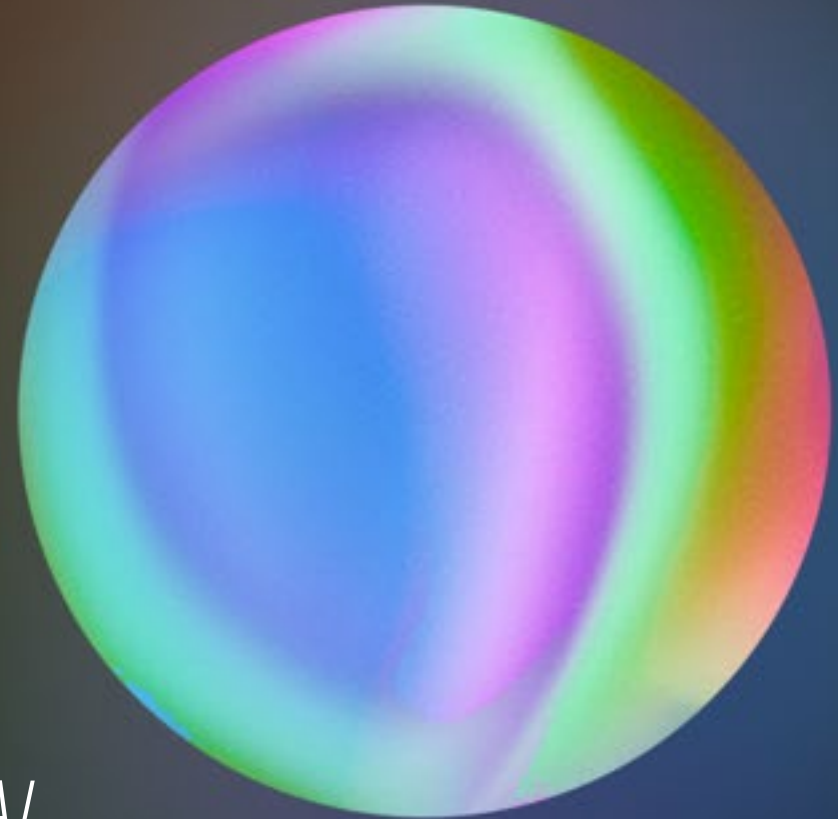


AI for Forensics & Legal

Navigating the Path to Trustworthy AI



May 2024

Agenda

1. Introduction
2. What is AI?
3. Why should I care?
4. AI in Forensics & Legal
5. Trustworthy AI
6. Q&A

Introduction



Meet the Presenters



Amy Dove

Partner

Forensic Discovery & Analytics



Maggie Zheng

Associate Director

Forensic & Risk Analytics

How often do you use ChatGPT?

Every day

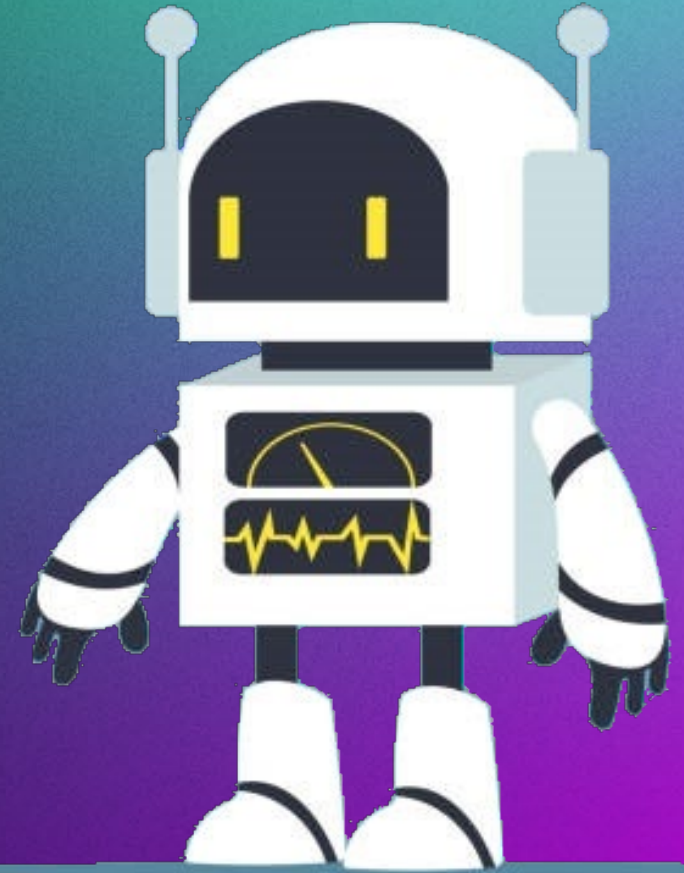
Once in a while

I've never used it

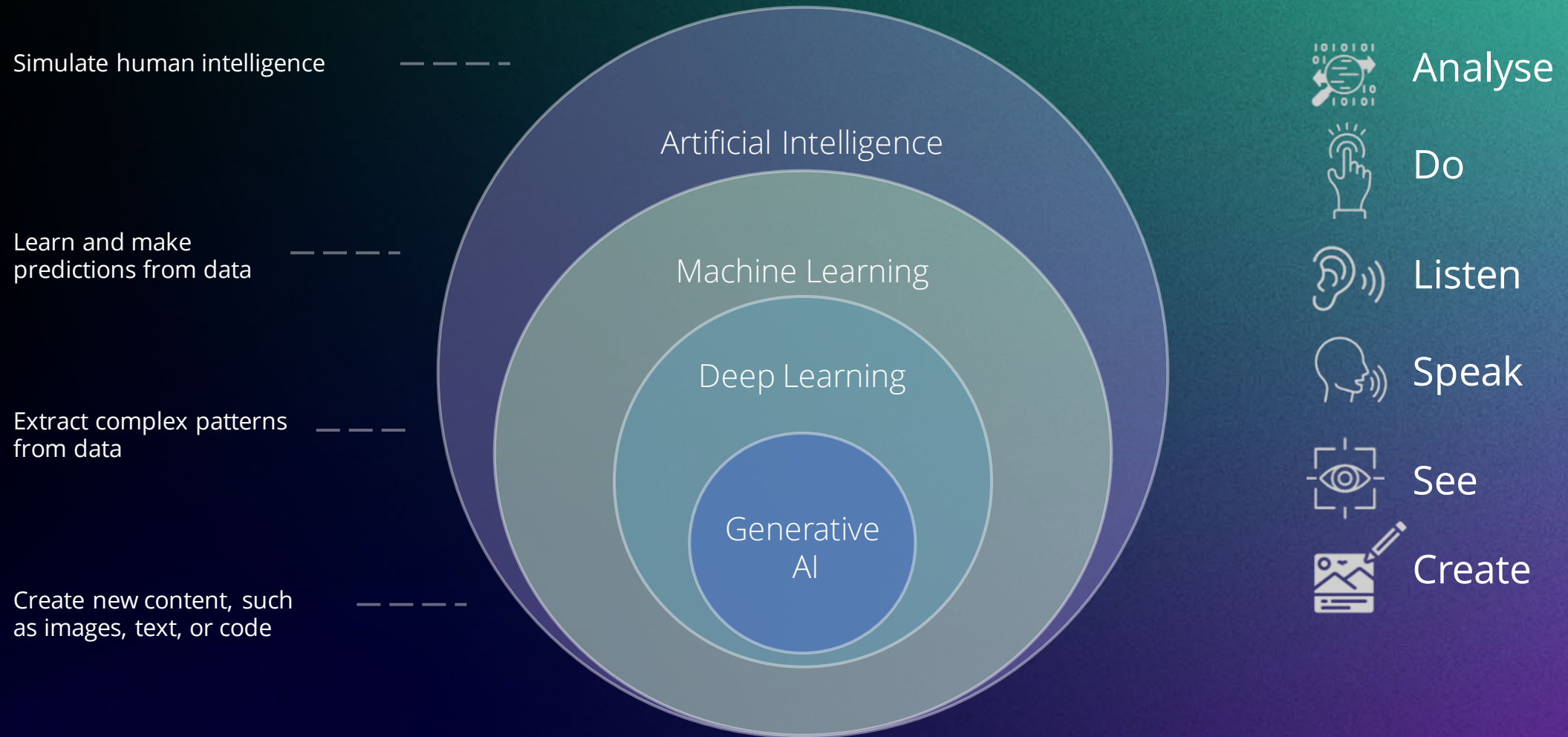
What is AI?



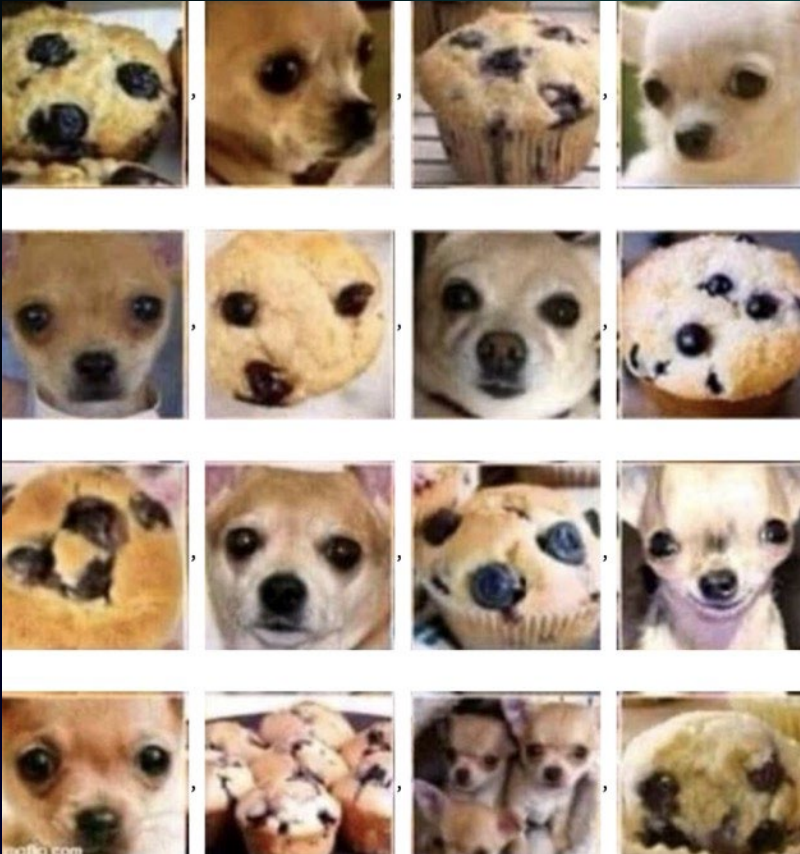
The attempt to get computers to
do things that make us say
'Gosh that was intelligent'



What is it?



From prediction



To creation



What is Generative AI?



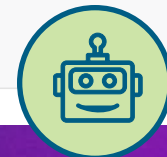
Define generative ai for me as though I am a five year old child



Generative AI is like having a magical computer friend who can make things up all by itself!



Just like when you play pretend and make up stories in your head, generative AI can create new things like pictures, music, or even stories all on its own, without needing someone to tell it what to do. It uses special rules and patterns to come up with these new things, and it can even learn from the things it creates to make even better things in the future.



How Generative AI works

The AI is trained on a Large Language Model

AI are trained on massive amounts of data to understand human communication and natural language.



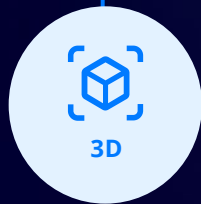
The user submits a prompt

A prompt is text submitted as part of a chat, an image, a video or a design etc.



AI generates new data

AI algorithms analyse the prompt and, in response, use prediction to generate new content across **different modalities**.



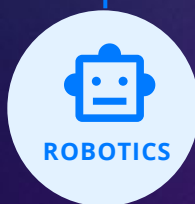
3D



CODE



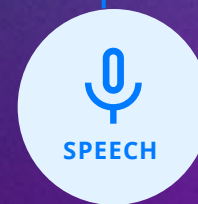
IMAGE



ROBOTICS



SPECIALISED



SPEECH



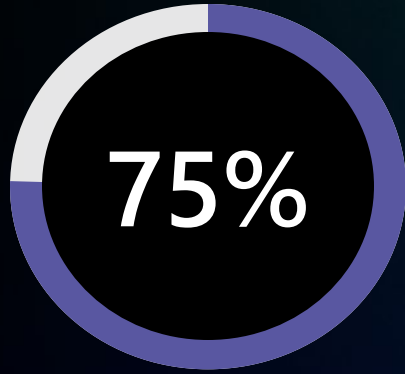
TEXT



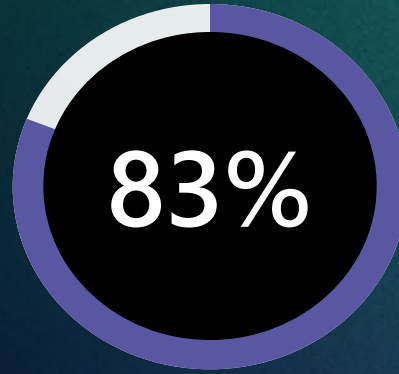
VIDEO

AI – Why should I care?

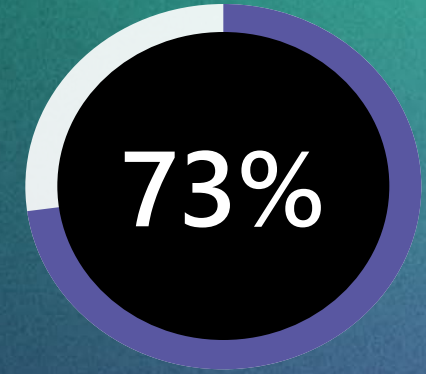
It is changing the way we work



General counsel expect to **increase their use of AI** and automation in the coming year¹



Organizations that expect to implement **Generative AI over the next two years** for their anti-fraud programs²



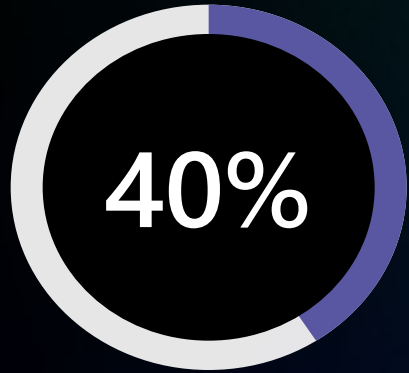
Expect to use **Large Language Models/Generative AI in 2024** to create new workflows in eDiscovery³

¹ Based on *The General Counsel Report 2024* global survey of Chief Legal Officers by FTI Consulting and Relativity

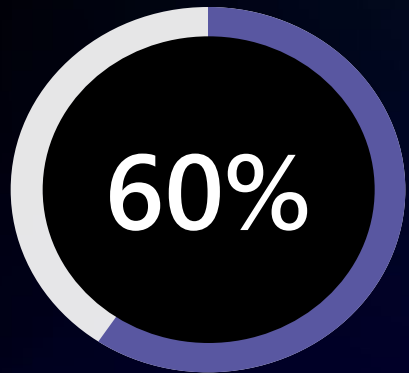
² Based on *Association of Certified Fraud Examiners 2024 Anti-Fraud Technology Benchmarking Report* – survey of 80,426 ACFE members

³ Based on *eDiscovery Today 2024 State of the Industry Report* – survey of 444 eDiscovery professionals

...but that's not the only reason



Improvement in the productivity of highly skilled workers – right now.¹



Activities of knowledge workers will be automated by 2030.²

1: [How generative AI can boost highly skilled workers' productivity | MIT Sloan](#)

2: [Economic potential of generative AI | McKinsey](#)



AP photo
Elementary school teachers picket against use of calculators in grade school
The teachers feel if students use calculators too early, they won't learn math concepts

Math teachers protest against calculator use

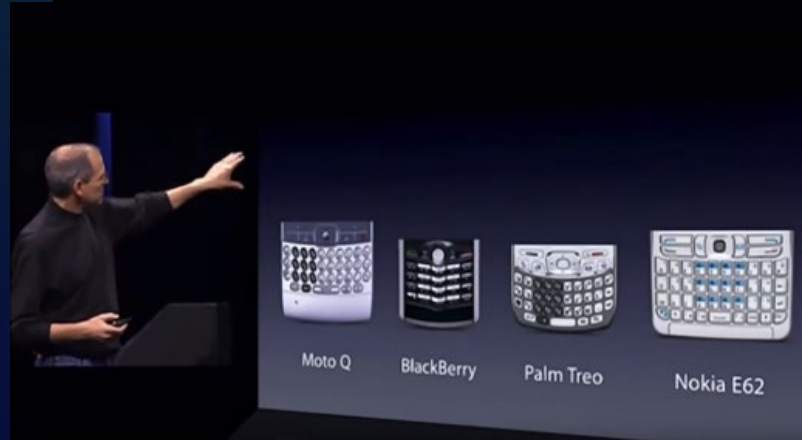
Internet 'may be just a passing fad as millions give up on it'

THE Internet may be only a passing fad for many users, according to a report. Researchers found that millions were turning their backs on the

By James Chapman
Science Correspondent

Woodgate, director of the society, said: "We are often presented with a mix-

Net loss: Two million Britons have logged off the Internet



"There is no chance of the iPhone ever gaining significant market share. no chance". Steve Ballmer, CEO of Microsoft, 2007.

“AI won’t replace you –
but someone that uses AI will”

The barrier to entry has never been lower.
The bar for differentiation has never been higher.

So what do we do about it?

**UNLOCKING
PRODUCTIVITY**

**Do Things
Differently**

Human performance accelerators
and new tools in a worker's toolkit

**UNLOCKING NEW
POTENTIAL**

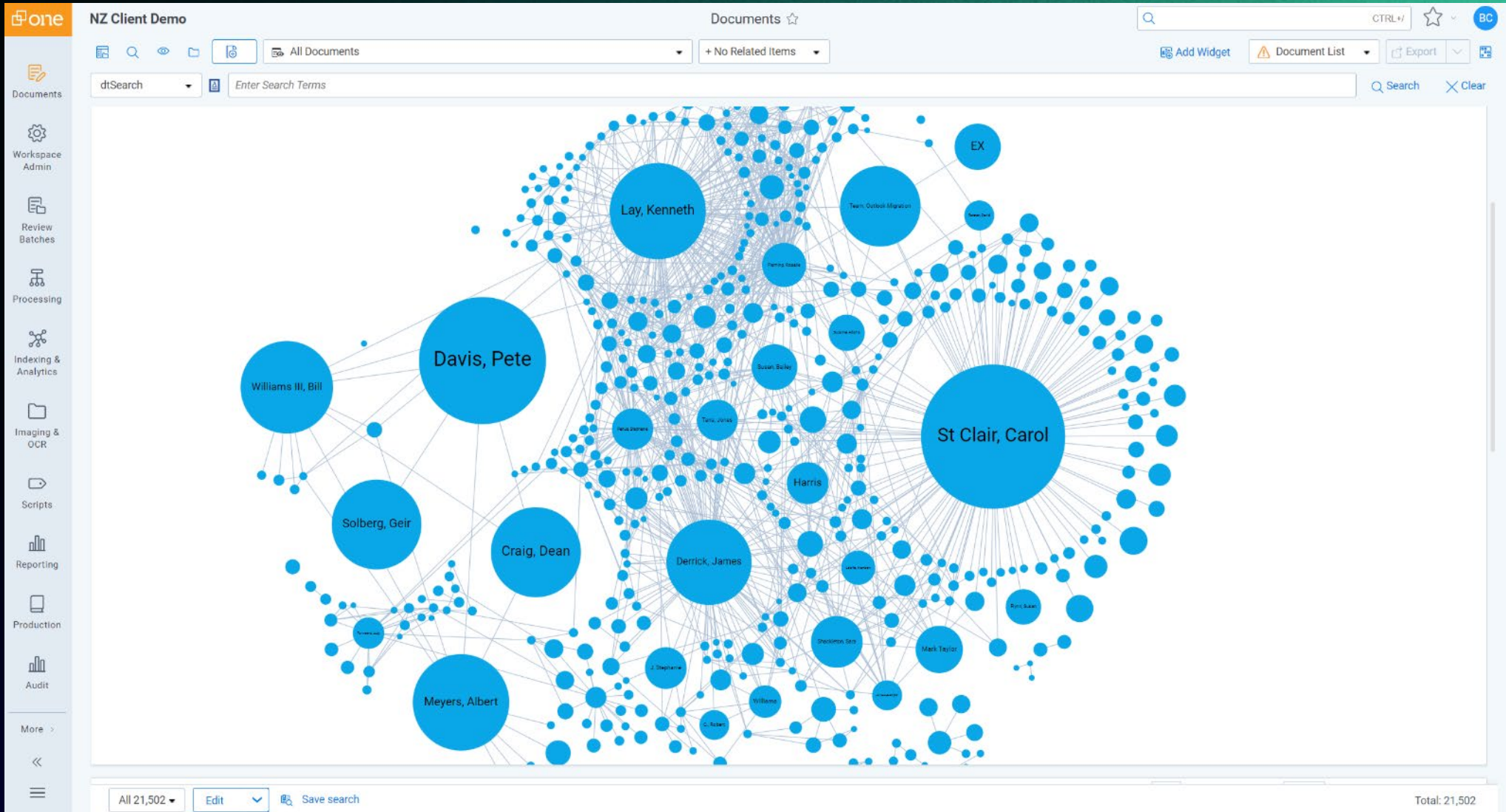
**Do Different
Things**

Rapid innovation for sectors, products
and services that leverage data

AI for Legal



Social network analysis

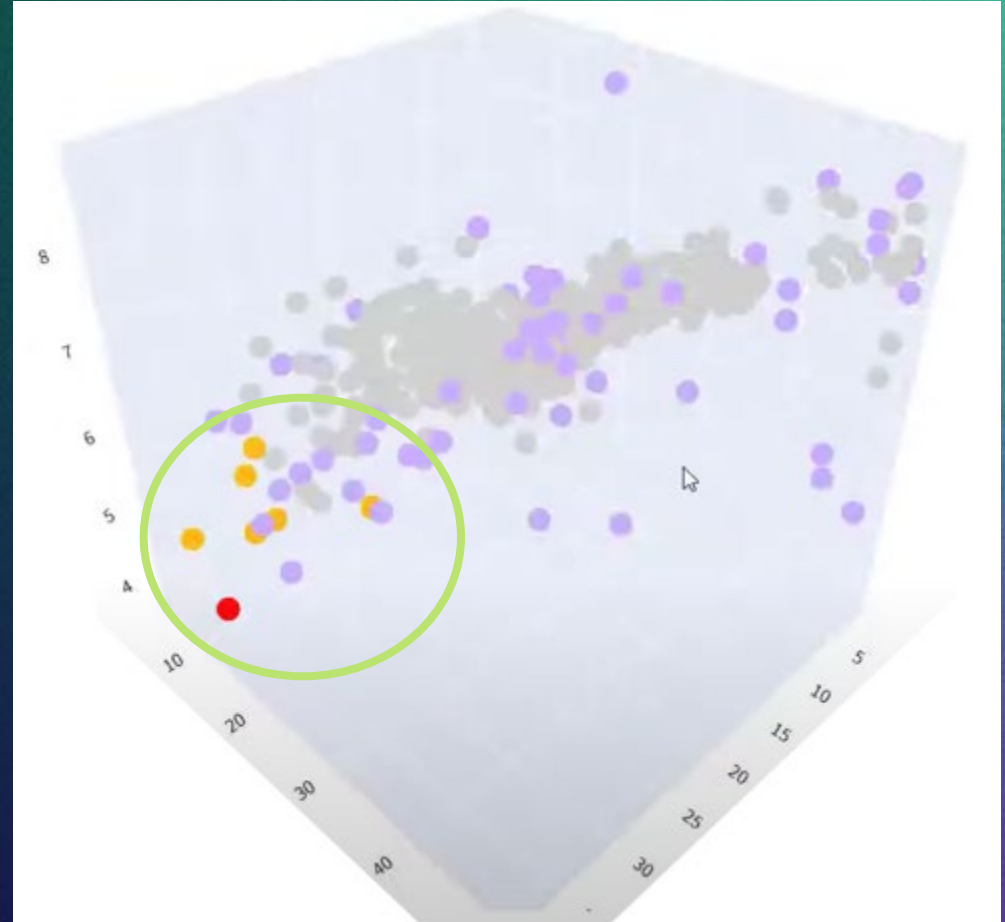


Outlier detection

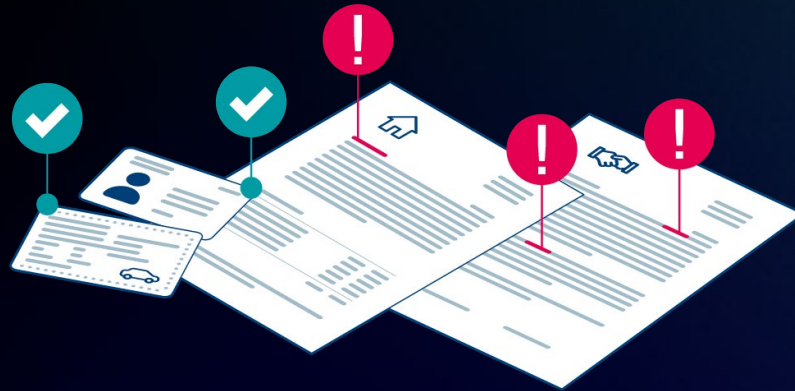


Rule-based anomaly detection finds behaviour you know to look for

Outlier detection finds behaviour that sits outside the norm



AI driven document verification



Multiple detection indicators

Score: Warning
Manual Inspection Recommended

Detailed Indicators
Evidence of changes guides verification

Edits highlighted in the document
For easier inspection

Secondary Indicators Confirm Modification
False alarms reduced by better reasoning

Whole Story
Modification traces confirm the person of interest's approach

© 2024. For information, contact Deloitte Touche Tohmatsu Limited.



FILE.PDF

HIGH RISK

AUTHOR:
ADMINISTRATOR

CREATION DATE:
05/03/2019, 17:07:22

MODIFICATION DATE:
23/10/2019, 15:24:57

RISK INDICATORS

PDF was modified [show in document](#)

This document was produced in several versions. This version of the document differs from the original.

Edited using editor [show in document](#)

This document contains tags used by PDF editors, which indicate modifications of the document.

Contains duplicit object definitions

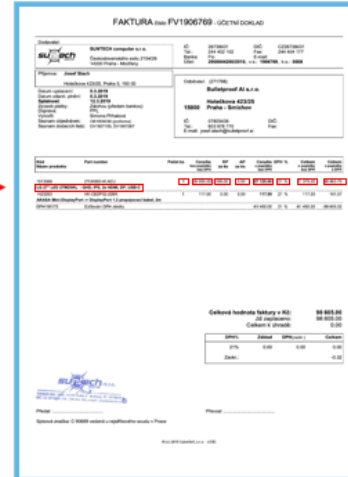
This document contains multiple objects with the same id. This may indicate that the original content was replaced by a new one.

Evidence of modification based on document metadata

Time of creation is 2019-03-05 17:07:22+01:00 but the time of modification is 2019-10-23 15:24:57+02:00. This may indicate that the document was updated or modified 231 days, 21:17:35 after creation.

Evidence of modification based on document trailer

File trailer, which describes the internal structure of the document, indicates that the file was incrementally updated or modified.



Sentiment analysis

Re: filling of wetland and ponds fup

Sent: Mon 16/01/2006 10:22:59 AM (UTC)

From: Trudy

To: Jeb Bush

My situation has gotten worse--The adjacent property owner is obviously continuing the drainage-without benefit of rainfall and has pumped water through his underground system to keep my 2 acres wet below the surface and continued as a large swamp area---that was a large upland grazing area for horses with a small pond in the center. I am spreading the word---**Jeb Bush is not for the Property Rights** but is directly against property rights---and I have the pictures and videos to prove my statement-----because of Jeb Bush and his agency-**I now have a condemned swamp!! My savings have gone into the swamp!!!!**

From: [Jeb Bush](#)

To: [Trudy](#) ; [Murray, Ted M.](#)

Cc: [Jim Roberts](#)

Sent: Tuesday, September 20, 2005 8:29 PM

Subject: RE: filling of wetland and ponds fup

where is this happening?

Jeb

Document ID JB0000402830

Document Title Re: filling of wetland and ponds fup

Document Type Email

Document Date 1/16/2006 10:22 AM

GAIT Response

Sentiment YES

Sentiment Type Anger; Frustration; Disgust

NavigAite_Sentiment_Score Anger: 80%; Frustration: 75%; Disgust: 70%

NavigAite_Sentiment_HTML	Sentiment Type	Sentiment Phrase
	Anger	Jeb Bush is not for the Property Rights
	Frustration	I now have a condemned swamp!!
	Disgust	My savings have gone into the swamp!!!!

Relevance & privilege review

PRIVILEGED AND CONFIDENTIAL

TO: Mark Haedicke
Peter Keohane

DATE: March 21, 2001

FROM: Chris Gaffney

DEPT: ECC – Legal

SUBJECT: Equity Investment, Services, and Wholesale Energy Supply Transaction with G6 Energy Corp.

ATTORNEY-CLIENT PRIVILEGE – DO NOT PRODUCE

LEGAL RISK RATING: 4

Factual Background

G6 Energy Corp. (“G6”) is an entity being formed to market retail electricity and services to small industrial and commercial consumers. The initial shareholders of G6 will consist of Enron Canada Corp. (“ECC”) and the corporate affiliates of four municipal electric utilities (“MEUs”): Hydro Mississauga, London Hydro, St. Catherines Hydro and Oshawa Public Utilities Commission. Pursuant to a Unanimous Shareholders’ Agreement (“USA”), Wholesale Energy Marketing Services Agreement (“Services Agreement”), Share Option Agreement (“Option Agreement”), master physical and financial trading agreements and other collateral agreements Enron Canada Corp. will:

- (a) hold a 20% equity interest in G6;
- (b) have an option to increase its equity ownership in G6 to 50%;
- (c) be the exclusive provider of wholesale energy marketing services to G6; and
- (d) be the exclusive wholesale supplier of electricity to G6.

Default Category	
Document ID	ENR.1078.0002.3596
Gait Response	
NavigAite_Relevance	YES
NavigAite_Relevance_Reasoning	The document discusses the equity investment, services, and wholesale energy supply transaction with G6 Energy Corp. It mentions Enron Canada Corp.'s equity interest in G6 and its exclusive role as the wholesale energy supplier. This is relevant to Enron's credit rating downgrade or decline in value.
NavigAite_Privileged	YES
NavigAite_Privileged_Reasoning	The document is marked as privileged and confidential, indicating that it contains privileged information.
NavigAite_Review_Summary	The document is marked as privileged and contains relevant information about Enron's equity investment and exclusive role as the wholesale energy supplier for G6 Energy Corp.

Trustworthy AI



MISCONCEPTION?




**“It is quick
and easy to
deploy”**

MISCONCEPTION?

**“AI is going to
take over the
world.”**



MISCONCEPTION?



**“It’s accurate
and unbiased
because it’s
based on data”**

What can go wrong with AI?

Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter

Attempt to engage millennials with artificial intelligence backfires hours after launch, with Tay Tweets account citing Hitler and supporting Donald Trump



What Air Canada Lost In 'Remarkable' Lying AI Chatbot Case

Marisa Garcia Senior Contributor @
I offer an insider's view of the business of flight.

Follow

Feb 19, 2024, 06:03am EST



ChatGPT: US lawyer admits using AI for case research

27 May



Meta AI Bot Contributed to Fake Research and Nonsense Before Being Pulled Offline

Nikki Main

Published 4 months ago: November 23, 2022 at 8:25 am - Filed to: ACADEMICDISCIPLINES



Bloomberg

Use TV Markets Economics Industries Tech Politics Businessweek Opinion More

Sign In

Subscribe

Nela Lofgren

HUMANS ARE BIASED.
GENERATIVE AI
IS EVEN WORSE

Stable Diffusion's text-to-image model amplifies stereotypes about race and gender - here's why that matters

By Leonardo Nicaletti and Dina Bass for Bloomberg Technology - Equality

Amazon warns employees against use of ChatGPT: Report



Photo Credit: Pixabay

Shradha Goled | 27 Jan, 2023



Like any progress, it isn't without risk

IP and copyright issues



Inaccuracy and hallucinations



Lack of transparency



Bias and discrimination



Data privacy and security



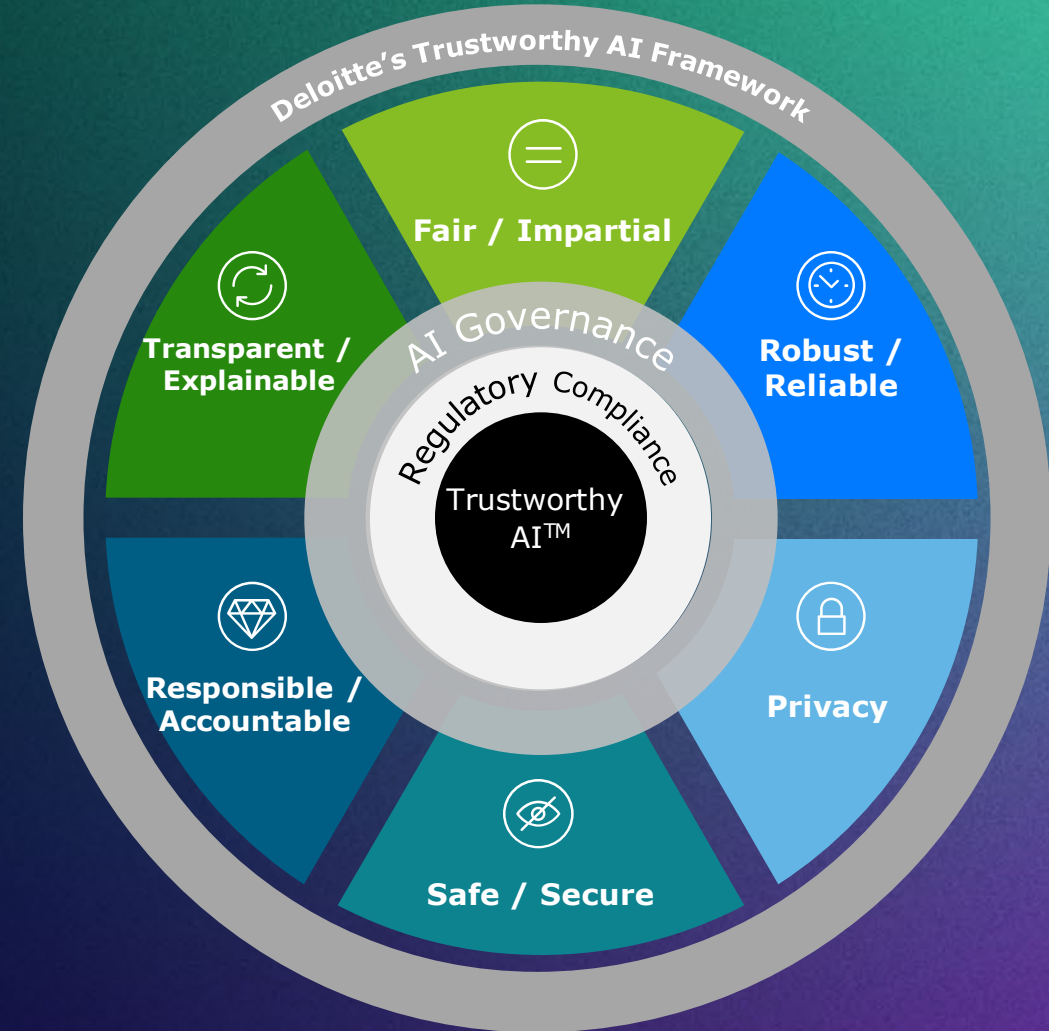
Fraud



A safeguard: Trustworthy AI

Trustworthy AI is a socio-technical ecosystem that is **ethical, lawful** and **technically robust**.

It is achieved through **governance** and **management of AI risks** across **people, processes** and **technology**.



Steps toward trustworthy AI



PLAN



BUILD



INTEGRATE



OPERATE

Design, assess and develop individual AI use cases

Use case **ideation and target definition**, perform **feasibility assessment**.

Proof-of-concept creation and **biases, privacy** leakage analysis.

Put the use case into **production**, and **industrialise components**.

Operate and ensure the performance and the **model's reliability**.

Establish trustworthy AI governance

Map the current ethical practice, set the bar and **build long-term vision** and ambition.

Translate the vision and ambition to AI **governance processes**.

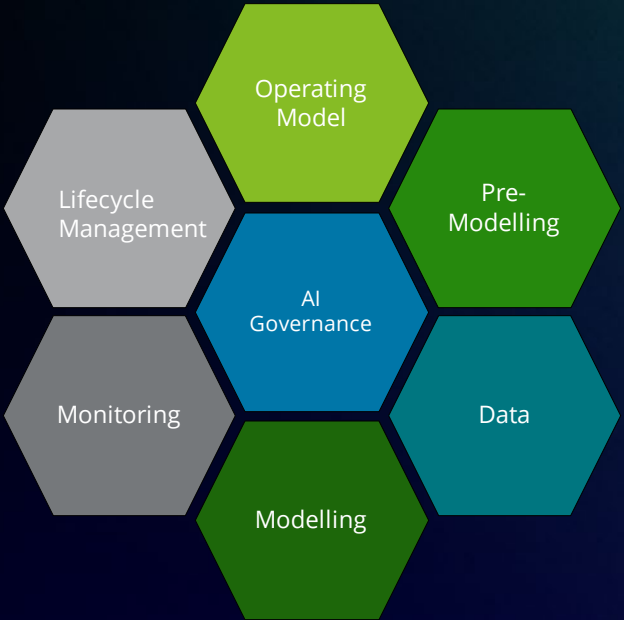
Enforce the AI governance and **industrialise processes**.

Validate and evaluate trust, regulatory compliance, processes and infrastructure.

Two layers: Overarching AI governance and specific model validation

AI Governance

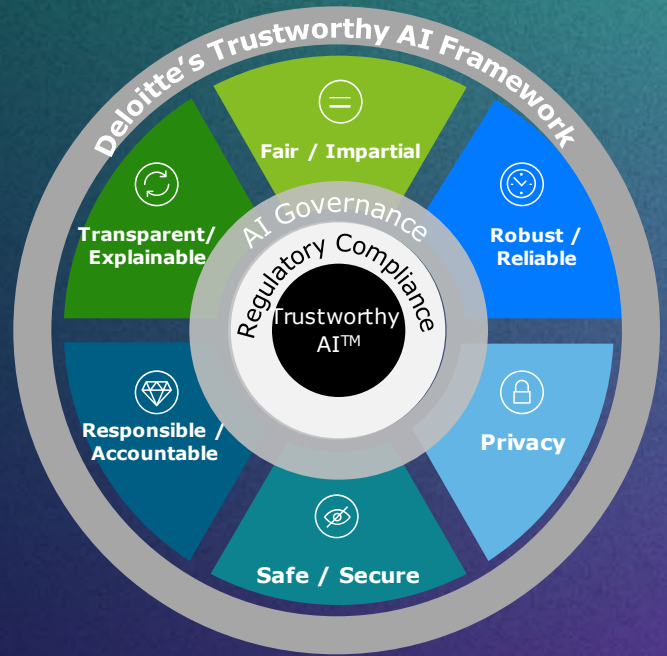
Define the AI Governance
 Ground your organization's AI ambitions in a practical roadmap along the development process based on identified risks



Execute the AI Governance strategy
 Deploy tools according to the identified risks and the agreed AI Governance

Model Validation

Execute Model Validation
 Perform an extensive model validation based on the Trustworthy AI Framework by using the aiStudio Tools and the audit catalogue



Raise awareness on AI and its potential risks
 Every involved stakeholder will become literate in AI and its risks

Guiding principles for trustworthy Generative AI



The Human Touch

Keep the **human in the loop!**



Empower Efficiency

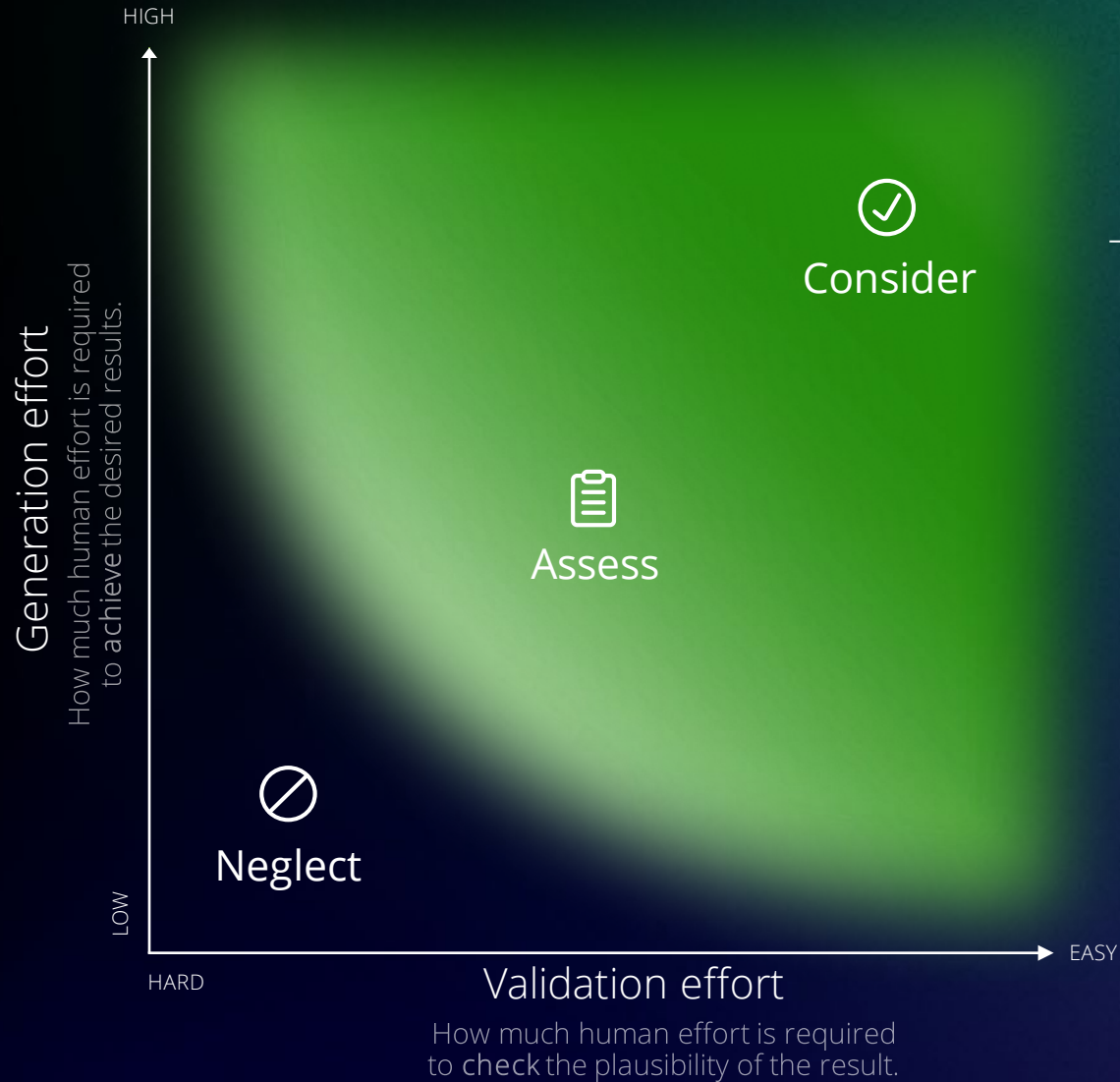
Repetitive, low complexity, low risk tasks are appropriate testing ground for generative AI solutions.



Truth Testing

To counter 'hallucination' risk, human **fact-checking** processes need to be in place to confirm accuracy of generative AI content.

Focus where human effort is high and validation is easy

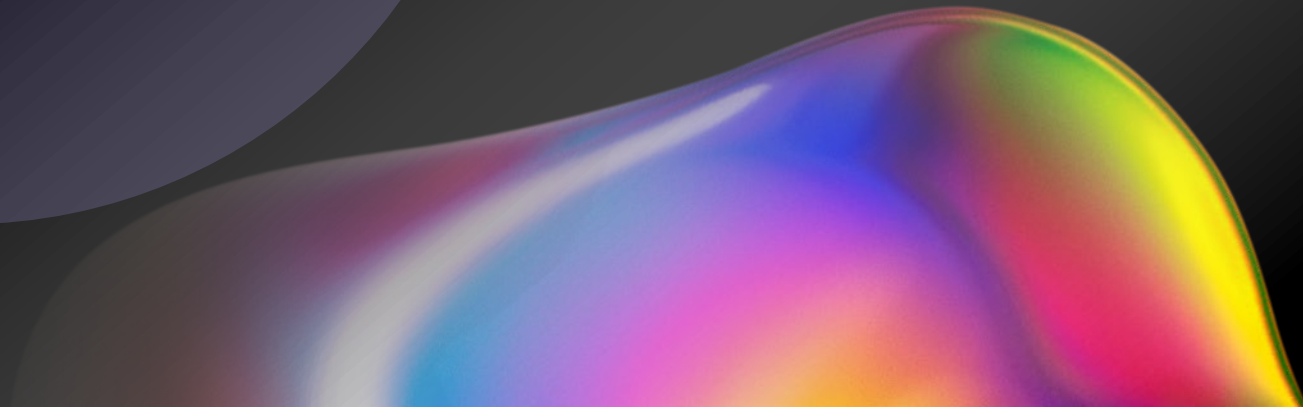


Working *with* AI: Easy first wins

- Draft customised communication
- Process documents (first pass)
- Provide customer service prompts
- Aid research and summarisation



Q&A



Guidance and resources

- [Government Use of Artificial Intelligence in New Zealand: Final Report](#)
- [On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?](#) (but also search for the follow-up and fallout from this article).
- [A new vision of artificial intelligence for the people](#)
- [Does GPT-2 Know Your Phone Number?](#)
- Privacy Commission [Generative Artificial Intelligence update](#)
- Digital.govt.nz [Interim Generative AI guidance for the public service](#)
- Deloitte AI Institute [The State of Generative AI in the Enterprise 2024](#)



Deloitte refers to one or more of Deloitte Touche Tohmatsu Limited (“DTTL”), its global network of member firms, and their related entities (collectively, the “Deloitte organisation”). DTTL (also referred to as “Deloitte Global”) and each of its member firms and related entities are legally separate and independent entities, which cannot obligate or bind each other in respect of third parties. DTTL and each DTTL member firm and related entity is liable only for its own acts and omissions, and not those of each other. DTTL does not provide services to clients. Please see www.deloitte.com/about to learn more.

Deloitte Asia Pacific Limited is a company limited by guarantee and a member firm of DTTL. Members of Deloitte Asia Pacific Limited and their related entities, each of which is a separate and independent legal entity, provide services from more than 100 cities across the region, including Auckland, Bangkok, Beijing, Bengaluru, Hanoi, Hong Kong, Jakarta, Kuala Lumpur, Manila, Melbourne, Mumbai, New Delhi, Osaka, Seoul, Shanghai, Singapore, Sydney, Taipei and Tokyo.

Deloitte provides industry-leading audit and assurance, tax and legal, consulting, financial advisory, and risk advisory services to nearly 90% of the Fortune Global 500® and thousands of private companies. Our professionals deliver measurable and lasting results that help reinforce public trust in capital markets, enable clients to transform and thrive, and lead the way toward a stronger economy, a more equitable society and a sustainable world. Building on its 175-plus year history, Deloitte spans more than 150 countries and territories. Learn how Deloitte’s approximately 415,000 people worldwide make an impact that matters at www.deloitte.com.

Deloitte New Zealand brings together more than 1800 specialist professionals providing audit, tax, technology and systems, strategy and performance improvement, risk management, corporate finance, business recovery, forensic and accounting services. Our people are based in Auckland, Hamilton, Rotorua, Wellington, Christchurch, Queenstown and Dunedin, serving clients that range from New Zealand’s largest companies and public sector organisations to smaller businesses with ambition to grow. For more information about Deloitte in New Zealand, look to our website www.deloitte.co.nz.