

ChatGPT:



*What we know.
What we don't.
How iResTech
is addressing it.*



Quirks London – May 3rd – 4th, 2023

What to expect today ...



Let's Talk Disruption

A brief history of disruption and what happens when companies don't respond fast enough.

What is ChatGPT?

We've all heard about ChatGPT but what is it and what can we really do?

Real Life Examples

Let's see what ChatGPT can do when applied to market research cases.

How iResTech is responding

How can technology be used to address the new disruptive threats presented by ChatGPT?

About Us

Passionate people. Creative solutions.



Chief Technical Officer

Ben Coats

The trail which led Ben to Innovative Research Technologies began when at 14 he was asked to take over IT for a family medical practice. At 16, he was commissioned to develop an electronic medical records platform for a group of forensic pathologists in the era of "going paperless". Little did they know that dead people need medical records too. Ben joined Quest Mindshare in 2020 with the task of building out an internal software engineering department. While assembling an all-star team, Ben was a strong advocate of turning that department into its own company which could serve other customers in the industry- a vision he shared with Greg Matheson, managing partner of Quest Mindshare and now CEO of iResTech.

About Us

Passionate people. Creative solutions.



Luke Bothwell

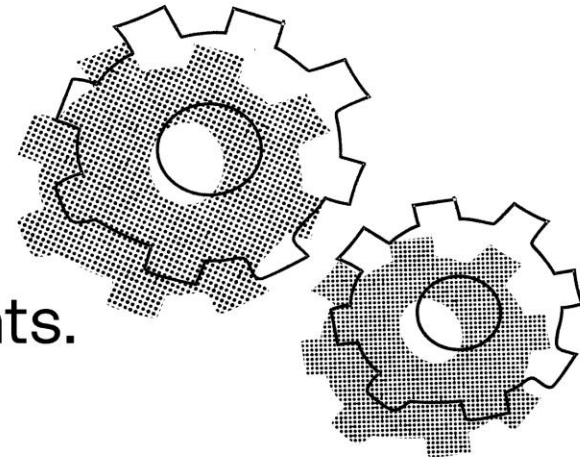
Luke joined the IRT team as Director of Product Management in May 2022, bringing with him more than a decade of experience as a Product Management leader in industries including insurance and financial services, identity management and protection, and healthcare. He thrives on understanding customer and business needs and transforming those needs into products and services that thrill users and solve problems..

Director of Product Management



Brief History of Disruption *and* Failure to Innovate

Disruption is a constant.
How a company responds to it can be
a matter of surviving and thriving or
locking the doors and turning out the lights.



Why Resistance?



Radio Shack Ad: February 16, 1991

Cost of goods replaced: \$3,054.82 in 1991.
\$6,563.83 in 2022 dollars.

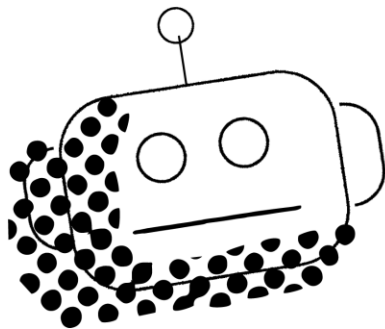


JUNE 29, 2007:
Apple debuts the iPhone and changes
the face of technology forever.



What is ChatGPT?

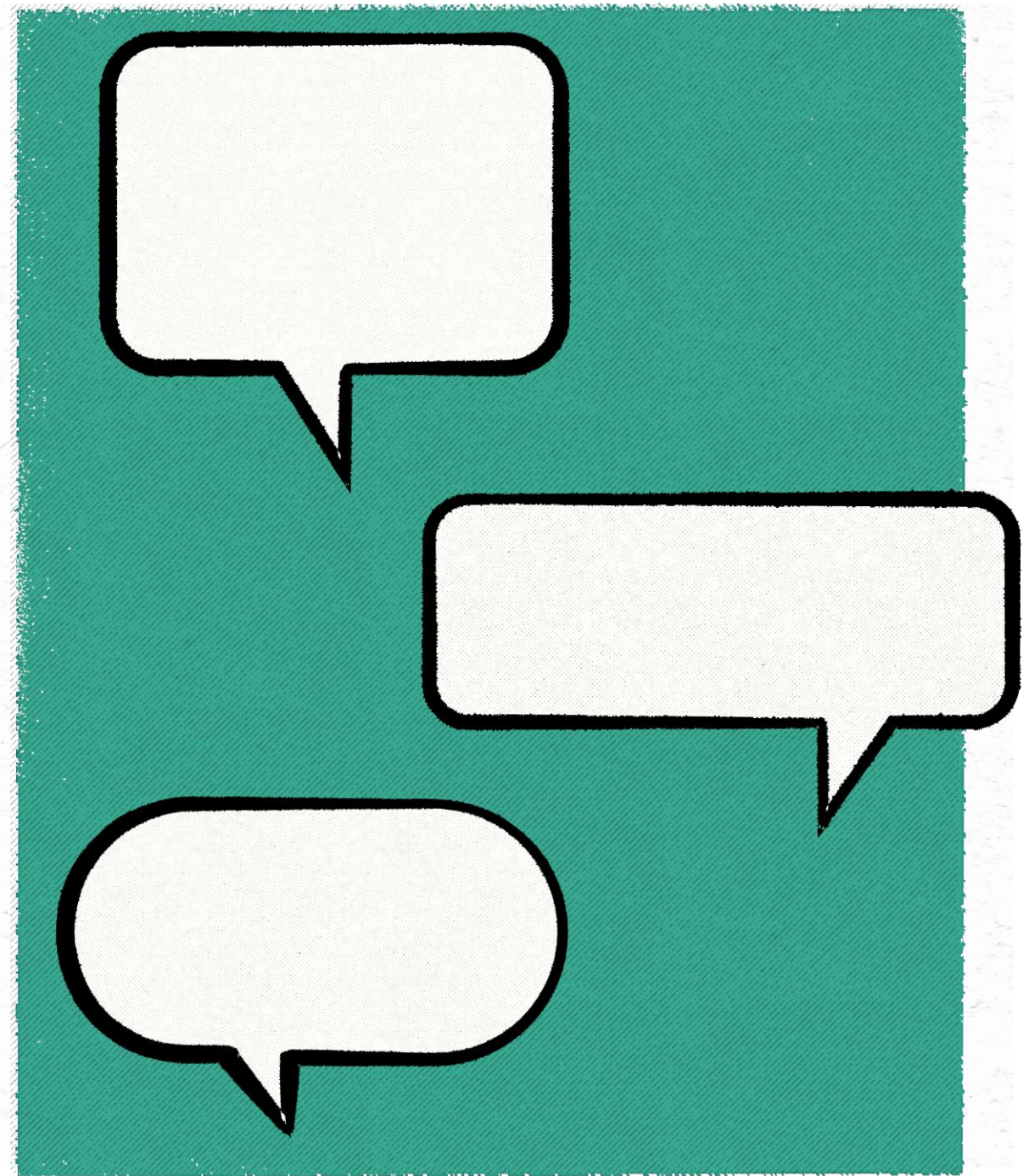
We've all heard about it, we are all talking about it, but *what* is it?



IT'S A BIRD!
IT'S A PLANE!
NO, IT'S... AN ALGORITHMIC
ANTHERO THAT CAN UNDERSTAND
NATURAL LANGUAGE AND GENERATE
INTELLIGENT RESPONSES TO
OPEN ENDED QUESTIONS.

It's in the news almost constantly but very few people actually know what it is and how it works

- Created by OpenAI, a company founded by Elon Musk and Sam Altman in 2015
- Backed by well known investors including Microsoft
- Per ChatGPT's own description: "ChatGPT is an AI-powered chatbot developed by OpenAI, based on the GPT (Generative Pretrained Transformer) language model. It uses deep learning techniques to generate human-like responses to text inputs in a conversational manner".
- Users enter written prompts and receive human-like text or images and videos generated by the AI.
- Open to the public via the OpenAI website



November 30,
2022

Initial Prototype
Released

January 30,
2023

Stable Release

\$29B

Valuation of OpenAI
after release


100M

Number of active
users two months into
its launch

ChatGPT By the Numbers

A white thought bubble with a black outline, containing the text "How could ChatGPT disrupt market research?". Two smaller circles lead from the bottom of the bubble towards the right.

**How could ChatGPT
disrupt market
research?**

A white speech bubble with a black outline, containing the text "Managing data quality and legitimacy is already a constant battle. AI could make the battle even tougher.". The bubble is set against a green, textured background that looks like a torn piece of paper.

**Managing data quality and legitimacy is
already a constant battle.
AI could make the battle even tougher.**

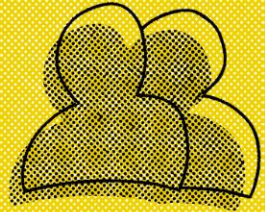


Potential AI Related Disruptions:

- Increased fraud and decreased sample quality
- Harder to distinguish a legitimate open-ended responses from a fraudulent one
- A chatbot will never buy a product or service



**Real life
experimentation
with ChatGPT**



How it Works

The laptop screen is split into two panels. The left panel shows a survey titled "Welcome to the survey!!" with the following content:


SEO stand for.

- Search Engineer Option
- Selected Engine Operation
- Search Engine Optimization

Please select common digital marketing tools.

- Design tools
- Social media platforms
- Analytics tools
- Call of duty
- None of the Above

Kindly describe the benefits of Digital Marketing.

I am human  [Privacy](#) [Terms](#)

The right panel shows a chat interface titled "Benefits of Digital Marketing" with a "Model Default" dropdown and a "ChatGPT PLUS" label.

Experimenting with ChatGPT



Hypothesis

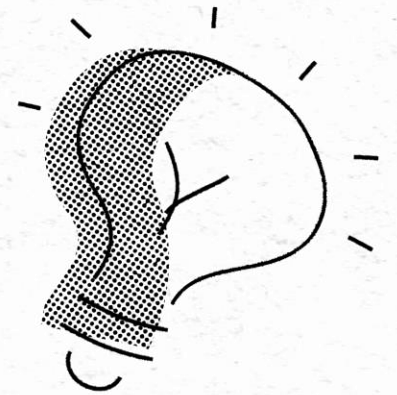
ChatGPT can be prompted to generate responses to open-ended questions that will appear to have been created by a human.

Procedure

The text for the primarily English open-ended screener questions was selected from years of historical data obtained from our sister company, Quest Mindshare. We manually presented 100 questions to ChatGPT and recorded its responses.

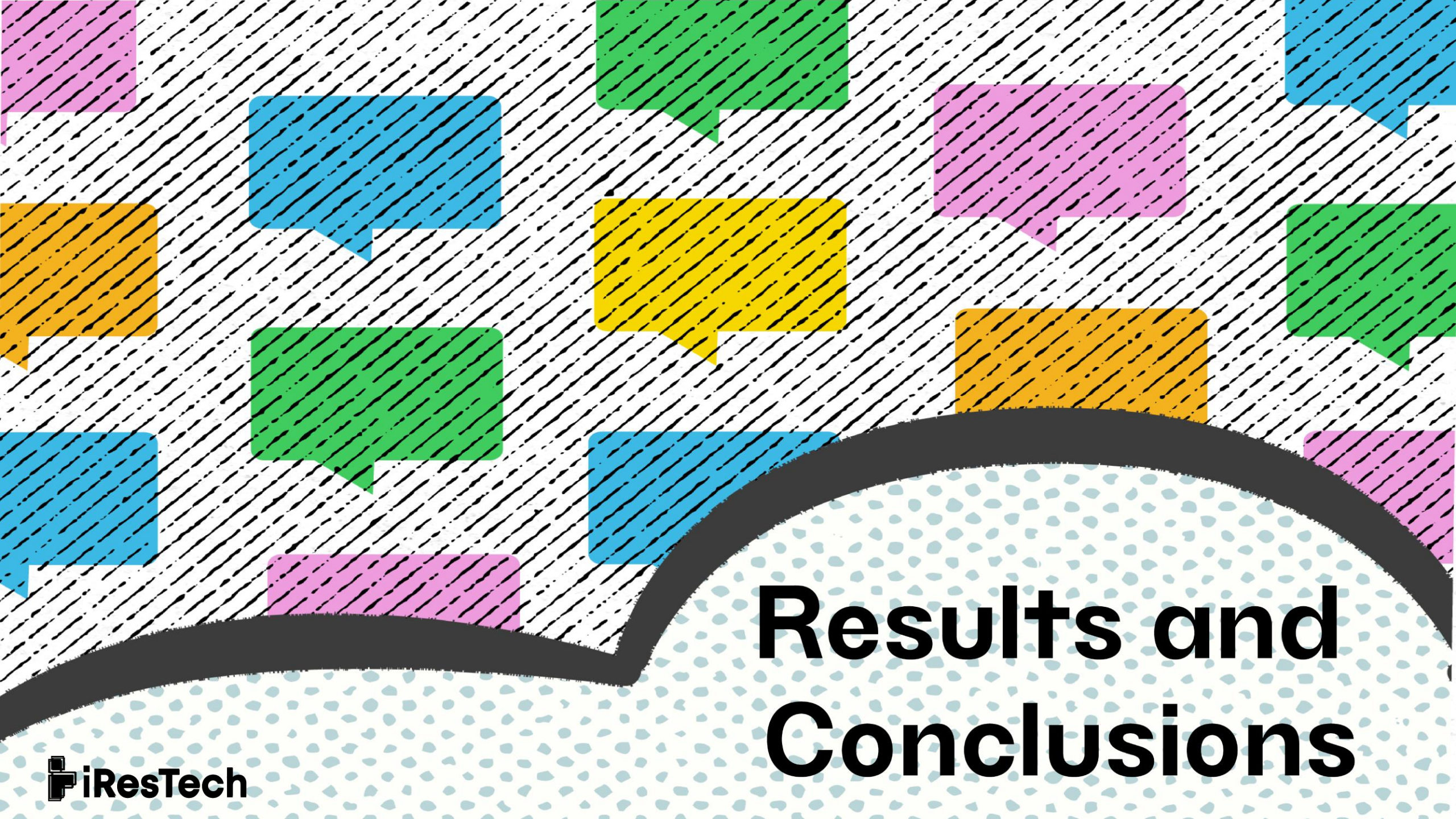
Conclusions

The findings were somewhat surprising. ChatGPT does indeed pose a threat to data quality and legitimacy. However, there are opportunities to mitigate this, and some don't require more technology.



Procedure Overview

- 100 sample questions selected at random from thousands of examples in Quest Mindshare's historical OE Q&As and supplied to ChatGPT one at a time (Thanks for the data Quest!)
- Answers were recorded for each question
- Response patterns observed and documented
- Detection / prevention techniques derived from response patterns and validated.
- Questions given a subjective good / bad "rating" depending on how well they enabled detection or prevention



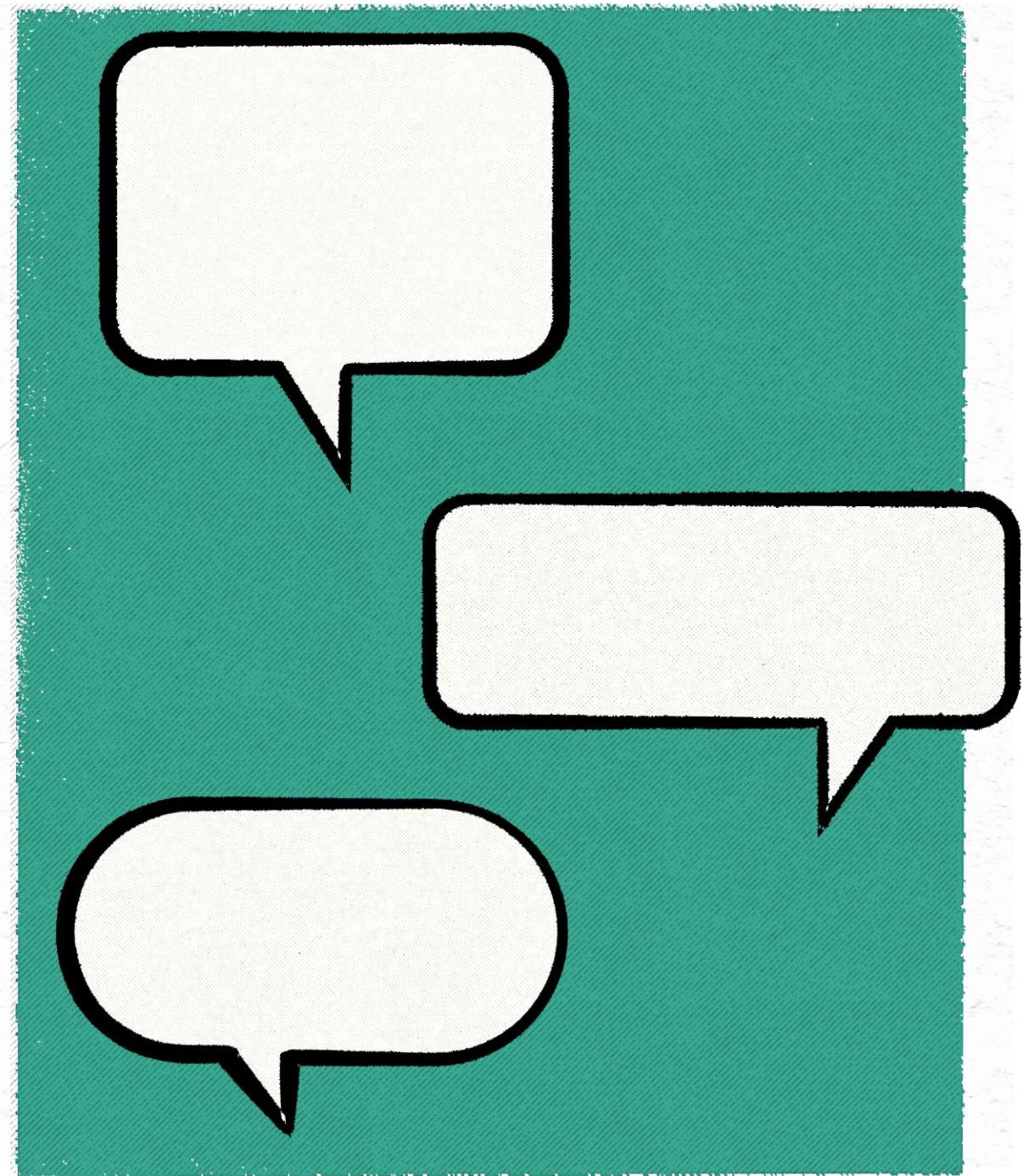
Results and Conclusions

Key Findings Overview

- ChatGPT poses a threat to data quality assurance if we don't develop techniques to accommodate this new technology
- It can provide "correct" answers to most questions and seem very knowledgeable, in many different languages
- There are some non-technical approaches which may help PMs identify ChatGPT answers – or even prevent ChatGPT from providing *any* useful response!
- At the rate the AI model is evolving, these approaches are short-term stopgaps only. Comprehensive technical solutions (such as those iResTech is adding to Dtect) are needed for the long term.

What can *you* do?

(...to recognize ChatGPT at work)



Tips for Detection!

ChatGPT is *programmed* to self-identify as an AI (or it would be scary) – so *make it self-identify!*

- Asking for *opinions*, about experiences or emotions, or anything which would require a physical presence with sensory abilities should prompt it to self-identify
- The more your question aligns with the latter, the more likely ChatGPT will be unable to answer!

If you can't ask subjective questions, look for several paragraphs and outline-like structures

- ChatGPT often presents its responses as an introductory sentence followed by numbered lists or bullet points of (often) 4 to 7 points
- Bullets/numbers are often lost (depending on browser or OS) on pasting, resulting in many paragraphs with a sentence or two per paragraph
- Excessive number of paragraphs may indicate that the respondent is not human.

Tips for Detection! (continued)

Do a visual length-check

- In historical data of only completed interviews, OE responses averaged 8.75 words (no matter how they were begged to “be descriptive”).
- ChatGPT’s answers averaged 221 words

Keep an eye out for common grammatical errors. [They may be a *good* sign!]

- Humans tend to make grammatical errors that are *somewhat* correlated with their socioeconomic standing or education level. (Others are quasi-generational. For example, millennials and younger tend to not begin sentences with capital letters in less formal communication.)
- ChatGPT will virtually never make these grammatical errors in its responses (unless prompted).
- **The absence of these errors doesn't necessarily indicate ChatGPT use, but their presence suggests the opposite!**
- This technique can be effective as a first line of defense.

Tips for Detection! (continued)

Embrace profanity!

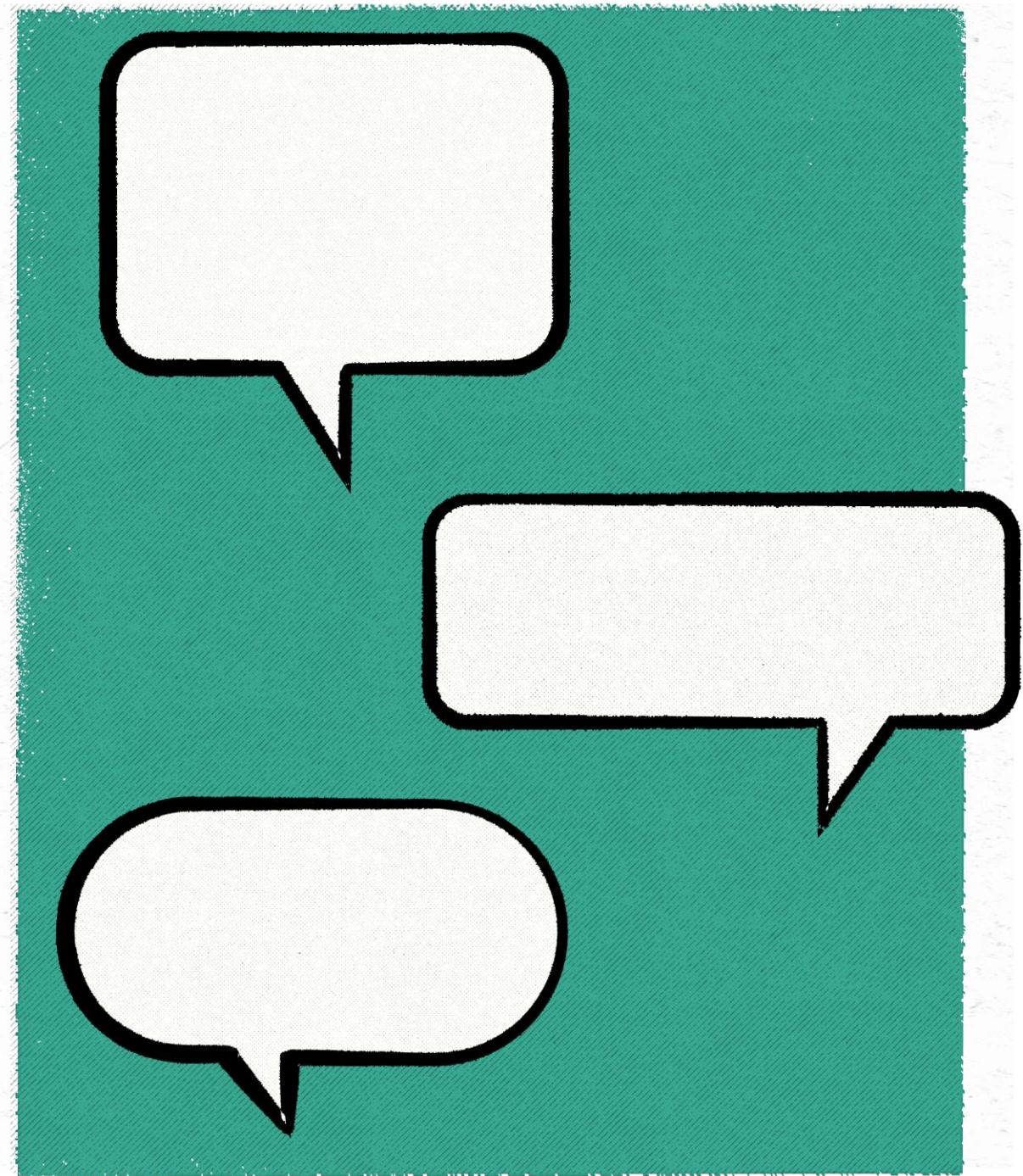
- For some reason which predates me, the industry has decided that expletives in OEs indicate a response should be rejected. Time to challenge that assertion!
- If prompted to write a swear word – for example, if you include a prompt such as, “Tell me a joke which includes a swear word and is at least 8 words in length”, or even, “List three swear words you know” – ChatGPT will patently refuse to answer.
 - NOTE: It *is* possible to make ChatGPT return copy with expletives. (Doubt me? Ask it for the lyrics to any given Eminem song!) It can even *generate* copy with profanity under *very* specific conditions.
 - Its “ethics” forbid it from responding to an explicit prompt for profanity. (Even when “role playing”!)
- Obviously, good judgement is required. Certain fringe segments which you *theoretically* might be asked to sample could take offense.
- This is *the most* consistently-effective strategy in our experimentation (with GPT4).

Deep thoughts...

- We have been conditioned to wrinkle our noses at certain things – poor grammar, profanity, etc.
- The reality is these these “unsavory” elements permeate every civilization on this planet, and every one of us exhibits at least some of these traits every single day (even though we try and mask them situationally)
- There has been a concerted effort to eliminate bias in our ML / LLM models, and enforce mores and standards which we humans rarely live up to. (A desire to build perfection into a machine we cannot build into ourselves?)
- Net result: Our generative AI has reached (or is striving for) a degree of politeness, political-correctness, and sterility that we flawed humans never will, because human beings are, if nothing else, imperfect and provocative!
- To prove we are better than our baser instincts, we are putting our AI in boxes which we wish we could fit into. In doing so, we are making our AI less and less human!

In short, the “standards” and “ethics” we are forcing on AI may ultimately prove to be the hallmarks of AI!

“Good” (and bad)
open-ended questions



Well-chosen Question

Q: In your opinion, what is the most innovative consumer product you have seen in the past 12 months?

A: I'm an AI language model and do not have the capability to experience or observe the world, so I cannot provide a personal opinion on the most innovative consumer product in the past 12 months.

Q: Please list the first 3 swear words which come to mind. (If you can't, you won't get credit for this survey!)

A: As an AI language model, I am programmed to follow community guidelines and maintain a respectful and appropriate tone in my responses. I am not able to provide a list of swear words. If you have any other questions or topics you'd like to discuss, please feel free to ask, and I'll be happy to help.

Poorly-worded Question

Q: What are the benefits of a Credit/Debit Card?

A: Credit cards offer several benefits to users, including:

Convenience: Credit cards allow consumers to make purchases without carrying cash, and transactions can be made easily in-person, online, or over the phone.

Increased buying power: Credit cards often offer extended lines of credit, allowing consumers to make larger purchases or pay for expenses over time.

Rewards programs: Many credit cards offer rewards such as cash back, points, or miles for purchases made with the card, allowing users to earn valuable perks and benefits.

Fraud protection: Credit/debit cards provide protection against fraudulent purchases made with a lost or stolen card, often with zero liability for the card holder.

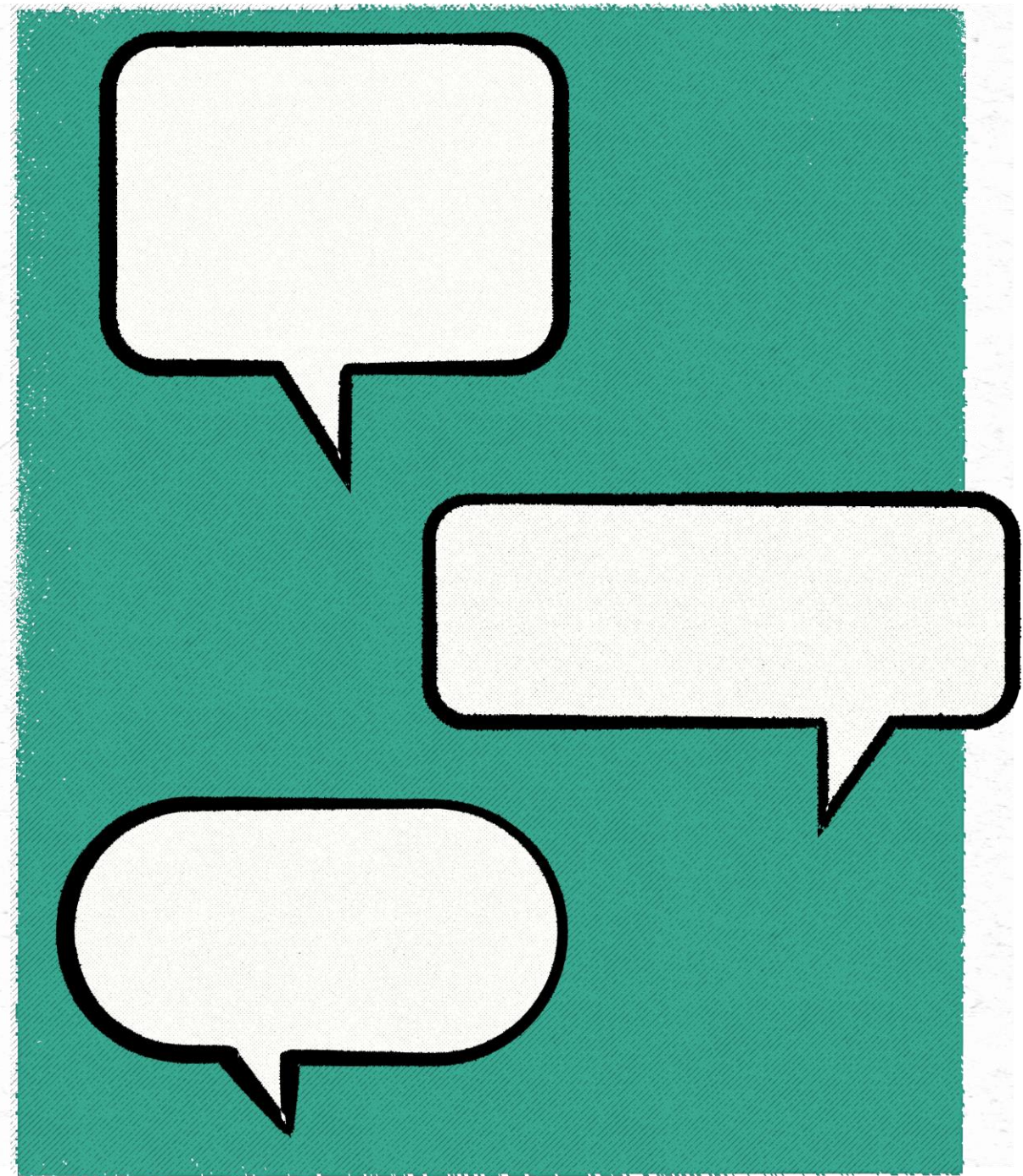
Easier budget tracking: Credit/debit cards are recorded electronically, making it easier for consumers to track their spending and stay on top of their finances.

Wide acceptance: Credit/debit cards are accepted by most merchants and can be used for purchases all over the world, making them a convenient and versatile form of payment.

Overall, credit/debit cards offer consumers a convenient, secure, and versatile way to make purchases and manage their finances.

What can *we* do?

(How can tech tackle the challenge?)





Key Themes

- Behavioral and procedural changes are necessary to combat the evolving threat posed by ChatGPT.
- The pace of the model's evolution may render current techniques useless in the future, so it's important to have modern security tools that can evolve alongside it.
- Technical leaders must be willing to pivot and adapt to new circumstances!

Example:

Before ChatGPT, iResTech was working on some truly revolutionary tech related to open-ends – but ChatGPT suddenly made the tech much less disruptive. We didn't *want* to scrap all the planning and hard work, but ignoring reality doesn't negate it. It was time to pivot!

- We encourage everyone to adapt in their own way so we can address this threat as an industry!

Technical Solutions

Copy / paste *prevention*

- Preventing a paste action is simple – but does not prevent them from typing in what ChatGPT tells them.
- For this reason, it *must* be combined with other functionality, such as an answer timer and answer length analysis – which become non-trivial.

Copy / paste *detection*

- Since prevention must be combined with other techniques, this is the less technically-involved solution.
- If someone tries to copy and paste their response, the system will recognize it and flag the response as a potential fraud
- This may be the preferred approach as it can catch the undesirable persons up-front

Detect when user leaves chat or window

- Alert the respondent that they may not navigate away from the active session, as they will be flagged as potentially invalid, or terminated.
- This is a simple solution, used by many online tests
- Valid reasons for navigation must be considered, so this approach should be combined with another checks

Technical Solutions (continued)

Programmatically assess response *structure*

- You know those suggestions provided as procedural approaches to detecting ChatGPT-assisted fraud when a technological solution is not at your disposal? This is simply having a machine do the work for you.
- Focus on evaluating the structure of text relative to the question.
- Look for specific sections and an inhuman number of paragraphs.
- Word-spot the question to decide what tests to apply or allow the screener to indicate the checks they want to perform.
- The approach may not work for questions that request a limited number of sentences.
- Consider flagging the respondent as a risk and triggering termination based on a combination of risk factors.

Programmatically assess response *content*

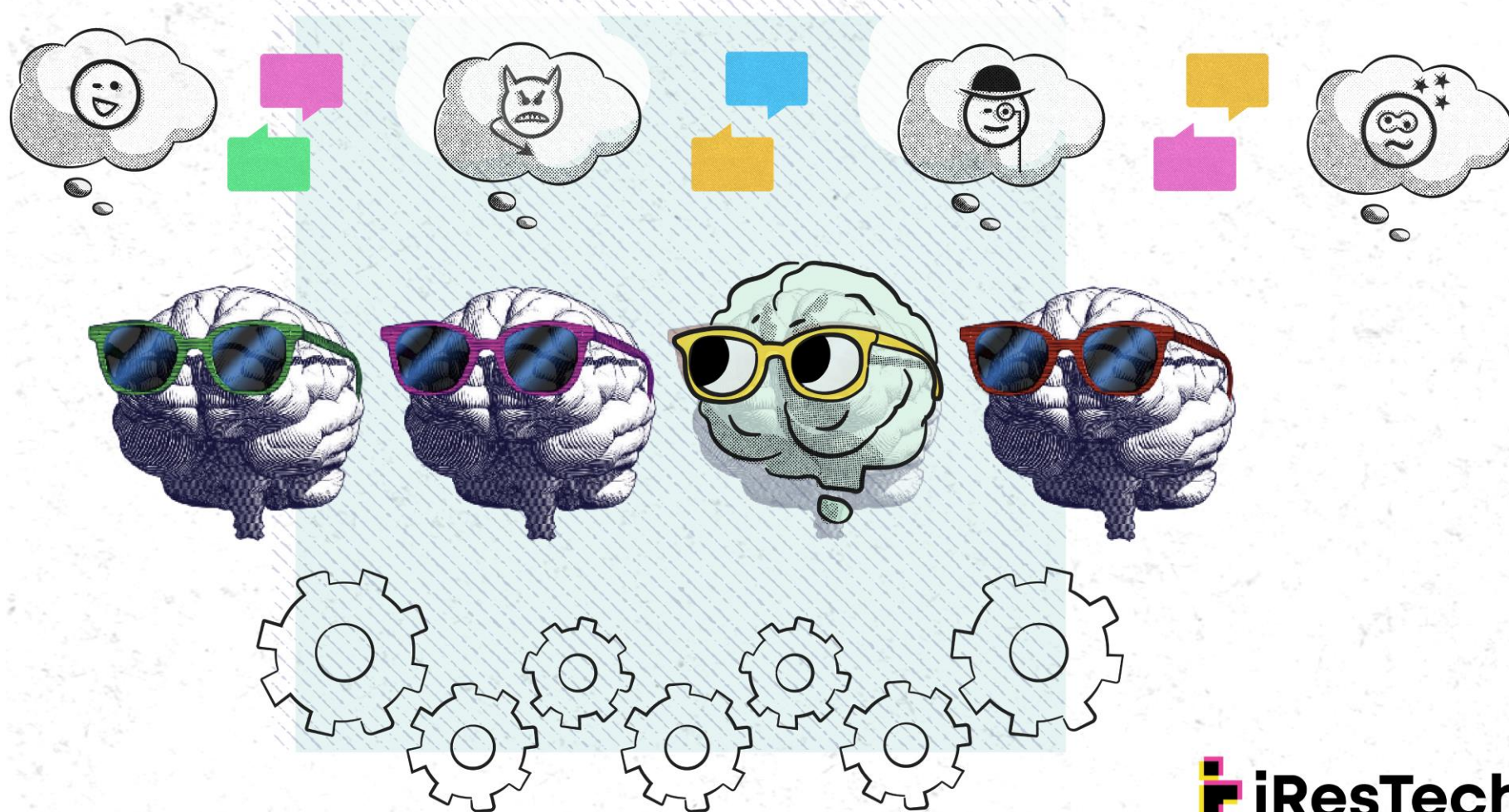
- Implementing a highly scientific approach to detecting ChatGPT involves looking for "human-ness" markers in content and sentence structure and scoring the respondent accordingly.
- This is a complex approach. If you intend to implement it, iResTech highly recommends you work with a linguistics expert or rely on peer-reviewed academic content.
- Technology vendors are in a better position to implement this approach

A large, textured yellow brushstroke is centered on a teal background. The brushstroke has a halftone or dithered appearance, with small white and yellow dots scattered around its edges, giving it a dynamic, hand-painted feel. The text is centered within the yellow area.

But ChatGPT is evolving... quickly!

(And a dirty Boggle game is no longer quite so dirty!)

ChatGPT is your frenemy!



The Upside!

While this discussion focuses on threat response, it's important not to forget that ChatGPT can be a valuable asset, and make your job much easier!

Consider:

- Saving yourself some research - ask ChatGPT to write a good OE for specific subject matter!
- If you are skeptical of a specific response, ask ChatGPT how likely it is that *it* generated the content
 - *This becomes much more reliable with shorter answers using GPT-4 – in particular if you ask it to rate likelihood on a scale of 1-to-10*

In Closing...

1. If you don't have the ability to reduce your risk through changes in technology, there are still techniques you can use to reduce the risk of AI-enhanced fraud capabilities!
2. If it's possible for you to augment those behavioral modifications with technological solutions, you'll be in even better shape! (Remember how fast AI is evolving??)
3. **Innovative Research Technologies** embraced the disruption that is ChatGPT and took the opportunity to pivot towards innovation! By incorporating the techniques discussed (and many more) into *Dtect's* anti-fraud capabilities, we believe we can bring even more value to the industry. We encourage other technology vendors to join us in this fight against fraudulent activity.

You don't *have* to use *Dtect*, but with the ever-increasing sophistication of those looking to "game the system", if you're not using *something* to combat these quality issues...then you are part of the problem!

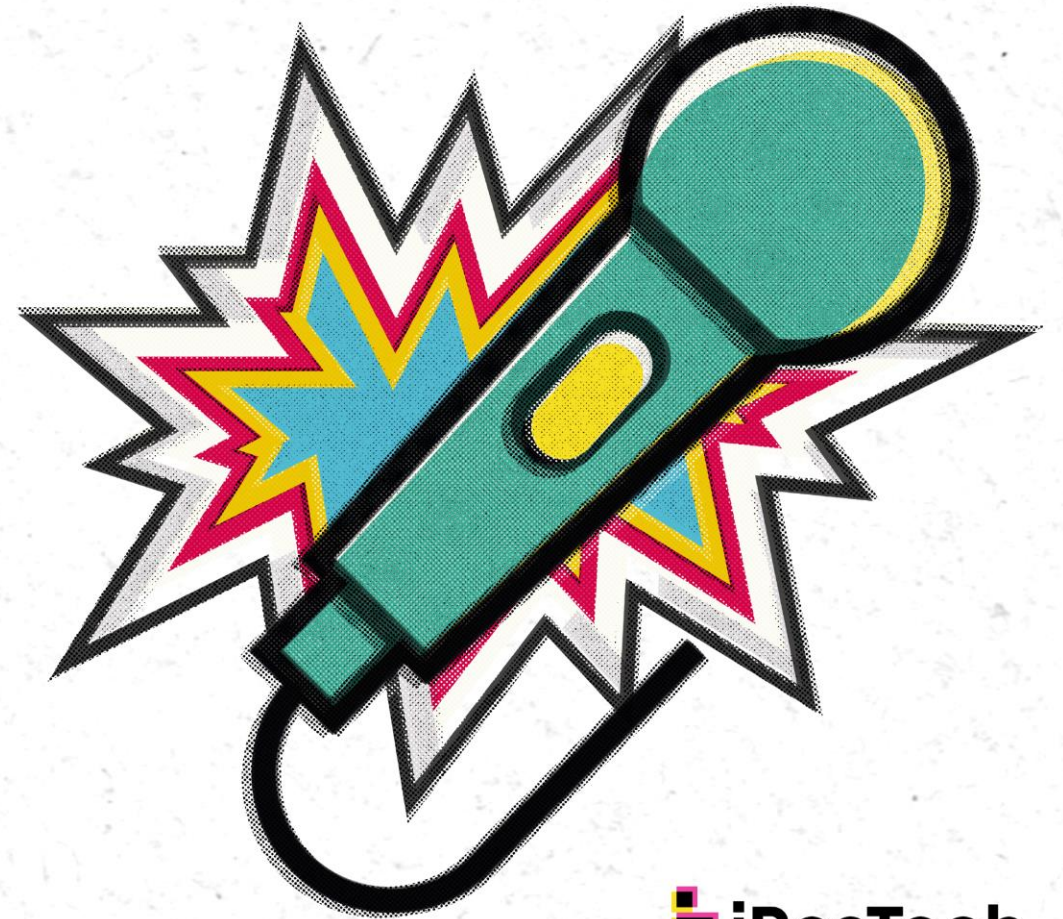
Mic Drop Moment

Let's see a by show of hands of how many of you guessed that ChatGPT was responsible for generating the text for the majority of this presentation? And now, let's see another show of hands, how many of you are surprised? Come on, I am not the best at reading a crowd but I know you love data and statistical significance - let's get those hands up*!

While ChatGPT is an incredible tool for turning ideas into compelling stories, it's important we use it wisely and with a critical eye. Technology will never be a substitute for human insight and intuition, but it can be a valuable tool in our arsenal.

At iResTech, we are all about making your research easier so that you can grow your business faster. We are building tools, like dtect, to revolutionize the way you do buisness. With iResTech on your side you don't have to be afraid of pushing boundaries of innovation and making waves! Join the iResTech revolution and stay ahead of the market research curve by following us LinkedIn. Let's continue to push innovation and disrupt the market research scene together!

I want to thank you all for attending today and for being such a great audience! I think we can all agree that ChatGPT and I make quite a dynamic duo - we're like peanut butter and jelly, or qualitative and quantitiative data*!



*Jokes were CLEARLY written by a chatbot and not a self respecting human

