OpenAI | Commonwealth of Pennsylvania

MARCH 2025

Lessons from Pennsylvania's Generative AI Pilot with ChatGPT

Report Summary

Overview

Throughout the Commonwealth's ChatGPT Pilot program, over 175 employees accessed ChatGPT Enterprise, with **136 providing direct feedback to the Commonwealth on their experience**. These employees volunteered to explore this new technology and ranged in age, job type, seniority, experience with AI technology, and included participation from union covered employees. Notably, **48% of participants had never used ChatGPT** before this pilot program. Overall, **85% of participants had a positive experience** using ChatGPT. While this population is not representative of the Commonwealth's workforce at large, participants broadly reported that they experienced efficiency gains and improvements to their work product.

ChatGPT users in this pilot varied in terms of the extent to which they relied on ChatGPT, however**, pilot superusers did not have one consistent profile.** Successful use was a function of how well employees could navigate common adoption challenges such as inaccuracy, habit forming, lack of learning time, learning curve, and concerns about data & privacy.

Driving Government Impact:

The most frequent use cases of ChatGPT stood out because Commonwealth employees found ChatGPT most useful and effective in completing work that is unique to the public sector.

Many government teams are immersed in complex policies, procedure, routines, stakeholders, and processes. ChatGPT provided a tool to tackle these challenges. **ChatGPT particularly enabled participants to succeed in government work as:**

Innovation Engines: Empowering participants to generate ideas, explore solutions, and accelerate problem-solving despite various constraints.

Bureaucracy Hackers: Creating documentation, demystifying polices, and navigating processes that are primarily manual, outdated, or text-based.

Strategic Communicators: Improving the clarity, consistency, and impact of communications between government stakeholders operating across a wide domain of topics.

The Commonwealth is grateful to Carnegie Mellon University's Block Center for Technology and Society for their consultative support and expertise throughout the pilot engagement

Top 5 Findings

01 Most employees described their experience using ChatGPT at work as “very positive.”

02 Users estimated saving 95 minutes per day with ChatGPT.

03 ChatGPT is not a substitute for the nuance and experience employees bring to their work.

04 There was no singular profile of successful ChatGPT user, a range of employees found both challenges and opportunities with adopting ChatGPT.

05 The most frequent uses of ChatGPT enabled employees to drive impact amid common government challenges.

Contents

Overview

Background

In September 2023, Governor Shapiro **signed Executive Order EO 2023-19 –Expanding and Governing the Use of Generative Artificial Intelligence Technologies Within the Commonwealth of Pennsylvania**. This executive order laid the foundation for the Commonwealth to proactively explore Generative AI tools through targeted programs such as the Commonwealth's pilot with OpenAI.

We can't ignore new technology — we have to educate ourselves and be proactive to minimize the risks and maximize the benefits of innovation and that's the approach my Administration is taking here in Pennsylvania. We have the potential to become a leader in responsible, ethical use of generative artificial intelligence."

- Governor Josh Shapiro

EO 2023-19 – Expanding and Governing the Use of Generative Artificial Intelligence Technologies Within the Commonwealth of Pennsylvania

"Agencies shall weigh the design, development, procurement, and deployment of Generative AI technology based on the following core values. The Office of Administration (OA) shall proactively consider these core values in all its activities, policies, and processes that leverage or intersect with technologies that enable automated decision-making."

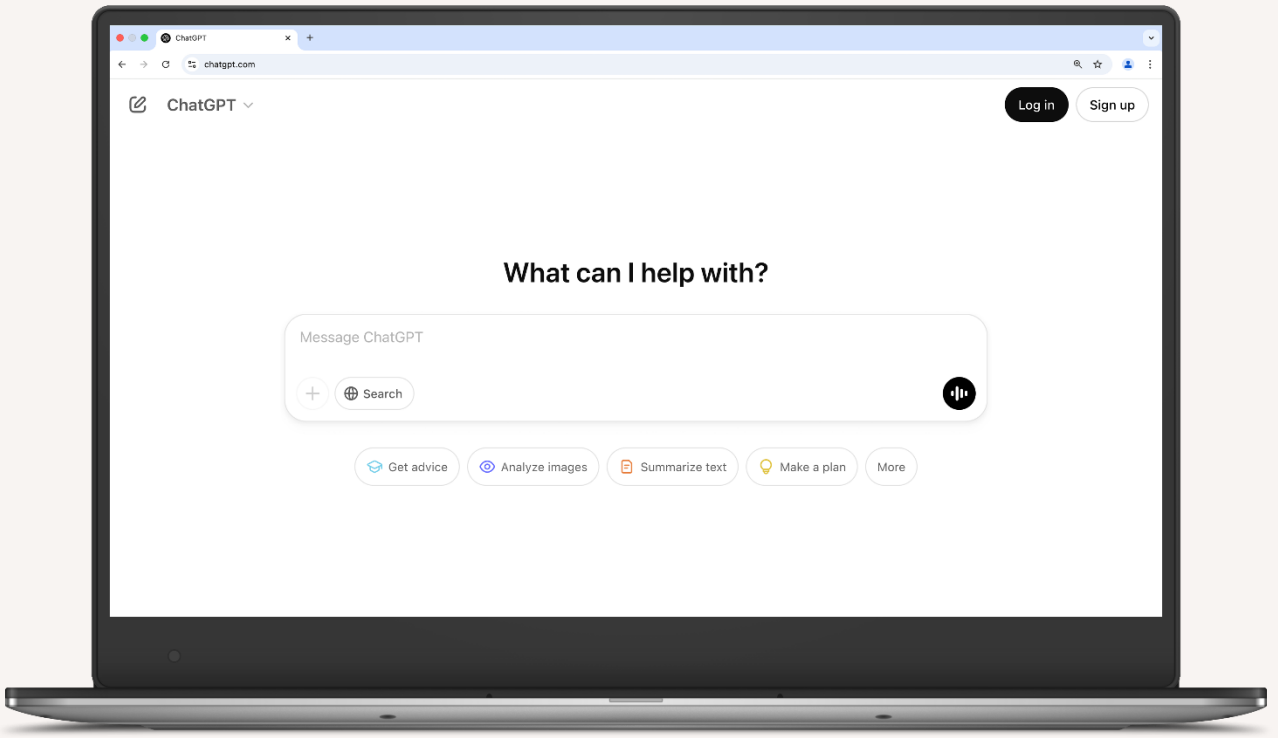
Values:

- Accuracy
- Adaptability
- Employee Empowerment
- Equity and Fairness
- Innovation
- Mission Aligned
- Privacy
- Proportionality
- Safety and Security
- Transparency

Background

In March 2025, the Commonwealth concluded a yearlong pilot exploring Generative AI (GenAI) technology.

The Commonwealth equipped 175 employees to use ChatGPT Enterprise, a leading GenAI tool, and studied how they used it in day-to-day work.



OVERVIEW

Pilot Scope

12 months of research and observation	5 cohorts of employees signed up for the pilot	175 employees equipped with ChatGPT Enterprise
48% participants who had never used ChatGPT before	599 surveys submitted by pilot participants	14 state government agencies represented in the pilot

OVERVIEW

Timeline

Of participants who provided feedback, five employee cohorts began the pilot in two phases from March 2024 – March 2025. Along the way, each cohort participated in bimonthly surveys and regular focus groups.

PHASE 1

1 cohort

56 employees

5 surveys

March – May 2024

PHASE 2

4 cohorts

80 employees

6 surveys

August 2024 – February 2025

END

Exit survey

March 2025

Please Note: Not every pilot participant provided survey-based feedback. The above represents the timeline for the 136 employees who provided direct feedback throughout the pilot.

OVERVIEW

Pilot Resources & Trainings

Employees were provided a range of trainings and resources throughout the pilot. These opportunities helped onboard participants to ChatGPT and allowed the Commonwealth to adapt to participants' unique training needs. Most participants engaged in a combination of:

Onboarding

- *Foundational ChatGPT use guidelines*
- *Live onboarding sessions*

OpenAI led trainings

- *ChatGPT enablement sessions*
- *Prompt engineering*
- *Live workshops*

Ongoing support

- *Weekly "office hours"*
- *Small group problem scoping sessions*
- *Targeted support to highlighted use cases*

General AI trainings

- *Foundational AI 101 sessions on request*
- *Monthly AI presentations to provide broader AI context*

About the Study

Recruitment

Phase 1: In January 2024, all employees in PA's Office of Administration (OA) were invited to participate in the pilot by email. A single cohort of employees with a range of job types was manually selected to begin the pilot.

Phase 2: All remaining volunteers from OA were added to the pilot on a rolling basis. Employees from the several other agencies were also invited to participate as capacity allowed in four separate cohorts.

Pilot participants: Pilot participants tended to be 35-54 years old and have a postgrad degree. Beyond these similarities, participants ranged widely in terms of job function, job tenure, prior GenAI use, and expectations for the pilot.

Methods

Pilot participants reported on their experience using ChatGPT at work through three channels:

Surveys: After beginning the pilot, participants received surveys every other week for ~8 weeks, plus an exit survey at the pilot's end.

Focus groups: All pilot participants were invited to weekly focus groups conducted by the research team.

Interviews: Select participants spoke one-on-one with the research team during Phase 1.

Please Note: Pilot participants were not a representative sample of Commonwealth employees. Their experiences provide insights into understanding generative AI tools but are not predictive for all Commonwealth employees

Findings

Research Findings

Throughout the pilot, the research team collected observations under five key areas:

- ChatGPT Use Cases
- User Experience
- User Types
- ChatGPT Challenges
- Barriers to Use

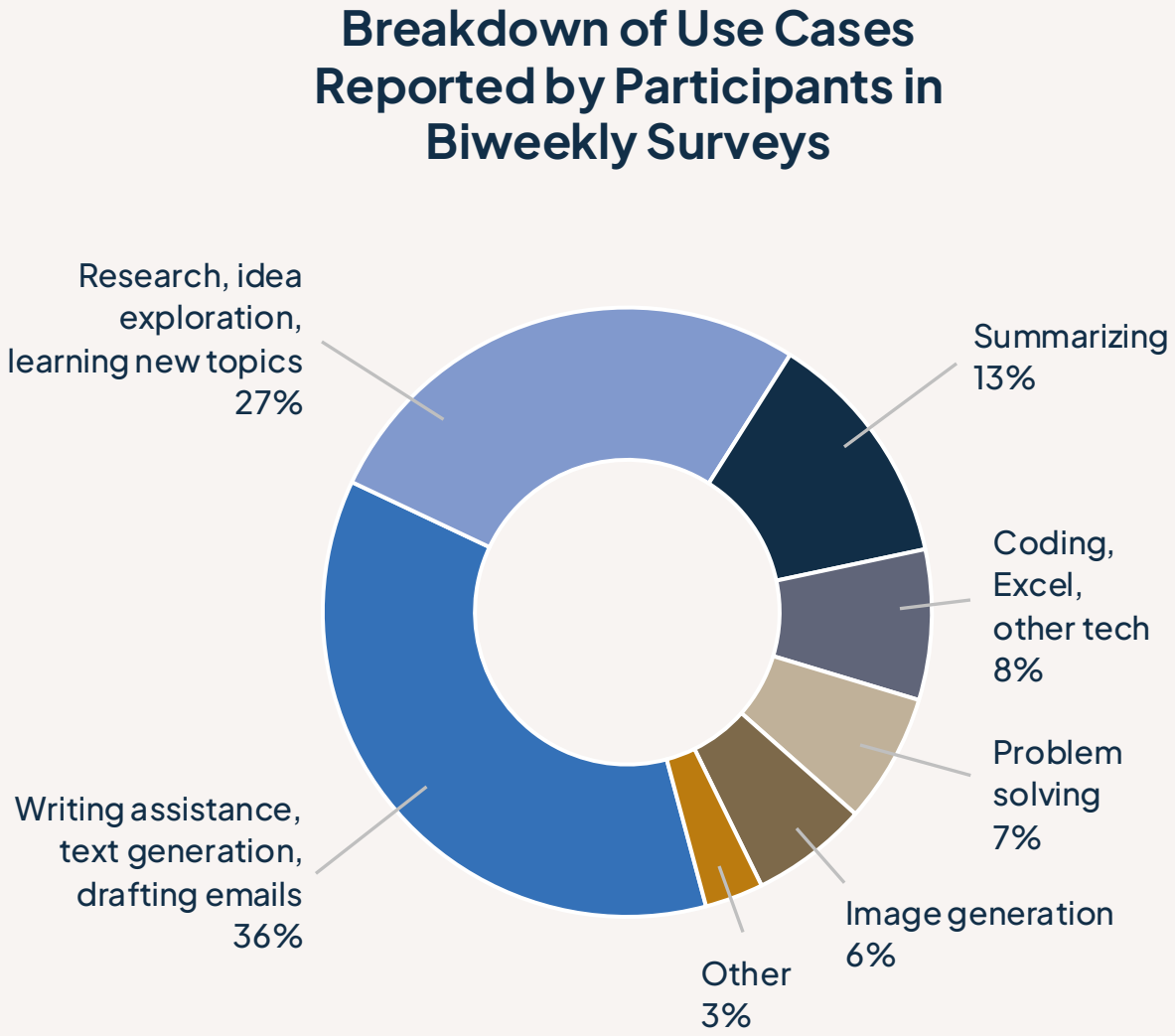
ChatGPT Use Cases

What did pilot participants use ChatGPT to do?

USE CASES

Most frequent uses

Pilot participants turned to ChatGPT when tackling a wide range of work tasks. The most frequently reported often involved navigating text-based information such as research, brainstorming, writing, and summarizing.



Most Frequent Uses Address Government Challenges

Government Challenges

The most frequent use cases of ChatGPT stood out as being areas that are particularly relevant to **navigating the features of public sector work.**

Many government teams are immersed in complex procedure, routines, and processes with feedback from a variety of stakeholders. ChatGPT provided a tool to tackle these challenges.

ChatGPT enabled employees to succeed as:

Innovation Engines: Empowering participants to generate ideas, explore solutions, and accelerate problem-solving despite various constraints.

Bureaucracy Hackers: Creating documentation, demystifying polices, and navigating processes that are primarily manual, outdated, or text-based.

Strategic Communicators: Improving the clarity, consistency, and impact of communications between government stakeholders operating across a wide domain of topics.

USE CASES

Innovation Engines: Brainstorming

Problem:

A Commonwealth communications professional had a tight turnaround for a new public health marketing campaign that discussed a new commonwealth program.

Solution:

The communications professional used ChatGPT to quickly brainstorm ideas and put together initial drafts that helped storyboard what the whole campaign could look like, saving this employee hours of planning work.

USE CASES

Innovation Engines: Brainstorming

> *The tool greatly increased the speed in which I was able to "see" what something would look like and helped to rip through rough ideas quickly.* ***The tool was not suitable to create a finished product... but it was extremely useful in the brainstorm process*** *to give a visual direction and a solid reference point to build on.*

USE CASES

Innovation Engines: Problem Solving

Problem:

Ensuring Commonwealth agencies are creating and implementing accessibility plans is essential but challenging work given the amount of managed content.

Solution:

ChatGPT helped provide roadmap outlines and templates that could be shared and scaled to agencies to jumpstart accessibility auditing. This saved time and created consistent documentation and practices across teams.

USE CASES

Innovation Engines: Problem Solving

> *What the prompt provided was a bare bones starting point that we could use to expand upon and create a simple roadmap that is easy to use and accessible to everyone...**Just having GPT create the starter that we worked from saved a lot of time up front, even if there was a lot of human time invested once we got our starter**, same as with the first example above.*

USE CASES

Bureaucracy Hackers: Research & Summarizing

Problem:

The terms and conditions for IT contracts required vendors to comply with the Commonwealth's 93 IT policies, which spanned over 500 pages. This was arduous on employees and external parties.

Solution:

Pilot participants used ChatGPT to aid in consolidating 93 policies into 34. These 34 policies continue to be updated and rewritten to align with common compliance frameworks that most vendors follow already.

***Learn more about this work:** Getting Stuff Done: Shapiro Administration Modernizes IT Procurement | Office of Administration | Commonwealth of Pennsylvania*

USE CASES

Bureaucracy Hackers: Summarizing

> *[We are] trying to align our policies better to the guidance provided by NIST, so I asked chatGPT to draft a policy for privileged user access management based on NIST 800-53.* ***I then compared the output to our current policy. For the most part they aligned fairly well, but there were some things in the chatGPT policy that we did not have in ours and we ended up incorporating some of those***

USE CASES

Bureaucracy Hackers: Summarizing

Problem:

HR classification is a complicated and tedious task. The HR team faces challenges in reviewing and navigating thousands of job classifications and position descriptions.

Solution:

An HR professional created a custom GPT capable of evaluating 3,600 position descriptions when presented a job classification question.

USE CASES

Bureaucracy Hackers: Summarizing

> *For me**, using GenAI is about getting closer to a solution faster**...I can use it to sift through lots of information to make suggestions and ask it why if I disagree. From there, I can use my own expertise and consult with others to validate and make a decision.*

USE CASES

Strategic Communicators: Writing Assistance

Problem:

A technical IT professional often needs to communicate critical information or project requirements with non-technical stakeholders associated with the work.

Solution:

Using ChatGPT to quickly review communications and suggest language that will ensure technical jargon is clear and more accessible to their audience.

USE CASES

Strategic Communicators: Writing Assistance

> *I'm not good at editing my own work...and having ChatGPT as a partner for this kind of task is a godsend!* ***It made the final product more fitting for the target audience****, and in much less time than it would have taken me.*

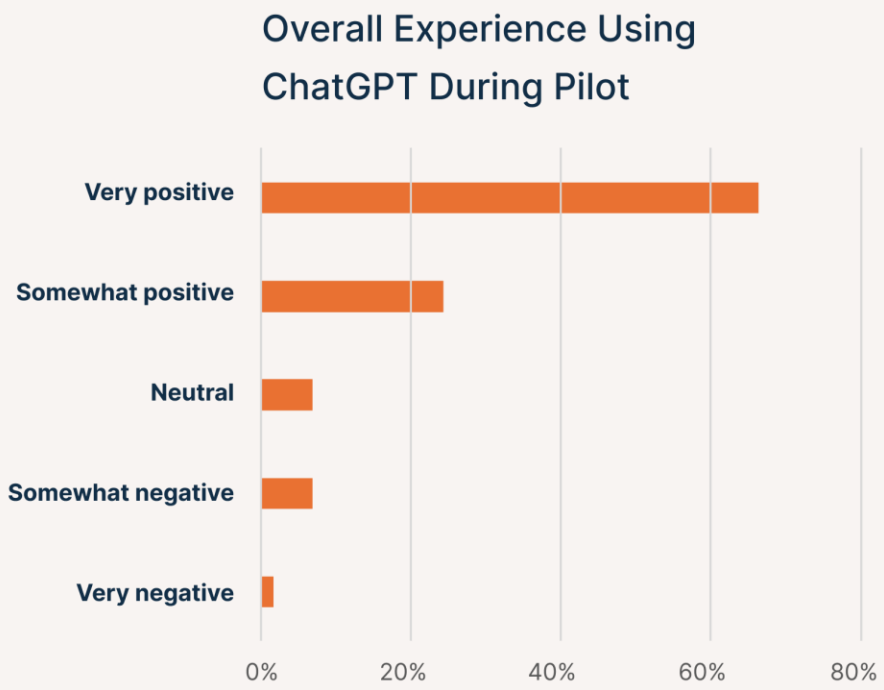
User Experience

How did pilot participants feel about using ChatGPT overall?

USER SENTIMENT

Employees Enjoyed Using ChatGPT

Generative AI is an emerging technology that roughly half of participants were unfamiliar with, yet over 85% of participants reported a “somewhat positive” or “very positive” experience using ChatGPT.



*“It's been a **wonderful** project. I regret that it is coming to an end...”*

*“ChatGPT is a **valuable tool** that has made nearly every part of my job easier, faster, and more efficient.”*

*“Participating in the pilot program and using ChatGPT in my work has been an **incredibly positive** experience for me.”*

USER SENTIMENT

Benefits and Drawbacks

While participants shared positive impressions overall, their experience was not uniform. Participants reported a wide range of benefits, drawbacks, and tradeoffs with ChatGPT.

*"ChatGPT is such a **useful and time saving tool** and I have had only minor hiccups in its responses so the overall experience has been extremely positive."*

*"ChatGPT felt like having an **incredibly reliable and resourceful** counterpart in the workplace."*

*"I have come to see both the **usefulness and the potential for misuse** of the technology."*

*"I see the benefit [of] having and utilizing Chat GPT. However, **I don't see how it can fit into my daily work tasks** on a regular basis."*

*"[ChatGPT] still gave **wonky results** that I could more easily get and verify on my own with less bias."*

USER SENTIMENT

Employees' Most Helpful Use Cases

When discussing their most helpful use cases employees reported ChatGPT:*

- Improved their final work product
- Reduced task complexity
- Required minimal review
- Cut working time to under 15 minutes in 58% of cases.

So what?

While ChatGPT can be used for a wide range of use cases, employees benefited from seeking out and investing in uses that were most helpful to their work. Once identified, these most helpful use cases created higher than average gains.

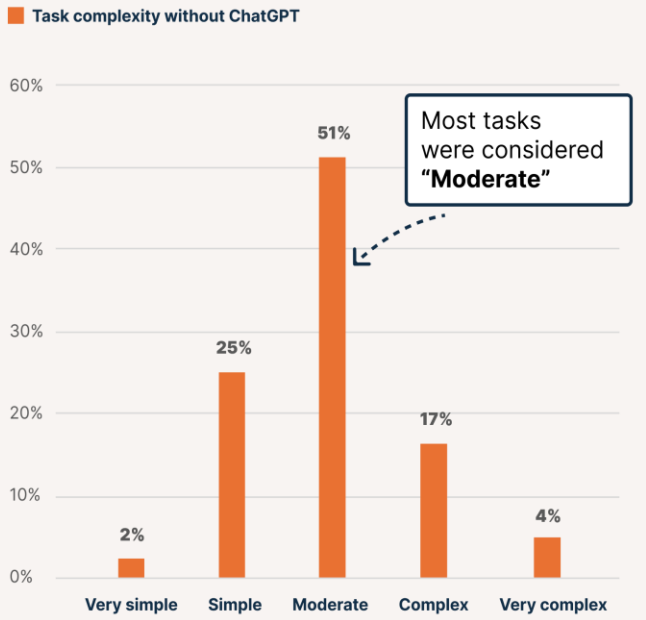
Identifying high-value use cases is key to driving impact.

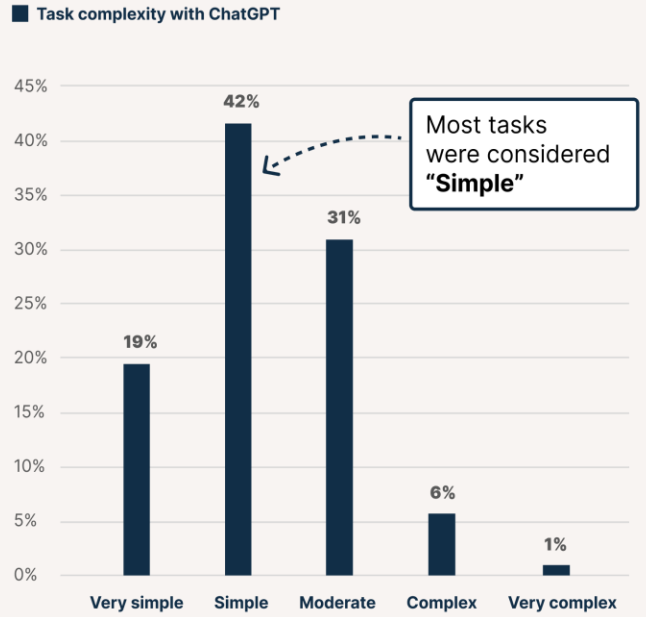
*Reported by participants in biweekly surveys

USER SENTIMENT

Employees' Most Helpful Use Cases

When employees discussed the ChatGPT uses that were most helpful to them, the average complexity of those tasks **shifted from being considered moderately complex to simple.** In addition to saving time and improving the quality of outputs, ChatGPT helped make these tasks less complex and more manageable for employees.





USER SENTIMENT

Efficiency Gains

When employees were able to effectively incorporate ChatGPT into their workflows they found noticeable efficiency gains.

35

average daily minutes **spent** using ChatGPT*

95

average daily minutes **saved** using ChatGPT*

*"[ChatGPT] has been a great help with my workflow, and **it has saved me so much time** helping with tasks that normally would take hours to do."*

*Estimates reported by participants in exit survey

USER SENTIMENT

Human Nuance Is Essential

ChatGPT reduced, but did not eliminate, human input into final work products. Pilot participants reported having to adjust ChatGPT's output throughout their use, including instances where ChatGPT was being most or least helpful.

The most common edits required were to add additional nuance from the employees' experience, **underscoring the importance of human judgment and expertise in AI-assisted work**.

So what?

ChatGPT is not a substitute for the nuance and experience employees bring to their work

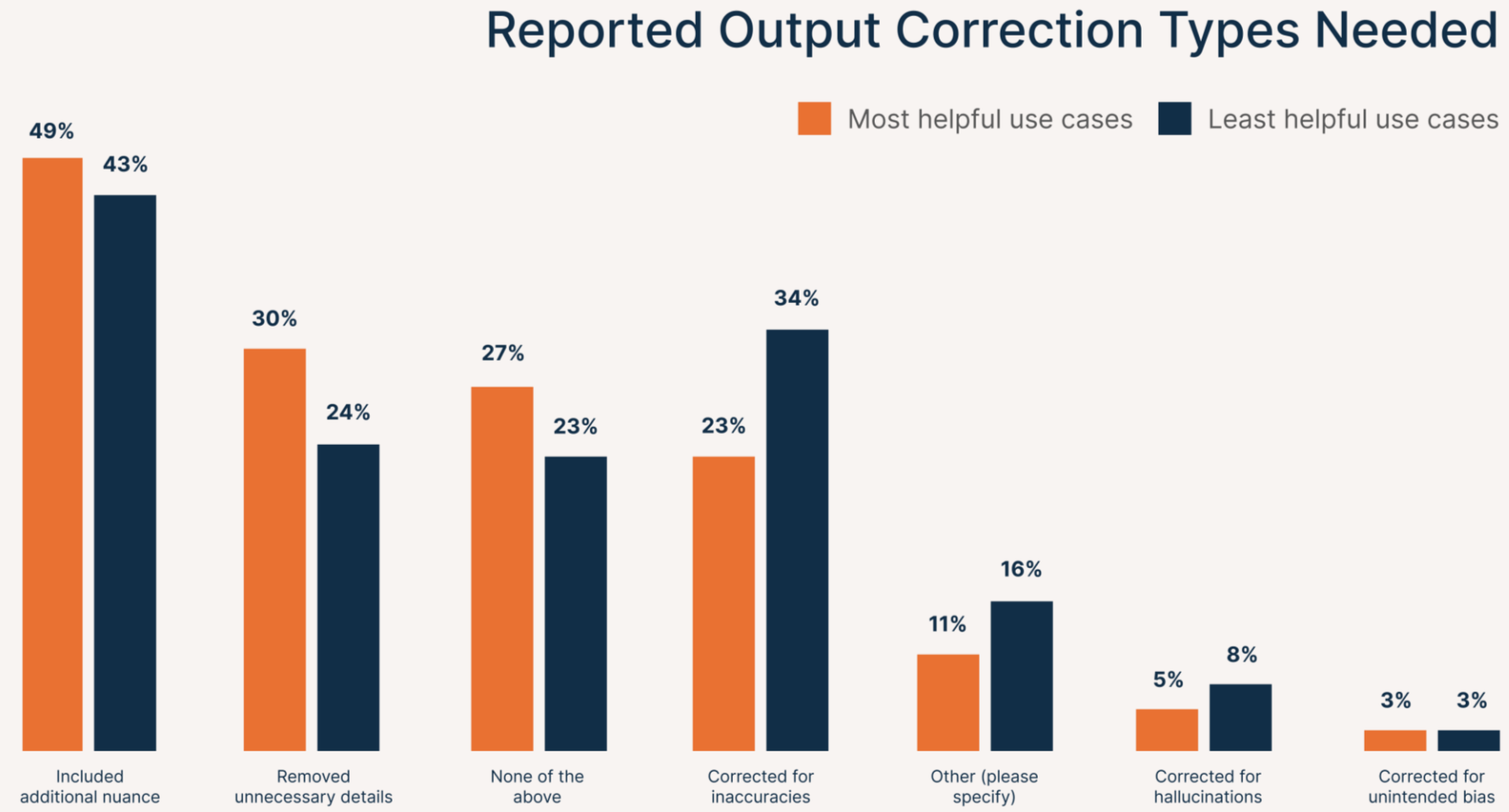
Successfully using ChatGPT at work requires a human-in-the-loop approach.

*Reported by participants in biweekly surveys

USER SENTIMENT

Human Nuance Is Essential

Of all types of prompt corrections, employees most commonly said they **needed to add their additional nuance.**



User Types

Who used - and didn't use – ChatGPT during the pilot?

USER TYPES

ChatGPT Use Patterns

Over the course of the pilot, participants split into four general user archetypes.*
The archetypes didn't map to particular groups of COPA employees; participants with varied roles, education levels, and demographics fell into each.

Sporadic, General Users

- *Used ChatGPT as needed, not for any particular task*
- *48% of participants*

Specific, Consistent Users

- *Used ChatGPT for one or two particular work tasks*
- *18% of participants*

Superusers

- *Used ChatGPT consistently for many different tasks*
- *30% of participants*

Nonusers

- *Didn't use or stopped using ChatGPT*
- *4% of participants*

* Classifications are based on participant responses to exit survey. The share of Nonusers is likely an underestimate, as employees who stopped using ChatGPT may have ignored survey requests.

USER TYPES

"It's a good resource to have more than not...I do still use it on occasion. I like having it. It's much better than googling. ***Sometimes it's like having an 'expert' colleague nearby to explore ideas****."*

-Sporadic, General User

*"****I am finding it a challenge to, first of all, think to use GPT in my work*** *- and second of all, to determine HOW to use it in my work."*

-Nonuser

"I have mainly used ChatGPT throughout the pilot for research purposes and to learn more about certain topics. ***This purpose has remained consistent throughout the pilot."***

-Specific, Consistent User

"I use it with purpose in many areas of my work. I use ChatGPT to help me write better, to analyze data and provide results, as a search engine to obtain information on multiple topics, to compare old and new documents and identify differences, etc. ***ChatGPT has saved me more and more time as I've expanded the ways and reasons I use it."***

-Superuser

ChatGPT Challenges

Where did ChatGPT struggle to meet employees' expectations?

ChatGPT Functionality Challenges

Pilot participants reported many ChatGPT wins with writing, brainstorming, and summarizing. **However, participants reported these same use cases most frequently when asked about times when ChatGPT didn't meet expectations.**

So what?

GenAI isn't perfect. It makes mistakes. Employees must be responsible for their AI-assisted work product. AI training should include how to anticipate and check for errors in GenAI outputs.

CHALLENGES

Particularly Challenging Cases

Please Note: ChatGPT updates and products that addressed some of these concerns were not readily available throughout the entire pilot. Specifically, no Deep Research usage is captured in the pilot results.

ChatGPT struggled in a few special situations, including generating images, citing sources, and extracting text from PDFs. An all-purpose GenAI tool like ChatGPT may not be the best bet for these use cases. It's worth investigating specialized software.

*"I wanted to play with the **image creation** so decided to try and make an avatar of myself. Every time I tried to adjust the photo, it would add a septum ring to the image. At one point, it changed my gender. When I asked it to remove the piercing for the 3rd time, it [removed] a good portion of the lips."*

*"And the only adjustment I had to my prompt...was requiring it to **cite the work it was using**...we had to go through a few renditions of that...because it would produce some hallucinations with links not being accurate or being made up...Which means for the manual edits that I had to do...It was a matter of fact checking."*

*"The **text extracted from the PDF** includes a lot of formatting artifacts, such as excessive whitespace and broken text. This might make it difficult to number the sentences clearly and sequentially. I ended up just doing the task manually."*

Barriers to Use

Why wasn't ChatGPT used as widely as expected?

BARRIERS TO USE

Adoption Barriers

In some ways, ChatGPT use was lower than expected:

- At the start of the pilot, participants predicted that they would use ChatGPT for 6 tasks on average but went on to report half as many use cases in biweekly surveys.
- At the end of the pilot, 4% of participants reported not using ChatGPT at all.

What happened? In focus groups and surveys, participants mentioned five barriers to adoption:

1. Inaccuracy
2. Habit
3. Lack of learning time
4. ChatGPT's steep learning curve
5. Doubts about data

BARRIERS TO USE

Inaccuracy

Some employees did not trust ChatGPT to produce accurate or reliable output. Hallucinations were of particular concern.

*"I noticed that **AI had hallucinated how many years of experience as a lawyer you had to have to qualify for some of the positions.** So that's why I always double check what chatGPT gives me."*

*"Consistently, the results would indicate a case (or more) existed that was directly on point; however, **when I checked for the cases they did not exist or did not contain the law specified**. I would modify the questions to be more specific and indicate that the prior answers were incorrect. While it apologized, **when it came to legal issues, the answers were almost never correct."***

So what?

Set appropriate expectations with first-time GenAI users. Remind them that GenAI outputs should be considered a draft with imperfections, just like a draft generated by a human.

BARRIERS TO USE

Habit

Some participants who intended to use ChatGPT failed to do so. They described difficulty changing ingrained workflows or simply remembering to give ChatGPT a try.

*"**It's been a process to rewire our brains to use ChatGPT** other than google or relying on ourselves like we are used to."*

*"I am still struggling with THINKING to use ChatGPT. I've been doing my job for a long time and am comfortable with most aspects - and even when I get tripped up or think about doing research or something, **ChatGPT doesn't come to the forefront of my mind quickly**! I'm not averse to using it, just keep having to remind myself it's there!"*

So what?

A small amount of initial training goes a long way to spur further learning. Encourage GenAI-savvy employees to seek out teachable moments with their peers and uplift successful use cases from different employees.

BARRIERS TO USE

Lack of Learning Time

Some participants were too busy to figure out how to weave ChatGPT into existing workflows, even if they expected ChatGPT to save them time in the long run.

"There are many other possibilities where ChatGPT can be useful in my work, ***the trick is finding the time to inve[s]tigate them while still meeting the demands of my current workload.*** *"*

"I have had an influx of work over the last 2 weeks that has limited my time to explore how ChatGPT might help. I am caught in this situation where ChatGPT would probably be very helpful in getting through this workload, but ***I am still struggling in learning how to use ChatGPT and it takes more time than I have right now to figure out how to build the appropriate tool to assist in my work.*** *"*

So what?

Experimenting with GenAI is essential. GenAI training opportunities should be folded into regularly scheduled work activities or designated time should be created to practice these skills.

BARRIERS TO USE

A Steep Learning Curve

ChatGPT overwhelmed some participants who found it to be too complicated to learn.

"I did not have the patience to wrestle with the tool and at that point ***it just felt quicker to continue on my own.***"

I feel like I'm missing something very basic in getting started*...For example, in learning Excel, a person should know how to open a new workbook, add sheets, rows, columns, and put data into the cells, before getting to the point of putting in formulas for data analysis****. I feel like we jumped into the data analysis part before learning the very basics of how to start the workbook.***"

So what?

Learning new tools is hard, even more so when starting from square one. Give employees easy, impactful uses of GenAI to help shrink the learning curve.

BARRIERS TO USE

Privacy and Use Concerns

Participants were reluctant to use ChatGPT due to uncertainty about how it processes and stores input data. They asked for clear policies and guidelines for using ChatGPT with potentially sensitive information.

*"I'm not 100% comfortable with the limits of use**. I have some ideas, but not sure what I'm allowed to upload into the tool.** "*

*"I need more creative examples and **limits of what I am permitted to do**."*

So what?

Proactively provide employees clear "GenAI dos and don'ts" so they can use new tools with confidence.

Please Note: these concerns existed despite existing policy guidance and Commonwealth-OpenAI Terms protecting user data

Next Steps

What should happen post-pilot?

NEXT STEPS

The Post-pilot Path

The Commonwealth's ChatGPT pilot provided crucial insights into the factors that help and hinder adoption of GenAI tools. It's now time to iterate and scale these lessons to enable Commonwealth employees to successfully leverage Generative AI in their work.

Create pathways

Drive Value

Set up for Success

Investigate Impact

Sharing PA's AI playbook

NEXT STEPS

Create Pathways

There is an ever-evolving landscape of Generative AI features, tools, and opportunities. The Commonwealth looks to continue to create pathways for employees to use generative AI tools that are safe, effective, and additive to work.

To get started, the Commonwealth should continue to:

- Provide Generative AI tools that meet the needs of users so that they can best leverage Generative AI broadly and in domain or function specific ways.
- Create flexibility in tooling options over time that account for a dynamic generative AI ecosystem.

NEXT STEPS

Drive Value

To meet the challenges in front of them, public servants need to be equipped with the best tools available to solve shared problems. The Commonwealth should continue to highlight and enable teams in areas where there is clear value from using generative AI tools. **Users should**:

- Lean into opportunities to use AI as innovators, bureaucracy hackers, and strategic communicators.
- Promote ownership of generative AI enabled work products. Results are the best when users engage with the quality, accuracy, and appropriateness of the outputs.
- Explore new use cases that enable teams to collaborate on shared goals and challenges.

NEXT STEPS

Set Up for Success

Some pilot participants struggled to get started with ChatGPT. They were overwhelmed by new tech, unable to change their work habits, or simply too busy to try out a new tool.

Future GenAI initiatives must shrink the learning curve for such users by:

- Creating communities of practice for employees to learn and share ideas.
- Providing clear resources with step-by-step instructions for low-effort, high-impact uses of GenAI.
- Training “AI ambassadors” to serve as embedded subject matter experts within departments or teams.

NEXT STEPS

Investigate GenAI Further

As generative AI tools continue to develop, having a user-informed approach is essential. As more employees engage with Generative AI, it is important to continue engaging with them on opportunities and challenges. **Future AI feedback and research should include**:

- Which types of GenAI training resonate most with employees?
- Which employees would most benefit from what types of generative AI systems?
- How do best practices and adoption change when teams, not just individuals, add GenAI to their workflows?
- Where do opportunities to leverage AI tools spark ideas for more complex implementations?

NEXT STEPS

Share PA's AI playbook

Pennsylvania's GenAI pilot was the first such program in the nation. We hope to share what we learned with other states that follow in our footsteps. **We're particularly keen to share our user-centric approach** to equipping government employees with GenAI technology and love learning about similar efforts. If you work in government and want to learn more, don't hesitate to reach out.

Commonwealth
Office of Digital Experience

Questions?

RA-OACODEPA-UXR@PA.GOV

ARTIFICIAL-INTELLIGENCE@PA.GOV