Journey

The Trusted Identity Platform

Avoiding GAI's Privacy and Regulatory (GDPR) Risks

Emmy Sobieski, CFA

Executive Summary:

"Whether person or bot, once they see and learn off personally identifiable information (PII), it's almost impossible to unsee, and unlearn. The only solution is to not show them the information in the first place. This can be accomplished using our zero knowledge technology." - Brett Shockley, CEO, Journey

Journey's zero knowledge technology avoids the complex and evolving risks inherent in the GDPR. This is especially true when these risks are magnified by the new and fast growing Generative AI (GAI) tools.

While language-based GAI escalates the risks of deep fakes and bad actors, the biggest current business risks may be regulatory.

- The majority of employees using the web based chatGPT for work are not telling their bosses. GAI systems like ChatGPT learned off the internet which contained PII*.
- Employees may be using GAI systems like ChatGPT to run analyses of customer trends, thereby sharing customer PII with OpenAI.
- Nearly half of all customer service centers are already using GAI bots, and it is highly likely these bots learned off data sets containing PII.

**PII for the purposes of this paper includes all forms of personal information. PII in this paper incorporates PII (personally identifiable information, as well as PFI (personal financial information), and PHI (personal health information).*

GDPR and CCPA state that a company must be able to remove personally identifiable information (PII) that are in its systems. Just like a person who learns something after reading a story that contains personal information, a bot cannot unsee what it has learned from. With GAI, the time to remove PII is before it is used. Zero knowledge technology allows GAI to learn and improve without being exposed to PII.

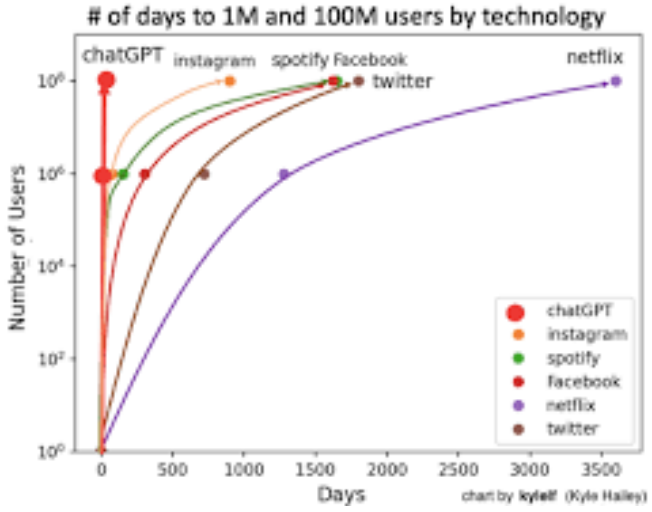
This paper explains how GAI and ChatGPT seemingly came out of nowhere to be used by half of customer related activities and more generally in over half of all companies, what the regulatory risks are to using these systems, and how these can be mitigated using zero knowledge technology.

1. ChatGPT is in the building. It's in *your* building.

ChatGPT currently has 100 million users and 1.6 billion visitors per month.

A Korn Ferry survey of professionals, 46% say they are currently using ChatGPT as part of their workday, Open AI reports 80% of Fortune 500 companies have ChatGPT accounts (based on corporate registered emails), and Business Insider found that 70% of those using ChatGPT at work did so without telling their bosses.

ChatGPT is the fastest technology to be adopted ever



While ChatGPT is likely the fastest growing and most used Generative AI (GAI) and Large Language Model (LLM) system used by your organization today, there are a number of GAI systems, including ones by Google (BardAI), Amazon (Titan), Microsoft (Bing AI), and many more, including international players. The concerns raised in this paper are not limited to ChatGPT. This paper will refer to all these systems, including ChatGPT, as GAI.

For larger companies with huge customer service organizations, GAI driven internet bots are being used to increase self-service containment dramatically.

GAI bots can be a powerful addition to your customer service. These days, GAI bots sound more natural and human, especially in an agent assistance function where it is listening to the call and providing recommendations to the agent on what to say or type. This can improve both costs and service levels while gradually training and testing out the bots.

Think this talk of GDPR risks of GAI bots in your customer contact centers is something to worry about later? Think again.

Generative AI is already embedded into contact center software. Like ChatGPT in companies, using GAI in customer contact centers is so tempting, companies are implementing it despite the risks. It makes current agents better and more effective while providing massive efficiency gains (companies say they would have to hire more than twice the number of agents without the GAI bots). Therefore, GAI is getting incorporated into contact centers at a rapid pace, and without addressing the GDPR compliance risks GAIs create.

Nearly half (47%) of global companies plan to use GAI for their customer-related activities, according to Metrigy's Customer Experience Optimization:2034-24 global study of 641 companies.

Metrigy's survey found organizations are using GAI on a variety of platforms, including:

- Customer feedback (47%)
- Contact center (46%)
- CRM (44%)
- Unified communications and collaboration (41%)
- CPaaS (40%)

What are ChatGPT, Large Language Models (LLMs), and Generative AI (GAI)?

Computers "learn" to read language by assigning numbers to each letter, seeing how they are grouped together, over time building more contextual understanding.

AI got "smarter" by adding newer algorithms to the menu, accessing larger data sets and using faster computers.

One of the large successes of ChatGPT is its narrow focus on predicting the next thing.

The process of predicting the next thing is called Generative AI. GAI takes everything into context, i.e. it has ingested large portions of the internet as training data (thus it is called a Large Language Model, or LLM), then predicts the next most likely word it should write, or you should write. It uses this same generative ability to write computer code.

Although our brains are far more complex, we do think in a similar predictive manner.

“With each movement, the neocortex predicts what the next sensation will be. If any input doesn't match with the brain's prediction … this alerts the neocortex that its model of that part of the world needs to be updated.” Jeff Hawkins, author, A Thousand Brains, A New Theory of Intelligence

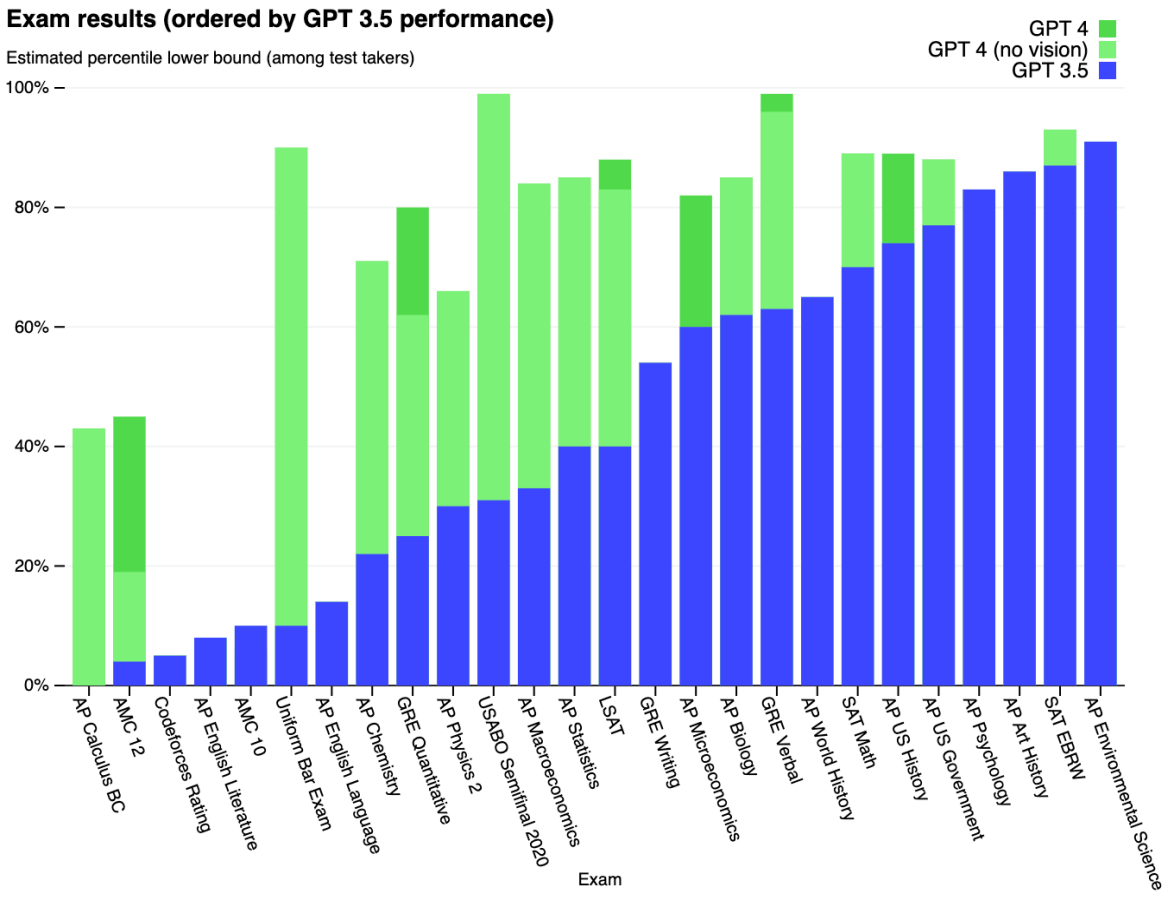
AI has been around for over 50 years, but the excitement since the November 30, 2022 release of ChatGPT is palpable and feels different.

Why is this moment different?

We have crossed a threshold.

GAI’s effectiveness and potential for efficiency gains are impossible to ignore. GAI can write papers, pass tests, then do your work when you get a job. In a world is based on trust and language, GAIs are poised to multiply both risks and opportunities.

“[GPT4] passes a simulated bar exam with a score around the top 10% of test takers; in contrast, GPT-3.5’s score was around the bottom 10%.” - Open AI website



This is Enterprise's iPhone moment, Part 2

As the iPhone rose in popularity, companies preferred using Blackberry, with its secure server architecture and ability to remotely erase the phone. Many financial services firms mandated use of a company-issued Blackberry.

But, executives broke the "rules" and started using their personal iPhones for company purposes. The iPhone made their jobs easier.

Companies were left playing catch-up.

As with the iPhone, stopping ChatGPT's use may not work. Accepting that it and other generative AI products (bots and web-based AI apps) are in use, what's the best path forward to assess risks and opportunities?

Common questions about GAIs GDPR risks:

- Did GAIs ingest Personally Identifiable Information (PII) to learn? (yes)
- Can GAIs identify people using non-obvious PII like digital footprints and identifiable patterns? (yes)
- When using GAIs at work, how secure is your corporate information and PII? (high risk that corporate information and PII are being shared with web-based GAIs in an unsecure manner)
- Is your company's use of GAIs violating privacy regulations like GDPR or CCPA? (high risk that the answer is yes)

2. GAI's GDPR Regulatory Risks

This paper is focused on GDPR, but similar issues exist for CCPA and other customer data privacy regulations around the world.

1. GDPR compliance was hard BEFORE GAIs like ChatGPT
2. The stakes are high. GDPR fines have been as high as €1.2 billion, while GDPR compliance with GAIs currently difficult to impossible, and half of US companies are already using GAIs like ChatGPT.
3. GAI has 3 main issues for GDPR compliance:
    a. GAI "learned" by training itself on portions of the internet containing PII. It didn't ask for each consumer's permission ahead of time.
    b. GAI generates new content based on what it read and learned. GDPR requires a company be able to identify and remove PII upon request by that individual. Bots, like humans, can't unsee or unlearn.
    c. Your employees are (inadvertently) sharing customer PII with the companies that own these GAI systems, like Google (BardAI) and OpenAI (ChatGPT) to "organize customer data, purchases, etc."

1. GDPR compliance was hard even before GAI entered the picture.

Here are the main requirements for GDPR compliance:

1) Anonymize or pseudonymize Personally Identifiable Information (PII) such as names, addresses, or contact details, before being used;
2) Obtain explicit consent from each person before using PII;
3) Limit data collection and retention of PII;
4) Implement appropriate security measures like encryption and access control to avoid security breaches;
5) Provide transparent privacy policies that detail how personal data is handled, processed and whom it will be shared with, and how long it will be retained;
6) Enable individual user rights over their data, so they can access, rectify, erase, restrict processing, and export personal data;
7) Conduct regular assessments and audits of your data handling practices; and
8) Collaborate with data processors to ensure they are GDPR compliant.

Before GAI, complying with GDPR was difficult, complex and expensive. Now, meeting these requirements are basically impossible.

2. The stakes are high - GDPR fines up to €1.2 billion

GDPR went into effect May 25, 2018. The fines for non-compliance have run into the hundreds of millions, to the record €1.2 billion fined against Meta this year. The largest 20 GDPR fines in 2021-2023 have been levied on social media, internet, and telecom giants, but others, notably the large clothing company H&M, are not immune.

Notable GDPR fines 2021-2023) - according to Data Privacy Manager

Fine (In Euros)	Company	Year	Issue
€1.2 billion	Meta	2023	Transfer PII, Inadequate data protection
€746 million	Amazon	2021	Used PII for Ad targeting without consent
€405 million	Meta	2022	Teenagers' PII displayed by instagram
€265 million	Meta	2022	Data breach leaked PII of 533 million users
€225 million	WhatsApp	2021	Lack of transparency of WhatsApp's use of PII
€50 million	Google	2019	Lack of control, information, consent regarding PII use for ads
€40 million	Clearview	2022	Collecting 10 billion facial images. No consent or right to remove

You can see that many of the GDPR fines focused on how companies treat PII. Failure to comply with GDPR, whether on purpose or inadvertently, is serious, expensive, and can lead to huge fines. And GAI takes GDPR compliance complexity to a whole new level.

3. GAI use has 3 main issues that risk GDPR non-compliance:

a. GAI systems used to power ChatGPT and Google's BardAI used training data that included portions of the internet. It is highly likely that this data contains inadvertent PII. According to GDPR, each person whose data was used should have been asked for their permission ahead of time. This is impossible given the scale of the internet.

b. These systems have "learned" and are now in use, the companies that own these systems cannot comply with GDPR by removing PII on request, because such systems incorporate PII in ways that cannot be undone

c. Your employees may be (inadvertently) sharing PII (as well as other internal proprietary data) with Open AI when using the web-based Chat GPT to "organize customer data, purchases." A public company CEO asked ChatGPT to write his earnings release - exposing non-public information to OpenAI over non-secure web channels.

GAIs like BardAI and ChatGPT "learned" by training data containing PII

The data ingested for the GAI to learn likely contained PII. After the model has ingested, transformed, and evolved itself off data and PII, it is next to impossible to give consumers the transparency and control that is required for GDPR compliance. You would be asking the model to "un-see" what it learned off of. And, unlike a human who has been exposed to PII, the AI cannot be trusted to not divulge what it knows is PII.

And today we think of PII quite simplistically, identifiable information that can be directly tied to a person, name, DOB, Address, SSN etc. This entirely misses the fact that people have a digital footprint that may be just as identifiable as these more obvious particulars. AI can infer and tie these digital footprints back to PII.

The regulatory pressure to remove PII from every-day use will only increase.

On August 28, 2023, Open AI released ChatGPT Enterprise which is Soc2 compliant. This is a great step forward in providing security for keeping private information secure, addressing a big portion of GDPR regulations.

But Soc2 compliance does NOT solve the issues we are highlighting here, that PII was seen without prior permission from each individual, is now part of GAI's knowledge, cannot be unseen or removed, both of which are requirements of GDPR.

Your employees may be (inadvertently) sharing PII with the owners of GAIs

When employees use GAIs like ChatGPT for business use, there are two levels of violations. First, GAI violates GDPR (because it likely ingested and learned off PII when training without previous permission from the individuals who own that PII and now without the ability for those same individuals to have that PII removed), and thus isn't allowed for business use. We covered this first issue above.

Second, there's a high risk that employees are inputting sensitive company data along with customer PII into the system as a matter of course.

It's incredibly tempting for an employee with a task of running data analysis on customer purchasing trends, for instance, by zip code, to input customer data, addresses, purchase history, into a GAI like ChatGPT and ask it to create charts and analysis. The job is often done faster and better, making the employee look great.

It gets easier and easier to accidentally incorporate PII. For example, by starting to make use of AI-enhanced spread-sheet technology as well as the multitude of "extensions" similar to Chrome extensions that lets your employees use the GAI for a specific purpose. These extensions are often created by different organizations than the GAI owner, and add to your GDPR risks.

The number of companies offering GAI powered business applications is exploding.

But when your employees enter this data, they expose that customer data to the owner of the GAI, like Open AI or Google, and the owner of the extension they are using, violating multiple parts of GDPR and creating a regulatory liability.

GDPR compliance in a GAI world seems impossible.

What kinds of GDPR violations do GAIs create?

As a reminder, here are the main GDPR regulatory requirements:

1. **Anonymize or pseudonymize Personally Identifiable Information (PII)**
2. **Obtain explicit consent before using PII**
3. **Limit data collection and retention of PII**
4. **Implement appropriate security measures**
5. **Provide transparent privacy policies**
6. **Enable user rights**
7. **Conduct regular assessments and audits.**
8. **Collaborate with data processors**

As evidenced by Open AI's recent Soc2 compliance, LLMs and GAIs can solve for a number of these GDPR requirements, such as #4, #7, and #8, and to some extent #5.

However, GDPR goes further when it comes to user rights. A company must obtain explicit consent before using PII, limit collection of PII, limit retention of PII, enable user rights, and provide transparent privacy policies, and enable user rights.

User rights to explicitly permit use of PII beforehand, know where the PII is used, and have the ability to request to have their PII removed isn't possible once a GAI has learned off data with PII in it. The PII is part of the GAI's knowledge base, and this PII has likely been transformed into a different form, in the same way we read data and then understand something.

How can we remove the thing we read? GAI's can't do that either once they have learned off data containing PII.

In a world rapidly adopting GAIs, these GDPR requirements exponentially increase your regulatory risks. The use of GAIs and LLMs that have ingested PII without prior explicit consent and transformed that PII by learning off of it create a situation where the PII cannot be removed.

Finally, GAIs often learn using automatic and "unsupervised" steps, making it difficult to provide complete transparency in privacy policies.

With so much regulatory scrutiny and risk, why not just forbid the use of GAIs and ChatGPT in your company?

Even if you are worried about massive fines and know you can operate today without using a GAI or ChatGPT, banning their use may not be possible. As with the iPhone, people are going to use GAIs because the benefits are too great not to.

We have already mentioned that nearly half of contact centers are already using GAI, as well as half of all US corporations, and that doesn't count the fact that a majority of employees use this technology without their employer's knowledge.

Using GAIs will be a competitive advantage in nearly every industry.

One founder uses ChatGPT as a strategist.

She "trained" ChatGPT by having it read and ingest more than 10 of her best articles she had written plus other documents representing her best ideas. Now she uses ChatGPT as her head of strategy (a mind that thinks like hers, for $20/month), assistant, and to write first drafts of all her future content. She shared that using ChatGPT instead of hiring for those roles extended her company's cash runway by 50%.

We have barely scratched the surface on the potential benefits to using GAIs and LLMs. Yet, in spite of this potential, there are also huge regulatory and reputational risks. Why? Fundamentally, we can't always say what a GAI is going to do with the PII it ingests.

How will GAIs use PII based on the tasks given to it by others? To protect customer PII, we need to separate this kind of data from the data used to make GAIs conversational.

How can we get ahead by using GAIs without then falling into regulatory hot water?

Journey has solutions.

3. How to use Zero Knowledge technology to leverage GAI while reducing your GDPR risks

"The best way to protect data is to never share it." - Brett Shockley, CEO, Journey

The old Mark Twain adage “Two people can keep a secret if one of them is dead” holds true in the digital world. How do we avoid the risk of GDPR violations while still leveraging the benefits of GAI tools?

First, don’t let the AI see the private data, ever.

With human agents, the Journey’s Zero Knowledge networking technology secures customer data and allows transactions to occur without sharing any data, or alternatively exposing a very small subset on a strict need-to-know basis.

If we replace a human agent or assist a human agent with a GAI-enabled tool, we must address the added risk presented by the new ways that sensitive data can be 'leaked' by current and future GAI-equipped tools and 3rd party GAI extensions. We don’t know where that data might end up in GAI, extensions, or in the case of breaches into a corpus of corporate-owned GAI data.

What is Zero Knowledge networking?

Say you go to a bar... and it’s a bar that requires patrons to be over 21. The bar employees verify your age by checking IDs. Those IDs also contain your weight, eye color, home address, and other PII that is not relevant to the task of verifying your age. Even the actual birthdate is not necessary.

How can you verify one thing (born before x date) without exposing additional, unneeded information?

If I can “prove” that I’m over 21 by verifying “born before x date” through a trusted third party (for example a secure database owned by the ID provider or a verifier), the bar employee can let me in without seeing my information. Technologies like today’s restaurant server credit card terminals provide a similar function for payments; neither the server nor the restaurant ever see or have access to the actual credit card or your personal financial information.

Zero Knowledge proofs allow you to prove the information is true (or false) without sharing the source information.

Zero Knowledge technology can be used to provide LLMs with training data that has been cleansed of PII, disconnecting it from identifiable people while preserving GAI’s ability to learn.

PII may be at risk in more than one way when introduced into an AI paradigm.

The first risk occurs when training the AI.

Customer stories and experiences are used to train LLMs. An anecdote by itself is just an anecdote, but thousands or millions of anecdotes when correlated and learned from are knowledge. Just like humans learn from experience and the historical experiences of their mentors and teachers, AI learns from large numbers of experiences.

Even if the GAI was trained on zero knowledge data, thus it never saw PII, the second risk in contact centers occurs when a GAI bot is used as a knowledge engine to deal directly with a customer or to assist a live agent in dealing with a customer.

Exposing an agent to PII, or having a GAI provide recommendations after it learned on data containing PII, creates GDPR risk.

Similarly, exposing a GAI-powered internet bot to this data creates GDPR risks. This is especially true if the GAI engine that powers the bot lives outside your company's or call center's enterprise data controls. Your agents could be exposing data to 3rd parties, unknowingly increasing your GDPR (and other) risks and reducing your ability to adhere to or prove GDPR compliance.

A zero knowledge network model that gives the agent necessary proof of things like identity, or payments, and enables sensitive data to complete transactions without exposure to the agent or the AI is essential.

To solve for GDPR compliance while using GAIs requires a zero knowledge model that goes beyond the obvious PII. It requires that AI engines themselves be in a Zero Knowledge network that prevents ANY sensitive information from leaking into the LLMs.

How Journey uses Zero Knowledge technology to solve for GDPR compliance while letting your company benefit from using GAI's

Privacy and GDPR compliance start at the beginning.

Ingest and train on data that is clean of PII, not because you trust the data engineers to remove each violation, or have purchased "clean" data that promises not to contain PII. Instead, the data is clean because the data is in a Zero Knowledge environment, so the AI program never gets access to PII.

Once the GAI models are trained, the second layer of privacy protection is needed.

Protect PII from internet bots.

Just like you don't want a human being to see sensitive PII unless required, you also want to be careful how much of that data is stored and available to the bots. When you do choose to expose small amounts of data on a need-to-know basis, make sure the granularity of the information is appropriate for the use case. For example, share the state someone lives in but not their city or street address unless those are truly relevant. Journey's technology can be used to control data residency and exposure to raw customer data.

Now your customer enters the equation.

Secure customer authentication is required to do meaningful transactions in self-service. As GAI bots become more capable, they will take over more of the conversation. This increases your need to prevent GAI bots from using PII while maintaining your ability to authenticate and help the customer.

As the percent of service calls handled without humans (self-service containment) increases beyond the 50% mark in 2023, on its way to 70% by 2027, according to Gartner, the need for the human agents to give a unique and much stronger Customer Experience goes way up. Using Journey's tech can help companies improve CX, operational efficiency, security, and customer data privacy.

The opportunities for efficiency, scale and new business is immense with GAI and LLMs. We are learning how to use these powerful technologies and how to navigate regulations. Regulating new technology is always a challenge.

Taking the Zero Knowledge approach BOTH with the training of the GAIs and LLMs, AND later with the data you load into them for specific results, can greatly reduce your regulatory risk while improving your customer experience.