pennsylvania

Employee-Centered Approach to Piloting Generative AI

Commonwealth of Pennsylvania

Office of Administration

Harrison MacRae – hmacrae@pa.gov

Emerging & Innovative Technologies

09.01.23 to 12.31.24

Executive Summary

Generative artificial intelligence (GenAI) has the potential to reshape how private and public sector organizations drive innovation, streamline operations, and enhance stakeholder engagement. Since the initial release of OpenAI's ChatGPT in fall 2022, GenAI tools have become widely known to the public through their ability to produce text and conversational interactions; in particular, features such as the ability to translate to other languages, convey information across mediums, and mimic human creativity. Government is in a unique position to benefit from this technology. However, there is currently no public sector consensus or best practices for what successful, safe, and sustainable GenAI implementation looks like.

Recognizing the importance of this moment, Pennsylvania Governor Josh Shapiro signed an executive order in September 2023 for the adoption and use of GenAI by Commonwealth agencies and employees. The order defines a balanced approach that emphasizes innovation and employee empowerment, as well as accuracy, adaptability, equity and fairness, mission alignment, privacy, proportionality, safety and security, and transparency. The order also established a Governing Board whose duties include making recommendations for the design, development, procurement, and deployment of GenAI.

Governor Josh Shapiro signs Executive Order 2023-19 - Expanding and Governing the Use of Generative Artificial Intelligence Technologies Within the Commonwealth of Pennsylvania on September 20, 2023 at the Carnegie Mellon University Block Center for Technology and Society in Pittsburgh, PA.

At the direction of the board, the Commonwealth of Pennsylvania's Office of Administration (OA) has embarked on a first-of-its-kind pilot program to explore how employees in a wide range of job duties can leverage GenAI tools in their day-to-day tasks. The pilot was launched in January 2024 and will continue through December 2024 with a focused footprint of 150 licenses for employee use. Participating employees provide their insights through structured user experience (UX) research that will help the Commonwealth make informed decisions on GenAI opportunities, risks, needed resources, and general employee sentiment toward the technology. In the initial results, employees have demonstrated immense enthusiasm, and self-reported efficiency gains and process improvements.

Idea

What Problem or Opportunity Does the Project Address: Public sector technology leaders need user-informed information on how to navigate this pivotal and novel technology to be successful in its adoption. Conducting an employee-centered pilot addresses the critical need for the Commonwealth and other government entities to have an informed approach about how, when, and where to leverage GenAI tools in the workforce.

Why Does it Matter: As many state agencies face the challenges of hiring talent in a competitive job market and losing a growing number of long-time public servants to retirement, they need to leverage every tool available to enable their workforce for success. Technology is a key piece of this equation and GenAI tools offer the opportunity to address critical inefficiencies within state operations, particularly focusing on manual and time-consuming tasks being performed by employees. The widespread gains envisioned through Gen AI can allow employees to focus on areas where their expertise is essential and deliver more effective results. This pilot is crucial for understanding where there is the most opportunity to transform state operations, reduce operational costs, improve employee productivity, and enhance customer satisfaction.

What Makes It Different: There have been a small number of GenAI pilots across tribal, local, state, and federal governments. However, none of them have focused with such depth on the perspective of, and impact to, their employees. For the most part, most government investments in GenAI are currently concentrated in technical teams looking to build products or solutions. The Commonwealth's pilot is a comprehensive approach rather than an isolated application of technology.

Additionally, the Commonwealth was the first U.S. government to sign a contract with OpenAI for its ChatGPT Enterprise product, and our pilot participants are the first government employees to officially use OpenAI's enterprise GenAI technology.

What Makes It Universal: The Commonwealth's work with GenAI is broadly applicable to other governments interested in adopting GenAI tools for their workforce. The initial pilot participant pool includes common government employee profiles that are widely applicable such as non-technical employees, union employees, and employees who had never used GenAI tools before participating in the pilot. This range of employees will allow the Commonwealth to scale the pilot to a wider range of users throughout the rest of 2024 and apply the learnings broadly to its workforce.

> **"***This is the first effort by the Commonwealth to involve employees in testing a new technology; I want to be part of that***."**
>
> **- Generative AI Pilot Participant**

Implementation

In addition to being a first of its kind use of new technology by state government, the methods and implementation of the pilot have also been innovative. Throughout this pilot, the Commonwealth has experimented with new ways to integrate Agile methodologies, human-centered design, and user experience research.

What Was the Roadmap: The GenAI pilot with Commonwealth employees seeks to gather detailed information in three categories:

- **Employee Experience**: Is this a tool that enables employees, and what training or resources are needed to do so?
- **Opportunities:** Where are there demonstrated, meaningful use cases that are valuable to employees and scalable across our workforce?
- **Risks:** What are the observed risks from employees using GenAI tools and what training, resources, or considerations must be taken to mitigate them?

The pilot was integrated into the broader enterprise framework with a clear, Agile management approach focused on iterative deployment and feedback to ensure alignment with the pilot's goals. This pilot is divided into three phases. The first phase is solely focused on the Office of Administration, which primarily supports HR and IT functions for state agencies. The second and third phases will then selectively spread to other OA employees and new agencies that have not yet participated. Additionally, in each phase, there are two pathways for employees to participate:

- **Collaboration group:** These employees progress through the pilot as a cohort, receive group training, and provide various types of quantitative and qualitative feedback about their experience.
- **Exploration group:** These are employees who have been given access to test a specific use case or provide visibility into a specific domain area (security, accessibility, etc.) and provide their input on an ad-hoc basis.

Integration of UX research into the implementation of the pilot played a pivotal role in optimizing the deployment of ChatGPT Enterprise and ensuring it effectively met the needs of the participants. This research component facilitated an in-depth understanding of employees' interactions with the tool and allowed the project team to configure the tool based on real-time user feedback, helping to drive engagement in the pilot, which is measured through the responses they submit through the various feedback channels.

How Did you Do It: The most significant barrier to implementing this pilot was to identify a path forward to procure a GenAI tool with acceptable terms and conditions and that can be executed in a timeframe relevant to a rapidly evolving technology. The Commonwealth leveraged a procurement directive that allows for piloting technology at a limited scale and

scope. This directive gave the Commonwealth the ability to contract with a supplier for less than a year and within a pre-defined cost. At the time of this procurement, there was only one enterprise product in the marketplace suitable for the goals and ambition of this pilot. Leadership from the Governor's Office and the Office of Administration helped to garner support from all key internal stakeholders within our technology, legal, procurement, and policy teams.

Impact

What Did the Project Make Better: There are three clear areas of progress that have emerged from the first phase of the pilot and contribute to the larger focus areas that define success.

1. **User Research Framework**: The User Research Framework created for the pilot can be scaled to other teams for a wide range of uses.
2. **Employee Experiences**: The pilot collected robust and candid feedback data that is enabling the Commonwealth to better understand GenAI and the specific opportunities and risks employees are encountering.
3. **Value Added to Employees:** The pilot has validated the types of use cases for GenAI and the value of these tools to employees.

How do you Know:

1) User Research Framework: In addition to the immediate results provided by this pilot, there is now a model for how the Commonwealth can collect user feedback from employees, constituents, and others. This new framework for gathering employee feedback had several key components:
    a) **Iterative and Participatory Approach:** Emphasizing an iterative research approach ensured that feedback was continuously integrated into the pilot, enhancing the GenAI tool's relevance and usability. This participatory approach also fostered a sense of involvement among employees, encouraging open communication and active feedback, which are crucial to the pilot's success.
    b) **Voluntary Participation:** Participation in the pilot was voluntary, with employees expressing interest by responding to a survey screener. This approach ensured that the feedback and data collected were genuinely reflective of employees' experiences and attitudes towards the technology.
    c) **Ethical and Legal Considerations**: Adhering to best practices in user research, the project maintained high ethical standards, including ensuring informed consent, voluntary participation, and strict adherence to legal and privacy regulations with respect to pilot participants. These measures safeguarded participants' rights and ensured the integrity of the research process.

d) **Diary Study:** A comprehensive diary study involved 57 participants who provided bi-weekly feedback through 15-minute surveys. These surveys captured their attitudes towards using GenAI and detailed examples of practical applications. This ongoing feedback mechanism helped the project team monitor and evaluate the tool's effectiveness and user satisfaction over time, making iterative improvements possible.

e) **Interviews and Focus Groups:** In addition to the diary studies, semi-structured interviews and focus groups were conducted to gather qualitative insights. These discussions allowed for a deeper exploration of user experiences, enhancing the understanding of the tool's impact on daily operations and identifying specific areas for improvement.

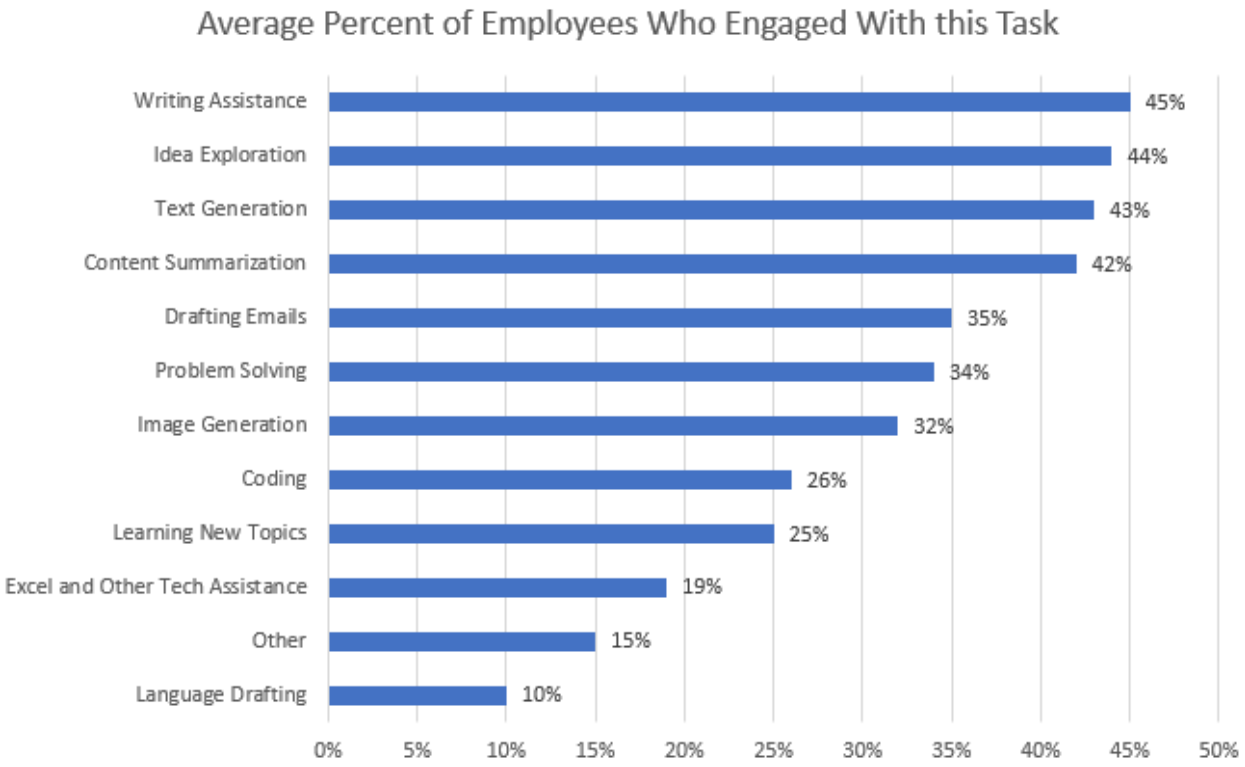
2) **Employee Experiences:** Throughout the pilot, a rich amount of feedback data is being collected through our user experience research methodology.

User Research Activity	**Number**
Diary study feedback entry submissions	206
Live audio recordings from Commonwealth employees responding to most helpful and least helpful use cases	53
Open text responses on use cases, training resources, and GenAI questions	300+
Recorded logs of employee prompt sessions that were identified as the most valuable use cases by the employee	69
Live employee feedback sessions	15
User interviews	12

3) **Value Added to Employees:** Employees self-reported using GenAI for a wide range of tasks. A high-level review of their feedback has allowed the Commonwealth to better understand the capabilities of GenAI they find most useful. In particular, the drafting of text is being leveraged by technical and non-technical employees alike to add to their productivity.

When surveyed about how much time they saved using GenAI in a day, the pilot participants self-reported an average time savings of 105 minutes, with qualitative feedback from participants underscoring improvements in service delivery and job satisfaction.

The above information can then broken down further to provide detailed insights on sub-populations of the pilot program. For example, in their first diary study entry, 13 employees reported using GenAI for various coding-related tasks, with half of them—7 out of 13—rating the tool as "extremely helpful." These employees reported a time savings of 42 minutes, underscoring the tool's efficiency in streamlining coding activities. When reviewing the open text responses and recordings of each employee, it became clear that there were two user populations who were using GenAI for coding, full-time developers who were performing advanced functions and program analysts who were exploring how to work with data. Such insights could lead to more tailored applications of GenAI tools that enhance productivity across diverse user profiles.



The following are examples from two pilot participants on their uses of GenAI and time saved:

> "*...take a set of High Level User requirements and break it down into items that can be delivered as part of an Agile IT delivery model which greatly reduced the time spent by BA's dev staff and Tech leads/architects*." - Self-reported time saving of 240 minutes.
>
> "*My team is taking over management of hundreds of multifunctional printing devices/scanners/faxes. I need a list of active devices and their IP address to run a scanning tool against them to pull an inventory since it was not up to date. Using this script it saved me from having to check each printer, one by one which would have taken days or possibly weeks. It was enough work I would have researched how to try to write the code myself."* - self-reported time savings of 300 minutes of work time.

What Now: The pilot's success so far has laid the groundwork for scaling GenAI across more departments. Areas of future focus will include:

- Exploring new agencies with a focus on employees hat have direct, resident-focused work.
- Seeking employees for the pilot who are historically underrepresented in the technology design process.
- Leveraging the pilot to test the effectiveness of lessons learned in employee training.
- Testing the efficacy of more specific use cases sourced from our pilot population.

These areas for future pilot exploration will allow the Commonwealth to continue to evaluate where, when, and how governments can scale GenAI tools in a safe, secure, and sustainable way. We look forward to sharing these insights with our peers and colleagues.

Conclusion

GenAI should be leveraged where appropriate as a tool that can provide better outcomes for government users, assist in digital transformation, and create no wrong doors to government information and services. GenAI tools are relatively untested in the market and in widescale employee use. To effectively leverage this technology responsibly, it is necessary to expose employees to opportunities to leverage GenAI in a safe environment. The pilot has enabled the Commonwealth to explore GenAI responsibly and be a leader in this space. The Commonwealth looks forward to leveraging GenAI in the future from a foundation that is built on employee-centered experiences through this pilot.