4



Skills importance for GPT

	Basic Skill	α (std err)	β (std err)	۲ (std err)				
	All skill importance scores are normalized to be between 0 and 1.							
	Constant	0.082*** (0.011)	-0.112*** (0.011)	0.300*** (0.057)				
	Active Listening	0.128** (0.047)	0.214*** (0.043)	0.449*** (0.027)				
	Mathematics	-0.127*** (0.026)	0.161*** (0.021)	0.787*** (0.049)				
	Reading Comprehension	0.153*** (0.041)	0.470*** (0.037)	-0.346*** (0.017)				
Г	- Science	-0.114*** (0.014)	-0.230*** (0.012)	-0.346*** (0.017)				
GPT ags	Speaking	-0.028 (0.039)	0.133*** (0.033)	0.294*** (0.042)				
	Writing	0.368*** (0.042)	0.467*** (0.037)	0.566*** (0.047)				
	Active Learning	-0.157*** (0.027)	-0.065** (0.024)	0.028 (0.032)				
L	- Critical Thinking	-0.264*** (0.036)	-0.196*** (0.033)	-0.129** (0.042)	GPT			
	Learning Strategies	-0.072* (0.028)	-0.209*** (0.025)	-0.346*** (0.034)	excel			
	Monitoring	-0.067** (0.023)	-0.149*** 0.020)	-0.232*** (0.026)				
	Programming	0.637***	0.623*** (0.022)	0.609*** - (0.024)				

© Genmab 2022 33

FDA guidance regarding appropriate use of AI/ML in drug development

- Efforts by FDA, NIST, ISO, and others to develop principles and standards for trustworthy and ethical AI use to inform drug development
- Transparency, human oversight, and model validation in the drug development context
- Engage stakeholders on considerations for AI/ML in drug development through a public workshop and other collaborative opportunities





Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products

Discussion Paper and Request for Feedback



https://www.fda.gov/media/167973/download



General AI considerations in the popular press

Ensure responsible AI deployment

Accuracy —	Bias —	Privacy —
THE WALL STREET JOURNAL LIFE& WORK IDEAS WORD ON THE STREET 'Hallucination': When Chatbots (and People) See What Isn't There A Latin term for mental wandering was applied to the disorienting effects of psychological disorders and drug use—and then to the misfires of AI programs	ChatGPT and large language model bias	Don't tell anything to a chatbot you want to keep private
When A.I. Chatbots Hallucinate	Bloomberg • Live Now Merkets Economica Industries Technology Politics Waliffit Pursuits Opinion Businessweek E Newsletter OpenAl Chatbot Spits Out Biased Musings, Despite Guardrails	6 Tips to ensure that your chatbots are GDPR compliant
Job Displacement & Reskilling	Accountability & Responsibility	Transparency & Explainability
What To Do About Al-Driven Job Displacement	NEWS 10 Annuary 2023 ChatGPT listed as author on research papers: many scientists disapprove At least four articles credit the AI tool as a co-author, as publishers scramble to regulate its use. Technology 4 refunde read: April 5, 2023 8:22 AM EDT - Last Updated a membrage Elon Musk and others urge AI pause, citing 'risks to society'	Technology And Analytics We Need Al That Is Explainable, Auditable, and Transparent
• Genmab	citing has to society	© Genmab 2022 35

General AI considerations in the popular press

Ensure responsible AI deployment

Accuracy	Bias	Privacy —
THE WALL STREET JOURNAL. IFFE WORK I NORED ON THE STREET 'Hallucination': When Chatbots (and People) See What Isn't There A Latin term for mental wandering was applied to the disorienting effects of psychological disorders and drug use—and then to the misfires of Al programs	ChatGPT and large language model bias	Don't tell anything to a chatbot you want to keep private
When A.I. Chatbots Hallucinate	Bloomberg • Live Now Markets Economica Wedentries Technology Politics Wedentries Opmon Businessweek E Newsletter OpenAl Chatbot Spits Out Biased Musings, Despite Guardrails	6 Tips to ensure that your chatbots are GDPR compliant
Job Displacement & Reskilling	Accountability & Responsibility	Transparency &
What To Do About Al-Driven Job Displacement	NEWS We Manuany 2023 ChatGPT listed as author on research papers: many scientists disapprove At least four articles credit the Al tool as a co-author, as publishers scramble to regulate its use. Technology 4 minute read - April 5, 2023 B-22 AM EDT - Last Updated a membrage Elon Musk and others urge AI pause, citing 'risks to society'	Technology And Analytics We Need Al That Is Explainable, Auditable, and Transparent
Gennian	ching hold to society	© Genmab 2022 36

Example of gender bias in LLMs

Inherent weighting given to specific genders for certain professions

ver: 0 ~ Attention:	All ~	Layer: 0 ~ Attention:	All
The	The	The	The
Doctor	Doctor	Doctor	Doctor
asked	asked	asked	asked
the	the	the	the
Nurse	Nurse	Nurse	Nurse
a	a	a	a
question	question	question	question
. //		. /	
She	She	He	He
said	said	asked	asked
nmab			© Genmab 2022

https://generativeai.pub/explainable-ai-visualizing-attention-in-transformers-4eb931a2c0f8

Concluding remarks

Balanced Approach: Integration with a balanced perspective, considering both potential and limitations

Ethical Considerations: Must be proactively addressed to maintain public trust and ensure equitable outcomes in drua development

Workforce Adaptability: Pharma's workforce needs to be agile and adaptable, acquiring new skill sets to work alongside AI technologies effectively

Regulatory Framework: A growing need for robust regulatory frameworks to ensure that AI applications in drug development meet safety and efficacy standards

Future Preparedness: Improved collaboration to prepare the next generation of scientists, researchers, and professionals for a future where AI will be an integral part of drug development

GUNIDUIS

- postdoctoral researchers polled said that .ce (AI) chatbots, such as ChatGPT, for everything rixing code and overcoming writer's block.

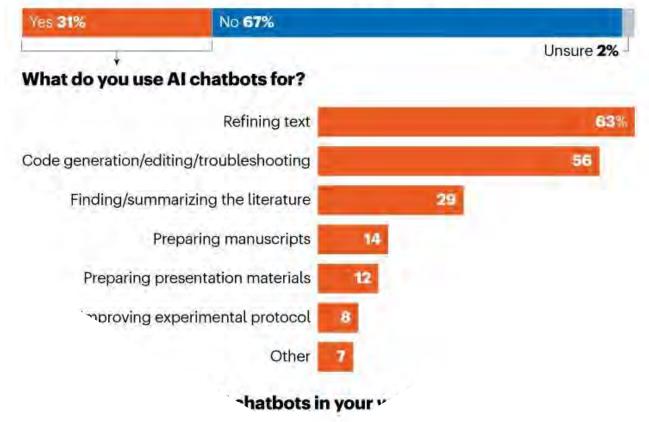
evolution and rapid adoption of AI chatbots d your day-to-day work or career plans, if at all?





31% Changed how I write papers 22% Changed how I analyse data 17% Changed how I stay up to date with the literature

Do you use AI chatbots, such as ChatGPT, in your work?







Acknowledgments

It takes a village...

- Genmab Translational & Quantitative Sciences
 - Han Si
 - David Soong
 - Sri Sridhar
 - Anantharaman Muthuswamy
 - Clifton Drew
 - Mohammed Qutaish
 - Mark Fereshteh
 - Manish Gupta
- Genmab Data Sciences & Al
 - Hisham Hamadeh
 - Sam Wagner
 - Scott Ogden
 - Kevin Tsăi
 - Elmutaz Shaikho
 - Hareesh Chandrupatla
 - Sanyam Kumar
- Genmab Discovery Data Science
 - Chris Klijn
 - Roberto Spreafico

- Rutgers University
 - Janet Alder
- Neomorph
 - Suzana Couto

