Information Systems Journal

WILEY

SPECIAL ISSUE ARTICLE OPEN ACCESS

An AI-Assisted Framework for Improving Innovativeness in Small Businesses: A Human–AI Collaboration Perspective

Kristijan Mirkovski[1] | Pete Williams[2] | Libo Liu[3] | Hao Liu[1] | Marta Indulska[4]

[1]Department of Information Systems and Business Analytics, Deakin Business School, Deakin University, Burwood, Australia | [2]Simply Headsets Pty Ltd, Melbourne, Australia | [3]School of Computing and Information Systems, Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia | [4]UQ Business School, the University of Queensland, Saint Lucia, Australia

Correspondence: Libo Liu (libo.liu@unimelb.edu.au)

Received: 1 July 2024 | **Revised:** 14 April 2025 | **Accepted:** 20 April 2025

Funding: This research was supported by the Australian Research Council, Discovery Project scheme (DP230102657).

Keywords: AI-assisted product innovation | human–AI collaboration | small business

ABSTRACT

Innovation is crucial for small businesses to remain competitive and adaptable in dynamic markets. Recent advancements in AI, particularly machine learning and natural language processing, offer promising tools for enhancing product innovation. However, small businesses often face significant challenges in adopting AI due to limited financial resources, data infrastructure, technical expertise, operational and cultural barriers. This paper presents a novel and holistic human–AI-assisted product innovation (HAIAPI) framework designed to address these challenges by integrating an advanced large language model approach across four key stages of the product innovation process: (1) AI-augmented problem articulation, (2) human expert problem selection, (3) AI-augmented solution generation and (4) human expert solution selection. Through an in-depth case study of an Australian e-retailer, this paper provides practical insights into how AI can enhance problem articulation and solution generation, while human expertise ensures relevant problem and solution selection. The detailed instructions on implementing this framework, including Generative Pre-Trained Transformers prompts, for small businesses are supported by a comprehensive resource toolkit and checklist detailing necessary financial, technical and human resources. Last, three key principles of human–AI collaboration are synthesised, offering further actionable strategies for small business managers/owners looking to effectively integrate AI into their product innovation processes.

1 | Introduction

Innovation is crucial for small businesses.[1] It drives growth, competitiveness and sustainability. In today's dynamic market, small businesses must continuously adapt and innovate to meet evolving customer needs and to stay competitive (Forsman 2011). Research shows that firms that prioritise digital innovation tend to perform better in terms of profitability and market share (Kohli and Melville 2019). For small businesses in particular, the adoption of new technologies is crucial to facilitate the development of products that differentiate them from larger competitors (Chan et al. 2019). Beyond creating new revenue streams, innovation improves operational efficiencies, reduces costs and enhances customer satisfaction (Mandviwalla and Flanagan 2021).

Recent data shows a 5% increase in global AI adoption by firms between 2022 and 2023, with North America leading at 61% (Thormundsson 2024). This increase reflects a growing reliance on AI to enhance operational efficiency and decision-making across industries. Importantly, AI has made considerable advancements in recent years, particularly through developments in natural language processing (NLP) and machine learning (ML), significantly transforming the product innovation process

(Bouschery et al. 2023). For example, Proven Skincare, specialising in personalised skincare, offers a rare case of a small business effectively utilising ML through its AI-powered Skin Genome Project (Prosser 2022). By analysing over 28 million customer reviews, scientific studies and ingredient data, Proven develops customised skincare products tailored to individual skin types, environmental conditions and lifestyle factors. This data-driven approach enables Proven to meet evolving consumer demands with personalised, science-backed skincare solutions, allowing the company to remain competitive in the beauty industry. This example illustrates AI's potential to automate and enhance various stages of product innovation, highlighting its importance as a tool for small businesses seeking to remain competitive and efficient (Verganti et al. 2020).

However, despite AI's transformative potential in product innovation, most small businesses struggle to adopt these technologies and establish effective human–AI collaboration. One of the most pressing barriers is limited financial resources, which often restricts small businesses from investing in advanced AI systems and the technical expertise required to implement and maintain such tools (Ransbotham et al. 2018). For example, hiring AI specialists or data scientists to design, train and manage AI models may place a financial burden on small businesses. Ongoing costs for maintaining and updating AI systems can also add to the financial strain, making long-term investment difficult (Ghezzi and Cavallo 2020). In addition to financial limitations, small businesses often lack the necessary data infrastructure to fully exploit AI's capabilities. AI systems depend on large datasets to detect patterns, make predictions and generate valuable insights (Kanbach et al. 2024). Yet, small businesses frequently have limited access to such data, either because they have not yet accumulated significant customer or product data or because they lack the tools to manage and analyse existing data (Duan et al. 2019). That is, while large corporations might have access to comprehensive customer databases, small businesses often deal with fragmented or incomplete datasets, reducing AI's ability to deliver accurate and actionable results.

Another significant challenge is integrating AI into existing business processes, which requires significant cultural shifts and organisational restructuring that can be daunting for small businesses. AI adoption often demands shifting from manual or human-driven decision-making processes to those that are more data-driven and automated (Bouschery et al. 2023). Specifically, small businesses typically rely on intuition-based, relationship-driven decision-making, which contrasts with AI's data-driven automation, creating resistance as employees fear job displacement or diminished roles (Brougham and Haar 2018). To add, unlike larger corporations, which often have dedicated change management teams, small businesses also operate with lean structures where employees wear multiple hats. This makes it challenging to implement AI-driven changes without disrupting daily operations (Schwaeke et al. 2024). Furthermore, building trust in AI outputs is essential, for both small business and large corporations, but often difficult due to the opaque nature of AI decision-making processes. Many AI models operate as 'black boxes' where the rationale behind certain outputs is not readily explainable. This lack of transparency can lead to scepticism among employees and decision makers, where personal experience and intuition have traditionally played a central role in decision making (Rudin 2019). For example, in a hypothetical scenario, if an independent local bookstore were to implement an AI-powered recommendation system, staff might resist, fearing it would replace their expertise in customer engagement. The AI's data-driven suggestions could conflict with traditional intuition-based curation, creating scepticism, while the owner might struggle with operational adjustments.

The rapid advancement of AI has created significant opportunities for product innovation, yet small businesses often struggle with its adoption due to financial, technical, cultural and operational barriers. While AI-driven frameworks typically cater to large organisations with extensive resources, small businesses require tailored, scalable solutions that prioritise accessibility and practical integration. Building on Bouschery et al.'s (2023) conceptual work on the use of large language models, specifically Generative Pre-Trained Transformers (GPTs), for product innovation and design, we aim to bridge this gap by developing a novel and holistic human–AI-assisted product innovation (HAIAPI) framework that integrates AI's analytical power with human expertise to ensure actionable insights by building on the Double Diamond design model (Design Council 2015). Our framework facilitates human–AI collaboration across the four distinct stages of the product innovation process: (1) problem articulation, (2) product selection, (3) solution generation and (4) solution selection. What sets our work apart is its emphasis on AI–human collaboration rather than automation alone, as evidenced by the four key stages of our framework: (1) AI-augmented problem articulation, (2) human expert problem selection, (3) AI-augmented solution generation and (4) human expert solution selection.

A major contribution of our paper is the development of a structured, resource considerate AI implementation guidance, complemented by a practical toolkit detailing the financial, human and technological resources needed for successful adoption. By incorporating GPTs, our framework allows small businesses to harness AI's capabilities without requiring extensive proprietary datasets or specialised technical knowledge. To validate our framework, we conducted an empirical case study with an Australian e-retailer, illustrating the practical viability of AI-assisted product innovation in small business environments. Based on our learnings from the case study, we distilled three key principles for effective human–AI collaboration, specifically addressing the challenges of small business. In our practical paper, moving beyond theoretical exploration, we offer concrete strategies that equip small businesses with the tools to integrate AI effectively, drive product innovation and enhance their competitive edge.

2 | Problem Statement

The increasing integration of AI in business operations presents substantial opportunities for enhancing product innovation, particularly for small businesses. However, despite the potential benefits, they often struggle to adopt AI-driven approaches due to limited financial resources (Ransbotham et al. 2018), inadequate data infrastructure (Kanbach et al. 2024), insufficient technical expertise (Duan et al. 2019) and cultural (Brougham and Haar 2018) and operational barriers (Schwaeke et al. 2024). These challenges hinder their ability to compete with larger counterparts that can leverage AI to streamline product development, improve efficiency and enhance customer satisfaction.

Research highlights the transformative role of AI in product innovation (e.g., Yin et al. 2023), yet its practical application in small business contexts remains underexplored. Large corporations have successfully implemented AI-driven solutions to automate their problem articulation and selection and solution generation and selection (BMW Group 2023; Creery 2024). However, these implementations typically require considerable resources, making them less accessible to small businesses (Duan et al. 2019). Consequently, there is a pressing need for frameworks that allow small businesses to integrate AI into their product innovation processes in a way that is both practical and resource-efficient. The significance of this issue extends beyond theoretical implications, as ineffective product innovation processes in small businesses lead to prolonged development cycles, increased costs and missed market opportunities (Heidenreich and Kraemer 2016). Traditional approaches to identifying and addressing product-related problems often rely on informal methods, such as anecdotal customer feedback or intuition-based decision-making, which lack scalability and systematic rigour, as evidenced by key insights from our case study. Without structured and resource-efficient AI-assisted frameworks, small businesses risk inefficient allocation of resources and delayed responsiveness to market demands.

Addressing these challenges requires a novel approach that integrates AI's computational power with human expertise to ensure both efficiency and relevance in product innovation processes. We thus develop and evaluate a comprehensive framework, designed to integrate AI capabilities and human intelligence across the four different stages of product problem and solution space in the product innovation process. By systematically leveraging GPTs for data-driven insights while preserving human judgement in critical decision-making processes, our framework provides small businesses with a structured yet flexible and resource-efficient model to enhance their innovation capabilities. Crucially, the practical significance of solving this problem is notable. By implementing an effective and efficient AI-assisted approach to product innovation, small businesses can enhance their ability to detect and prioritise customer problems, generate creative solutions, and make informed design decisions—all while reducing costs and development time. We contribute to business practice by offering a tested framework that small businesses can adopt to harness AI in a way that aligns with their unique constraints and operational needs.

3 | Background

NLP and ML are the main AI-assisted approaches for automating the first two stages of product innovation: problem articulation (e.g., our product's complexity is causing user frustration, leading to decreased satisfaction and higher return rates) and problem selection (e.g., prioritise product complexity issue over other concerns to improve customer satisfaction and reduce return rates). NLP tools can analyse vast amounts of unstructured customer feedback, product reviews and social media posts to detect recurring issues, such as product complexity leading to user dissatisfaction (Zachlod et al. 2022). For example, sentiment analysis, a subset of NLP, can identify specific complaints and gauge the intensity of customer emotions in user-generated data, such as customer reviews, social media posts, blogs and personal websites, helping to articulate problems such as frustration due to product design flaws (Aldunate et al. 2022). Moreover, through clustering techniques, ML can categorise identified problems based on their frequency and significance. For instance, unsupervised ML algorithms can group similar customer issues, highlighting major pain points such as product complexity, and then rank these issues according to their predicted impact on metrics such as customer satisfaction and return rates (Noori 2021).

ML is the key AI-assisted approach for automating the next two stages of product innovation: solution generation (e.g., the product design team held a series of brainstorming sessions to generate a wide range of potential solutions, combining ideas from different departments and stakeholder feedback) and solution selection (e.g., after generating multiple ideas, the product design team used multiple criteria to evaluate each solution's feasibility, cost and potential impact). Generative design tools, powered by algorithms, can rapidly create multiple design variations by exploring a wide range of possibilities based on predefined constraints, such as user preferences or cost limitations (Martorelli and Gloria 2023). For example, in addressing a problem such as product complexity, generative design tools, such as Autodesk Generative Design,[2] can propose various solutions that simplify the product's interface while maintaining core functionalities. In addition, ML is used to refine these generated solutions by predicting their performance based on historical data. For solution selection, multi-criteria decision analysis (MCDA) models, combined with ML, can automate the evaluation of each potential solution against criteria such as feasibility, cost-effectiveness and potential impact on customer satisfaction (Martorelli and Gloria 2023).

While these tools have tremendous promise, current AI approaches for automating product innovation processes, such as NLP and ML, are not ideal for small businesses, who often have limited resources. These methods often require large datasets, advanced technical expertise and substantial financial investment, all of which small businesses typically lack (Duan et al. 2019). GPTs, on the other hand, offer a more accessible solution for small businesses. GPTs are advanced AI models that generate human-like text by predicting and assembling words based on vast pre-existing language data. These models are designed to perform a range of language-related tasks, from answering questions to creating content, with minimal additional training (Orrù et al. 2023). Such tools are more suited for small businesses for a number of reasons. First, GPTs can function with smaller datasets, making it more feasible for businesses with limited data. Second, they are versatile and user-friendly, reducing the technical barriers that traditionally accompany AI implementation. Third, GenAI provides a wide range of applications, from automating content creation to facilitating brainstorming sessions, making it a flexible tool for innovation across various industries.

4 | Human–AI-Assisted Product Innovation (HAIAPI) Framework

The Double Diamond design model provides a comprehensive approach for understanding and effectively resolving complex

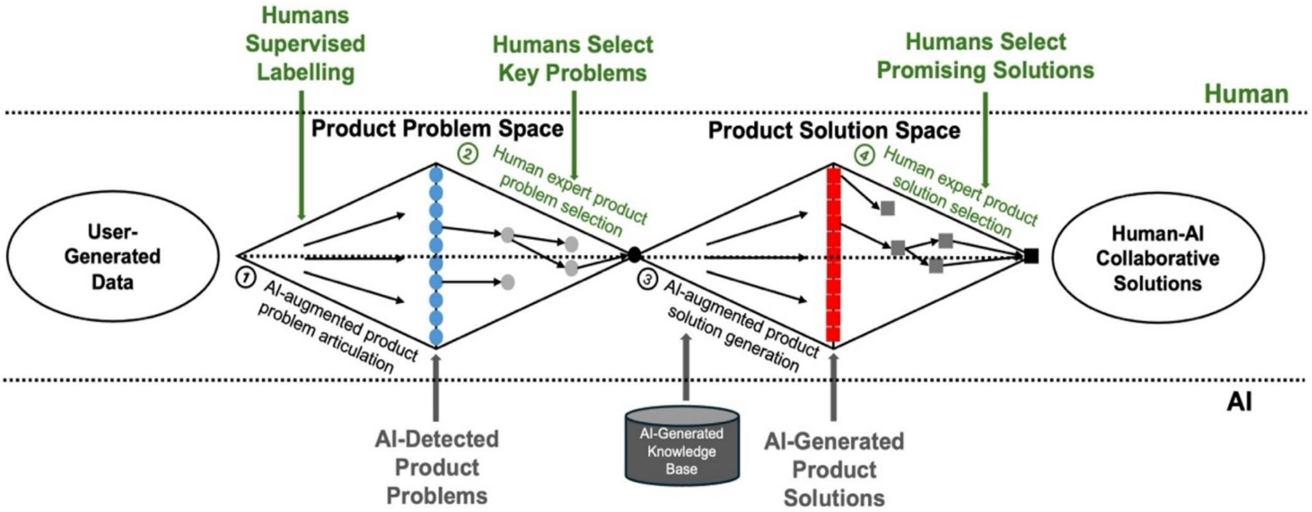

FIGURE 1 | Human AI-assisted product innovation framework.

problems, making it a valuable tool in design thinking and innovation (Design Council 2015). The model is divided into four stages: (1) discover (i.e., understanding the problem by gathering insights), (2) define (i.e., defining the problem by analysing insights), (3) develop (i.e., brainstorming to develop solutions) and (4) deliver (i.e., refining and delivering the final set of solutions), with each stage contributing to a structured approach to innovation. Bouschery et al. (2023) built on this model to propose a conceptual framework to demonstrate that GPTs can foster the innovation process, with their aim being to stimulate discussion, provoke thought and inspire further research on the integration of AI into human innovation teams.

Building on Bouschery et al.'s (2023) conceptual work, we developed a comprehensive and holistic framework—namely, HAIAPI—that details how AI can be effectively integrated into the four traditional design stages. In Figure 1, we illustrate our framework, which encompasses four stages addressing both the problem and solution spaces: (1) AI-augmented problem articulation, (2) human problem selection, (3) AI-augmented solution generation and (3) human solution selection. The *first stage*, AI-augmented problem articulation, involves leveraging existing resources, such as user-generated product reviews, to identify opportunities for innovation by pinpointing existing problems or complaints using AI. It comprises of divergent thinking, where product experts gather information to clearly understand the problem and identify underlying issues and unmet needs that may not be immediately apparent (Patnaik and Becker 1999). In this stage, we develop a new approach to articulating problems using GPTs for data labelling and problem extraction and summarisation. The *second stage*, human expert problem selection, remains primarily human driven. It involves critical thinking, empathy and contextual understanding, areas where human judgement excels (Dorst 2011). In this stage, human experts use their experience and intuition to evaluate the significance and feasibility of potential AI-generated product problems, ensuring that the focus is on the most impactful issues.

In the *third stage*, AI-augmented solution generation, AI can explore various combinations of design elements and rapidly produce multiple design alternatives based on predefined problems. This capability allows designers to consider a wide array of potential solutions early in the process, which can then be refined and iterated upon (Bouschery et al. 2023). In this stage, we propose a novel GPT-based approach for construction of a knowledge base of potential solutions and solution idea extraction, which in turn ensures that the ideation process is both expansive and efficient, fostering an environment conducive to innovation and creativity. The *fourth stage*, human expert solution selection, is predominantly human driven and again involves convergent thinking, focusing on evaluating and refining the generated concepts. It involves critical evaluation, where human expertise is essential to interpret AI-generated insights within the broader context of social, cultural and ethical considerations (Brown 2008). In this stage, through iterative testing and refinement by product experts, the most promising concepts are developed into detailed designs ready for implementation.

5 | Case Study

We present a case study[3] detailing the operationalisation and implementation of the proposed HAIAPI framework in collaboration with an Australian e-retailer—*Simply Headsets Pty Ltd*. Established in January 2008 by Pete Williams, Simply Headsets operates as the largest e-retailer in the Australian headset market. With a small team of approximately 15 employees, the company distributes renowned headset brands such as Jabra, HP Poly, Yealink and EPOS, serving around 90 000 customers nationwide.

Innovation in the headset retail industry is essential due to rapid changes in consumer preferences and advances in technology. Recognising this, Simply Headsets, with annual revenue surpassing AU$10 million, initiated plans to relaunch its private-label headset project. The company historically relied on a simple and informal approach to new product development, resulting in costly outcomes and lacking structure. Simply Headsets' approach to identifying customer issues was reactionary, based on short-term feedback from sales teams, rather than driven by long-term strategic goals. This led to elongated development cycles, typically spanning 6–8 months, resulting in costs of up to AU$10 000 for a new product.

TABLE 1 | Product development processes and stages at simply headsets.

Product problem space	Product solution space
Simply headset's existing process and stages	
Stage 1: Problem Identification	**Stage 2: Solution development and selection**
Step 1: Manual data collection and storage – Sales representatives and customer service staff manually record customer complaints, returns and feedback in a spreadsheet. – Order issues, product defects and support requests are logged in spreadsheets or written records. – Warranty claims and product failure incidents are documented manually. – Customer feedback from emails and calls are compiled into a centralised but non-automated system. Step 2: Manual extraction of key problems – Staff members review collected data to identify recurring issues, such as faulty microphones, poor battery life or connectivity problems. – Critical issues are flagged based on frequency, severity, and impact on customer experience. – Staff rely on intuition and experience to prioritise issues. Step 3: Group discussion about key problems – Team meetings are held to discuss and validate identified product issues. – Sales, customer service and technical support staff share insights on frequently reported problems.	Step 1: Brainstorming potential solutions – Gather sales, support, and store staff for a relaxed discussion—maybe over coffee. – Share real customer complaints or feedback to spark ideas on what needs fixing. – Look at what other retailers are doing—any better products or services we can adopt? – Identify quick, low-cost improvements (e.g., clearer product descriptions, better troubleshooting guides). Step 2: Feasibility check and selecting most practical solutions – Narrow down the best ideas with a quick team discussion—no overcomplicating. – Reality Check: Can we actually pull this off with our current budget, suppliers and stock? – Reach out informally to see if they have better models or warranty options. – Customer Impact: Would this change make customers happier or just complicate things? – DIY vs. Supplier Fixes: Can we solve this ourselves (better product listings, troubleshooting guides) or do we need a manufacturer fix? – Ask a few regular customers what they think—would this solution help them?

Recognising these inefficiencies and the rising costs associated with their new product development approach, founder Pete Williams sought to revamp the process by integrating AI-driven insights and collaborating with academic partners to implement the HAIAPI framework, aiming to reduce development time and costs while fostering innovation. Table 1 summarises Simply Headset's existing new product development process and its stages, while Table 2 summarises the processes and stages of the HAIAPI framework. In the following sections, we outline the operationalisation and implementation of the HAIAPI framework within the product problem and solution spaces.[4] Last, we discuss the evaluation of our framework.

5.1 | Product Problem Space

To operationalise and implement the product problem space of our proposed framework, we synthesised the collected customer reviews to identify key insights and define a clear problem statement, an approach small businesses can follow to inform their product innovation efforts. The objective in the product problem space is to narrow the broad spectrum of issues to specific, actionable problems that can be addressed through design, ensuring that design efforts are directed towards the most impactful and feasible challenges (Dorst 2011). The product problem space within our framework consists of two stages, broken down into five critical steps: (1) data access, collection and preprocessing, (2) data labelling, (3) atomic problem extraction, (4) atomic problem deduplication and summarisation and (5) problem selection and prioritisation. These steps, conducted by both AI and human experts, can guide small businesses in systematically identifying and prioritising product-related issues, ensuring that their design efforts are directed towards addressing the most critical and actionable product issues perceived by customers. By following these steps, businesses can streamline their product problem identification and selection processes; thus, enhancing product development and customer satisfaction.

5.1.1 | Operationalisation of Stage 1: AI-Augmented Product Problem Articulation

5.1.1.1 | Step 1: Data Access, Collection and Preprocessing. The data used to operationalise our framework was sourced from the 'Electronics' subset of the McAuley-Lab/Amazon-Reviews-2023 dataset (Hou et al. 2024). To identify headset products within the dataset, we searched for the keyword 'headset' in the 'main category', 'title' and 'categories' fields of the review metadata. Small businesses can access similar databases through platforms such as Kaggle (https://www.kaggle.com/datasets), the UCI Machine Learning Repository (https://archive.ics.uci.edu/datasets) and Papers with Code (https://paperswithcode.com/datasets). To obtain relevant data, businesses can search these platforms for industry-specific reviews, using keywords related to their product or market. They can also download the data in formats such as CSV, JSON

TABLE 2 | HAIAPI framework's processes and stages.

Product problem space	Product solution space		
HAIAPI's new process and stages			
Stage 1: AI-augmented product problem articulation (4 weeks)	**Stage 2: Human expert product problem selection (2 weeks)**	**Stage 3: AI-augmented product solution generation (6 weeks)**	**Stage 4: Human expert product solution selection (6 weeks)**
Step 1: Data access, collection and preprocessing. – 25 265 customer headset reviews were retrieved from McAuley-Lab/Amazon-Reviews-2023 dataset. – A final set of 21 072 textual customer headset reviews (816 795 words in total) were acquired after pre-processing. Step 2: Data labelling. – Product experts identified 14 headset aspects. – Product experts labelled 307 headset customer reviews and GPT-4 Turbo labelled all the customer reviews. Step 3: Atomic problem extraction. – GPT-4 Turbo extracted 17 045 atomic problems with a total of 310 650 words from large product review dataset. Step 4: Atomic problem deduplication and summarisation. – Using the K-Means++ clustering method to remove duplicate atomic problems, 92 distinct clusters of atomic problems were identified. – Final set included yielded 417 uniquely identified headset problems across 14 product aspects.	Step 5: Problem selection and prioritisation. – Product experts independently reviewed the 417 headset problems to identify the most critical issues – A focus group session was organised with the product experts to discuss and finalise the key problems. – Final set of 30 key headset problems were prioritised as the most urgent and feasible ones.	Step 6: Construction of AI-generated knowledge base. – Using GPT-4 Turbo, 7775 atomic ideas were extracted to populate the knowledge base. Step 7: Atomic idea retrieval. – For each problem identified by human experts, atomic ideas were retrieved with a cosine similarity score exceeding 0.5 using the Text-Embedding-3-Small embeddings. – For each problem identified in Stage 2, GPT-4 Turbo generated 10 potential solutions out of which five are concrete, specific and feasible ideas and five are groundbreaking, visionary and potentially revolutionary ideas. – A total of 4170 solutions, comprising 2085 concrete ideas and 2085 visionary ideas were produced.	Step 8: Solution review and selection. – Product experts were tasked to select 60 most promising ideas from a pool of 2085 AI-generated concrete, specific and feasible ideas and 2085 AI-generated groundbreaking, visionary and potentially revolutionary ideas (solutions) – Experts were also tasked to provide comments on how to improve the solutions to make them more feasible. – Three headset solutions were selected to move forward to prototype development.

or XML, or access it through APIs, making it easier to incorporate the data into their analysis workflows. Alternatively, small businesses can use spreadsheet tools such as Google Sheets (https://docs.google.com/spreadsheets).

We then filtered user reviews to include only those published between April and September 2023, yielding a total of 25 265 customer headset reviews. The period from April to September 2023 was selected to capture recent and relevant user feedback. This 6-month window balances the need for data recency with a manageable dataset size, providing insights applicable to the present market conditions. We further pre-processed the data by removing reviews with fewer than 10 words and splitting lengthy reviews at sentence boundaries into manageable segments of up to 100 words each. Small businesses can automate the removal of short reviews with limited insight using spreadsheet functions or a Python script, for which they need to hire or assign a programmer proficient in Python and experienced with NLTK or SpaCy to develop and test scripts for data preprocessing. For segmenting lengthy reviews, NLP libraries such as NLTK (https://www.nltk.org) or SpaCy (https://spacy.io) can be used. In the Simply Headsets case, completing data preprocessing resulted in 21 072 textual customer headset reviews, totalling 816 795 words. This step ensures that only relevant data are retained, improving the focus of the subsequent analysis.

5.1.1.2 | Step 2: Data Labelling. To assist with data labelling, three product experts at Simply Headsets identified 14 key product aspects that significantly influence consumer purchasing decisions, including aftersales support services, battery, comfort, connectivity, controls design/functionality, endurance, material quality, noise cancellation, price, sound quality, shape/appearance, microphone quality, portability and water resistance. Each aspect's definition and measurement scales were clearly defined by experts. Small businesses can replicate this task by first identifying the key factors relevant to their industry, using team discussions, customer feedback analysis or market research to determine which aspects are most important to their target audience. For example, in consumer electronics, aspects such as sound quality and comfort are critical, whereas in sectors such as home appliances, factors such as energy efficiency and durability may be more pertinent.

Once the aspects are identified, businesses should establish clear, measurable scales for each, such as numerical ratings (e.g., 1–5 for satisfaction) or descriptive labels (e.g., low, medium, high for durability). Precision in defining these scales is crucial to avoid ambiguity in the data labelling process and to ensure consistent application. For instance, a rating of 1 for comfort could represent 'uncomfortable for short use', while a 5 might indicate 'highly comfortable for extended use'. They can further refine these scales through iterative testing, where a team labels a sample dataset and adjusts the scales as needed to resolve any inconsistencies.

In the Simply Headset case, the experts initially labelled 307 customer reviews each (921 in total) for detecting headset-related problems. This task, which took five and a half hours, was reported to be time-consuming and resulted in information overload and inconsistencies. To automate this task, GPT-4 Turbo was prompted to label the reviews based on these predefined aspects and scales, and experts spent 2 h refining the prompts for ChatGPT-4 Turbo to label all the review data.[5] The initially labelled 307 reviews by experts were used to test the prompt. Small businesses can implement this AI-assisted labelling by first formulating clear prompts for GPT-4 Turbo, for which they should hire or assign a programmer proficient in Python and experienced with GPT-4 Turbo's API to develop and test scripts and prompts for large dataset labelling. Access to GPT-4 Turbo is available through OpenAI platform, where businesses can subscribe to use API access to integrate the model into their existing workflows. A well-structured prompt should direct the AI to analyse customer reviews systematically by providing clear, step-by-step instructions. This ensures a structured, evidence-based analysis, enhancing both the accuracy and transparency of the AI's assessment process. After creating the initial prompt, it is essential for businesses to test its effectiveness by running a small set of reviews through the AI to evaluate the accuracy and quality of its responses. If necessary, the prompt can be refined to improve clarity and consistency. Small businesses should also identify and hire or assign three suitable product experts to manually label sample datasets and refine GPT-4 Turbo's prompts.

To evaluate GPT-4 Turbo's performance, we tasked the Simply Headset experts to label 307 reviews, which were also labelled by GPT-4 Turbo for comparison. The accuracy showed an average agreement of 91.03%, demonstrating that GPT-4 Turbo's performance was comparable to human experts. This step reduces manual workload while maintaining accuracy, allowing small businesses to streamline the analysis of customer feedback augmented by AI.

5.1.1.3 | Step 3: Atomic Problem Extraction. Customer reviews often address multiple issues across various product aspects, necessitating a targeted approach to analysis. To enhance the relevance and focus of aspect-specific corpora, we applied atomic problem extraction. This task, facilitated by GPT-4 Turbo, identified and isolated discrete, standalone (atomic) problems related to specific product aspects within each review. This extraction approach improved the information density of the product review data, enabling a more precise and granular analysis of the user-generated data. We also iteratively optimised the prompt to maximise the correctness of the extracted problems to their source reviews by using GPT-4 Turbo.[6]

Small businesses can leverage atomic problem extraction using GPT-4 Turbo via its API to identify specific, unique issues related to product design or development, such as 'battery life too short' or 'uncomfortable ear cushions', for which they should hire or assign a programmer proficient in Python and experienced with GPT-4 Turbo's API to develop and test scripts and prompts for atomic problem extraction. To begin, businesses should start with clear instructions for GPT-4 Turbo. To improve the precision and relevance of the extracted problems, they should iteratively refine the prompt based on initial results. For example, if the AI's outputs are too vague, the prompt can be adjusted to be more specific. By testing variations of the prompt and reviewing the results for accuracy, they can optimise the extraction process, assuring that the atomic problems closely reflect the original review content and are aligned with the targeted product aspects.

In the case of Simply Headsets, the ChatGPT-4 Turbo generated 17 045 atomic issues with a total of 310 650 words, providing a rich dataset for further analysis and product refinement. By following this step, small businesses can effectively extract detailed product problems from customer reviews, facilitating more targeted product development efforts.

5.1.1.4 | Step 4: Atomic Problem Deduplication and Summarisation. The deduplication process was performed on the atomic problem dataset using the K-Means++ clustering method (Arthur and Vassilvitskii 2007) to remove duplicate atomic problems with the help of Eext-Embedding-3-Small embeddings provided by OpenAI, resulting in a total of 92 distinct clusters of atomic problems.

Small businesses should hire or assign a programmer proficient in Python and experienced with GPT-4 Turbo's API, Text-Embedding-3-Small's API, and Scikit-learn to develop and test scripts and prompts for atomic problem embeddings, clustering, deduplication and summarisation. To replicate the deduplication task, they should first convert text data, such as customer reviews or identified problems, into numerical vectors. This is done using text embedding models, such as OpenAI's Text-Embedding-3-Small, which capture semantic similarities between different text entries. These embeddings can be generated via OpenAI's API, providing a foundation for subsequent analysis by representing text data in a structured numerical form suitable for ML tasks. Once the data are converted into numerical vectors, the next task is to group similar atomic problems into clusters using the K-Means++ algorithm, which is accessible through tools such as Python's Scikit-learn library (https://scikit-learn.org). The clustering task is critical as it helps to identify and eliminate redundant feedback, thereby consolidating similar reviews or issues into unique problem groups.

In the case of Simply Headsets, we utilised GPT-4 Turbo to further summarise these clusters and ultimately yielded 417 uniquely identified problems.[7] Each summarised problem included a reference to the original set of issues that informed the conclusion. After clustering, businesses can utilise GPT-4 Turbo to summarise the core issues within each cluster. By feeding the clustered data, such as customer reviews or atomic problems, into GPT-4, concise summaries can be generated that capture the main themes or concerns from each group. This step reduces the redundancy and complexity of the data and also provides a clear direction for small businesses in addressing key product design or development issues.

5.1.2 | Operationalisation of Stage 2: Human Expert Product Problem Selection

5.1.2.1 | Step 5: Problem Selection and Prioritisation. In the Simply Headsets case, GPT-4 Turbo generated 417 key problems across 14 product aspects, categorising issues by factors such as price, quality and performance. To prioritise these identified problems, three product experts at Simply Headsets independently reviewed each issue, assessing them based on two primary criteria: urgency (the immediacy and severity of the problem) and feasibility (the ease with which the problem can be addressed). This assessment was conducted using a simple rating scale, for instance, a 1–5 scale where a higher number indicates greater urgency or feasibility.

Small businesses can adopt a similar approach by first ensuring they have a diverse group of experts who understand different facets of the product and customer experience. They should hire or assign three suitable product experts to select and prioritise urgent and feasible problems. These experts should independently evaluate each AI-generated problem, assigning ratings that reflect both the urgency and feasibility of addressing each issue. To facilitate this, they can use spreadsheet software (e.g., Google Sheets) to organise and record the ratings systematically in a well-documented, transparent and easily accessible way to stakeholders.

After the independent reviews were completed, a focus group session was held with the experts from Simply Headsets to discuss the identified problems, compare their ratings and reach a consensus on the most critical issues. This collaborative process allowed the experts to refine their individual assessments and agree on which problems required immediate attention. Following the session, the experts finalised a list of 30 headset problems that were considered both urgent and feasible to address[8]. For small businesses, this collaborative approach can be easily replicated by organising focus groups or workshops to strategically prioritise identified product problems. This step can help businesses to reduce an overwhelming number of problems into a manageable, targeted and actionable list.

5.1.3 | Insights From Implementation of Stages 1 and 2 at Simply Headsets

In the implementation of Stage 1, during the data labelling, the three product experts at Simply Headsets encountered ambiguity in defining and applying the 14 product aspects. Despite establishing clear definitions, some product problems were difficult to categorise. The root cause of this ambiguity lied in the complexity and overlap inherent in product features and customer feedback. For example, a customer complaint about volume control posed a classification challenge for the experts. While some viewed it as a 'sound quality' issue, others considered it more appropriate to categorise it under 'controls design/functionality'. This variation in interpretation reflected the underlying difficulty of assigning rigid categories to issues that are multifaceted and context dependent. The variability was noted by one product experts:

> I was trying to work out this and it took me some time. It wasn't very straightforward. Like, where should I fit a headset volume issue? I wouldn't put it under sound quality [...] There was a lot of room for interpretation.—Product Expert A

Discrepancies between human interpretations and AI-generated labels for customer reviews were observed, stemming from fundamental differences in how each approaches the labelling process. GPT-4 Turbo applied predefined labels strictly according to established rules, ensuring consistent application but without the ability to interpret nuanced contextual information or sentiment. In contrast, human experts relied on subjective judgement, which was influenced by the tone, sentiment and broader context of the review, leading to variability in classifications. This variability arose because humans are capable of interpreting subtle cues in language and considering the overall emotional tone of a review, whereas AI systems operate within the constraints of their algorithmic parameters. The root cause of this discrepancy lied in the divergence between algorithmic precision and human intuition. As two of the experts stated:

> Initially, I label one review as a 2. But then, after reading another review, you'd think, 'Oh no, this one should be a 2.' This made you question if you needed to go back and re-label the earlier reviews. There was no consistent standard.—Product Expert C

> The AI stuck to its parameters, but sometimes, that didn't align with the subtle differences I picked up in the reviews. It's like it was missing the context that a human would immediately understand.—Product Expert A

The process of refining the prompts for GPT-4 Turbo to accurately extract atomic problems from customer reviews required multiple iterations, which was regarded as demanding by the three product experts from Simply Headsets. Initially, the AI struggled to distinguish between overlapping issues within the same review, especially when customers raised concerns about different product aspects simultaneously. For instance, reviews often contained feedback about both 'battery life' and 'sound quality' making it difficult for the AI to treat these as distinct problems. Early iterations of the prompt often produced overly general or simplified responses, failing to capture the more nuanced problems conveyed by the customers. To address this, the product experts, together with the research team, repeatedly adjusted the prompt, testing different phrasings and instructions on the AI in separating these complex, multi-dimensional reviews into clear, standalone problems. Two of the experts reflected:

> We helped had to tweak the prompts a quite many times [...] the AI combined different issues, but with each adjustment, it got better at separating the problems clearly.—Product Expert A

> It took some trial and error to make sure the AI understood the difference between related issues. Its first pass would be too vague and we had to fine-tune the language to get closer to what the customer actually meant.—Product Expert B

In the implementation of Stage 2, the three product experts at Simply Headsets questioned the practicality of AI-generated product summarisation, especially when dealing with relatively small datasets comprising around 300–500 customer headset reviews. They highlighted a significant concern regarding AI hallucinations—instances where the AI fabricates issues not actually present in the customer review data. As one product expert put it, 'Humans would not say that', stressing the notion that AI sometimes identifies problems that are not explicitly stated by customer reviews. Another product expert added, 'It's like it knows what it's talking about but it's seeing problems where people haven't said there's an issue,' pointing to the AI's tendency to overanalyse and infer issues that might not exist. Furthermore, they criticised the AI for generating excessive and unnecessary details, a phenomenon they referred to as 'creating problems within problems'. This dilutes the focus and complicates the extraction of actionable insights, as argued by two product experts:

> The AI keeps producing layers of irrelevant information [...] making it harder for us to pinpoint the actual issues. It's like creating problems within problems and this just ends up muddling the overall analysis instead of clarifying it—Product Expert A

> I noticed that the AI's summaries often include too many details that don't really help me. Instead, this complicates the extraction of good actionable insights [...] This turns a straightforward problem into a convoluted mess of minor issues—Product Expert B

Despite the criticisms regarding small datasets, the three product experts acknowledged the efficiency of AI in handling large datasets, such as those exceeding 10 000 customer headset reviews. They noted that in these scenarios, AI problem summarisation becomes particularly valuable. The ability of AI to process vast amounts of data swiftly and consistently was praised, with experts suggesting that such outputs would be especially beneficial for novices lacking experience in the retail headset industry. 'AI doesn't suffer these distractions', one expert remarked, alluding to the AI's capacity to remain focused and objective, free from the biases and distractions that can affect human analysis. The product experts conceded that AI-generated problem summaries for large datasets might indeed be more objective and of higher quality compared to manually extracted summaries by humans, who might miss patterns or become overwhelmed by the volume of data, as explained by two experts:

> When dealing with thousands of reviews, it's easy for us to miss recurring patterns or become overwhelmed by too much data thrown at us. AI, however, could process everything consistently and perhaps objectively [...] highlighting trends we might overlook—Product Expert C

> I think, AI-generated summaries for large datasets should be more reliable because they aren't influenced by human biases or fatigue [...] From what I gather, AI could handle large volumes of data easily and spit out cool insights that might take us much longer to extract and put together manually—Product Expert B

The three product experts at Simply Headsets also inherently distrusted AI-generated problem summarisation. They expressed a need to double-check AI findings by cross-referencing the original reviews from which these summaries were derived. This scepticism stems from a desire to understand the logic behind the AI's decision-making process. As one product expert noted, there is a critical need to know how the AI arrives at selecting key headphone problems that need urgent attention. Without transparency in the AI's reasoning, the experts felt uneasy relying solely on AI-generated summaries. They emphasised that 'human intervention is necessary to make sense out of the problems', highlighting the belief that while AI can aid in problem identification, human expertise is crucial for contextualising and prioritising these issues. This is elaborated on by two product experts:

> Having the ability to refer back and verify that a problem is valid based on actual data, versus seeing information that has been misinterpreted, is essential [...] So, understanding the AI's reasoning and how it derived its conclusions about the problems would be a must—Product Expert A

> AI is not entirely up to the task yet; there still needs to be a lot of manual intervention. However, AI can generate good ideas, even if some problems it identifies aren't explicitly mentioned. For instance, people do experience issues connecting devices to their PlayStations. Whether that is a major problem is for people to decide, but there is definitely some usefulness in collaborating with AI.—Product Expert C

5.2 | Product Solution Space

To operationalise and implement the second section of our proposed framework, we build upon the selected and prioritised product problems to generate viable product solutions. The primary objective in the product solution space is to develop a range of product innovation ideas that are grounded in the most pressing product problems, yet are concrete, specific and feasible or visionary and groundbreaking to provide small businesses with actionable insights to inform their innovation efforts (Chiesa et al. 2009). The product solution phase of the framework is organised into two main stages, which are further divided into three key steps: (1) AI-generated knowledge base, (2) atomic ideas extraction and (3) solution review and selection. These steps, involving both AI and human expertise, enable small businesses to systematically and rapidly generate product innovation concepts. By following these steps, businesses can accelerate their idea generation while ensuring that product development is aligned with market needs, resulting in innovative products that directly address the most critical issues identified in customer feedback and increasing the likelihood of customer satisfaction and competitive advantage.

5.2.1 | Operationalisation of Stage 3: AI-Augmented Product Solution Generation

5.2.1.1 | Step 6: AI-Generated Knowledge Base. We conducted atomic idea extraction to populate a knowledge base of potential product solutions. *Atomic ideas* refer to individual suggestions or solutions that are aimed at guiding product improvement or innovation, with each atomic idea representing a distinct and standalone concept. In the Simply Headsets case, using GPT-4 Turbo, we extracted 7775 atomic ideas to populate the knowledge base.[9]

Small businesses should assign a programmer proficient in Python and experienced with GPT-4 Turbo's API to develop and test scripts and prompts for extracting atomic ideas and constructing a knowledge base. To replicate this task, they should utilise GPT-4 Turbo by inputting customer reviews and prompting the AI to identify atomic ideas for each product aspect. For example, if customers frequently express dissatisfaction with battery life, the AI might identify atomic ideas such as 'extend battery capacity' or 'introduce faster charging technology'. By continuing this task across multiple reviews, businesses can build a knowledge base containing numerous atomic ideas covering various product aspects.

Once the atomic ideas are generated, they should be systematically stored in a structured knowledge base to enable efficient idea extraction and generation in the next step. Businesses should first select an appropriate storage method based on the volume and complexity of their data. For smaller datasets, simple solutions, such as Google Sheets, may suffice. However, for larger datasets, more robust systems such as MySQL (https://www.mysql.com) or MongoDB (https://www.mongodb.com) databases are recommended. The next step is to define key data fields for organising the atomic ideas. Essential fields may include the *atomic idea* (the specific suggestion or solution), *product aspect* (the product category or feature it addresses, such as battery life or sound quality) and *a customer review reference* (a link to the original review from which the idea was generated). This structured step can assist small businesses to organise customer insights systematically and effectively utilise them to guide their product development efforts.

5.2.1.2 | Step 7: Atomic Idea Retrieval. After building the knowledge base, we next conducted atomic idea retrieval. We retrieved atomic ideas based on the product problems identified by human experts using a cosine similarity score exceeding 0.5, calculated with the Text-Embedding-3-Small embeddings model. This allowed us to measure the semantic similarity between the product problems and the atomic ideas stored in the knowledge base. Atomic ideas that exceeded the threshold were selected for further analysis. Once these relevant ideas were retrieved, they were input into GPT-4 Turbo to develop feasible, practical solutions as well as visionary, groundbreaking ideas. For each problem, GPT-4 produced 10 potential solutions, 5 of which were concrete and feasible and 5 that were more innovative and forward-thinking. In the Simply Headsets case, this step produced a total of 4170 solutions, comprising 2085 concrete ideas and 2085 visionary ideas.[10]

Small businesses should hire or assign a programmer proficient in Python and experienced with GPT-4 Turbo's API and Text-Embedding-3-Small's API to develop and test scripts and prompts for atomic idea embeddings, problem matching, atomic idea retrieval, and the generation of concrete and groundbreaking ideas. To replicate this step, they should first calculate cosine similarity scores between the identified product problems and the atomic ideas. Using Text-Embedding-3-Small (via OpenAI API), businesses can transform both the problem statements and atomic ideas into numerical vectors that reflect their semantic content. Once the embeddings are generated, cosine similarity should be calculated between the problems and the atomic ideas. A threshold of 0.5 or higher is recommended to ensure that only the most relevant atomic ideas—those closely aligned with the product problems—are retrieved (Mihalcea et al. 2006). By filtering out ideas below this threshold, businesses can focus their attention on the most relevant and actionable insights.

Next, the retrieved atomic ideas should be organised in a structured format for further analysis. Businesses can store these ideas in a spreadsheet or a database, categorised by the specific problem they address. At this point, the stored ideas are input into GPT-4 Turbo for solution generation. They should prompt GPT-4 to generate both concrete, specific solutions and visionary, innovative ideas. GPT-4 will generate solutions tailored to each problem, offering a balance between actionable improvements and innovations.[11]

5.2.2 | Operationalisation of Stage 4: Human Expert Product Solution Selection

5.2.2.1 | Step 8: Solution Review and Selection. A final set of 60 potential solutions produced by AI was sent to the three product experts at Simply Headsets for review and selection of the most promising solutions/ideas. The experts were tasked to select the 30 most promising ideas from a pool of 2085 AI-generated concrete, specific and feasible ideas. They were also presented with 2085 AI-generated groundbreaking, visionary and potentially revolutionary ideas, out of which they needed to select 30 and provide comments on how to improve them to make them more feasible.[12] A focus group was then organised to discuss and finalise the selection of the most promising ideas/solutions. Last, the three experts, along with the founder/manager and other design specialists from Simply Headsets, selected three product solutions to move forward with prototype development.

Small businesses should hire or assign three suitable product experts to review and select viable solutions. When conducting the review step, they should consider several key criteria: feasibility, relevance and potential impact. *Feasibility* refers to whether the idea can realistically be implemented with the business's current resources, technology and market constraints (Novak 1996). For example, a solution that requires advanced materials or technology that the business cannot access may be less feasible. *Relevance* focuses on how closely the idea aligns with the business's product goals, customer needs or market trends (Berry and Shabana 2020). A highly relevant idea directly addresses pressing customer issues or future market demands. *Potential impact* involves evaluating the degree to which the idea can affect customer satisfaction, product performance or market competitiveness (Alegre et al. 2012). High-impact ideas typically solve critical problems or introduce innovative features that can differentiate the business in the marketplace.

To reach consensus, businesses can use several tools and techniques. Voting systems such as the Delphi Method (Ehringfeld et al. 2023) can be employed, where experts anonymously provide ratings for each idea, followed by rounds of feedback to refine selections. Alternatively, they can use digital platforms such as Trello (https://trello.com) to facilitate real-time collaboration, where team members can review, comment and vote on ideas/solutions. Last, collaborative decision-making software such as StormBoard (https://stormboard.com/) can also help visualise the prioritisation process by allowing teams to rank ideas based on the agreed criteria (feasibility, relevance and impact). These tools encourage transparency and participation, helping businesses streamline the decision-making process and ensure that the most valuable ideas are selected for implementation.

5.2.3 | Insights From Implementation of Stages 3 and 4 at Simply Headsets

In the implementation of Stage 3, the product experts observed that the quality and specificity of the atomic ideas and solutions produced by GPT-4 Turbo were directly dependent on the AI's capabilities. The AI often missed crucial nuances or oversimplified customer feedback, leading to ideas that were too vague or lacked the necessary detail. This reliance on AI for generating solutions necessitated additional human oversight to provide the contextual understanding needed to enhance the relevance and applicability of the AI-generated ideas. The experts also mentioned that the AI often overlooked deeper, implicit customer issues. As noted by one of the experts:

> AI didn't always pick up on the context or implicit issues that were embedded in the reviews. Say, it might suggest 'fix battery life' but miss the underlying problem of customers using the headsets in extreme environments.—Product Expert C

In the implementation of Stage 4, the three product experts at Simply Headsets expressed a general distrust in AI-generated ideas/solutions. They emphasised the necessity of being able to trace these ideas back to the original product reviews from which they were derived. This traceability was considered crucial for verifying the validity and relevance of the AI's suggestions. Without reference to the actual reviews, there is a risk of accepting solutions that may not be grounded in real customer feedback. As one product expert noted, 'having the ability to refer back and verify that a problem is valid based on actual data, versus seeing information that has been misinterpreted, is essential'. Two of the experts explained:

> Yeah, I'd want to see the customer comments that led to that conclusion. You're identifying this as a problem, but if there are multiple answers for the same issue, it makes me distrust the results. I would want to double-check where the information is coming from. If the data appears duplicated or inconsistent, I won't trust it.—Product Expert A

> Focusing on the specific problems of the headset rather than the one the customer currently has is crucial. It's about how the AI interprets and relays

> that information. That's why I think it's necessary to double-check the AI's conclusions.—Product Expert C

The three product experts were also concerned about the occurrence of AI hallucinations—occasions where the AI develops ideas or solutions that were never mentioned in the customer review data. This led to confusion and mistrust among the product development team. One expert humorously pointed out, 'I'm pretty sure no one wrote a review about their headset exploding', highlighting the absurdity of some AI-generated issues. Such hallucinations underscore the importance of having human oversight to ensure the AI's outputs are accurate and relevant. Another expert added, 'these AI-generated problems can sometimes be so off-base that it makes us question the reliability of the entire system. If the AI is producing issues that have no basis in the actual customer feedback, we can't trust its other conclusions either'. This sentiment was echoed by another expert who emphasised the need for a verification process: 'We need to have a way to trace back every AI-generated suggestion to the original customer comments. Without this transparency, we can't confidently act on the AI's recommendations'. As noted by two of the experts:

> And if you're going to invest millions of dollars in developing a new headset or making a significant change, you want to ensure that it's relevant. So, yes, humans need to be included in the process.—Product Expert A

> Even if there was a review about the battery expanding, it was just one out of 300. If you expand the sample to 1,000 or 10,000, that issue might only appear 2 or 3 times. Do you really want to focus on such an extreme outlier? Without us or someone else actually reading through the reviews, you might end up focusing on the wrong problem.—Product Expert B

The three product experts at Simply Headsets also noted that while some AI-generated concrete ideas/solutions were specific and feasible, many were repetitive or incomplete. The experts stressed the need for AI to produce more refined and applicable solutions. Despite these drawbacks, they acknowledged that AI could still be useful in generating a large volume of ideas quickly, especially when dealing with extensive datasets. The efficiency of AI in processing and summarising information from over 10 000 product reviews was seen as a significant advantage in terms of remaining consistent and objective with ideas/solutions amidst vast amounts of data. The experts recognised that AI-generated ideas could serve as a valuable foundation for product innovation discussions. However, they emphasised that human interpretation is necessary to refine these ideas into practical solutions. AI can provide a broad range of suggestions, but it takes human expertise to evaluate, adapt and implement them effectively. This collaboration between AI and human experts can lead to more innovative and feasible product developments, as stated by two experts:

> Yeah, I mean, it's easier for AI to process 3,300 reviews and come up with 60 ideas, then narrow it down to 20 key ideas to focus on. AI does the hard work of sorting through the mass amounts of data and finding commonalities. Then, humans can step in to fact-check those findings, rather than starting with 3,300 reviews themselves.—Product Expert A

> Yeah, I think it would. The groundbreaking, visionary ideas are important, even though 90% of the commonsense stuff has probably already been considered. It's the remaining 10% of innovative ideas that we haven't explored yet. For example, self-healing material might not be possible, but it could spark other innovative ideas.—Product Expert B

When further refining the groundbreaking solutions, the experts highlighted AI as a starting point for truly innovative product development while also emphasising the need for human refinement and practical adjustments. One expert remarked on the value of AI-generated groundbreaking ideas/solutions as stepping stones to innovation, appreciating the creative prompts provided by AI, which can inspire feasible alternatives and improvements. Another elaborated that the breakthrough ideas and solutions were somewhat feasible and, with a bit of work, could be successfully implemented into new products. As explained by two of the experts:

> I would say they are a good stepping stone for you to maybe say, yeah, this sounds good, but maybe we don't have self-healing material. Instead, we have more robust materials that don't break as much.—Product Expert C

> Just things that looked, you know, they weren't so far-fetched that you couldn't see them becoming a reality. Things like self-healing materials for battery compartments are really beneficial, even if they're a bit out there.—Product Expert A

5.3 | Framework Evaluation

To evaluate the efficacy of the implemented solution, we assessed the performance of the proposed methods (see Table 3). GPT-4 Turbo was employed for its zero-shot learning capabilities to identify product attributes in each review. A random sample of 307 reviews was independently annotated by three domain experts external to Simply Headsets. Only reviews with consistent labels across all experts were retained, resulting in a benchmark dataset of 270 validated reviews. GPT-4 Turbo's performance was evaluated on this dataset using four standard multi-label classification metrics: Accuracy (0.910), Hamming Loss (0.090), Micro-AUC (0.878) and Micro-F1 (0.717) (Zhang and Zhou 2013). We also benchmarked GPT-4 Turbo against other models, such as Claude-3 and GPT-3.5, across metrics including processing time of labelling 307 reviews in a single thread (we parallelised 100 threads to label all reviews in

2 h). GPT-4 Turbo outperformed both models, particularly in generating precise, nuanced responses and relatively efficient handling of large datasets, making it better suited for complex, high-demand applications.

To evaluate the outputs of our framework, another group of design experts from Simply Headset was invited. Questionnaires were sent to three independent experts to evaluate the problems and solutions generated by AI. We received valid responses from all of them. Then, a focus group was organised post-evaluation to finalise the selected problems and solutions, which will be taken into consideration in the new private-label product design. In the questionnaire for problem evaluation, we included the 30 problems identified by AI and selected by product experts in the first and second stages and the 30 solutions identified by AI and selected by product experts in the third and fourth stages.

For each problem, the design experts answered six questions using a five-point Likert scale for evaluation. For each problem, we calculated an average score (ranging from 5 to 30) based on the evaluations of the human experts. The average score from the three experts was 25.4 out of 30 (with individual scores of 29.2, 27.6 and 19.5, respectively). This result indicated that the problems identified by AI and selected by humans received a positive evaluation from the design experts (see Table 4).

For each solution, the design experts answered four questions (see Table 4), which were adopted from Yin et al. (2023). A five-point Likert scale was used for the evaluation. We calculated an average score (ranging from 4 to 20) for each solution based on the experts' evaluations. For the four questions, the average score from the three experts was 15.2 out of 20 (with individual scores of 11.2, 14.2 and 20.2, respectively). This result indicated that the solutions generated by AI and selected and refined by humans received positive evaluations from the design experts.

6 | Discussion and Conclusion

In this paper, we propose the HAIAPI framework, designed to address the unique challenges small businesses face in AI adoption, such as limited financial resources, data infrastructure technical expertise, as well as cultural and operational barriers. The GPT-based framework integrates human–AI collaboration across four stages of product innovation: (1) AI-augmented problem articulation, (2) human expert problem selection, (3)

TABLE 3 | Performance evaluation of GPTs.

Model	Accuracy	Hamming loss	Micro-AUC	Micro-F1	Time (min)
GPT4-Turbo	**0.910**	**0.090**	**0.878**	**0.717**	113.1
GPT3.5-Turbo	0.877	0.123	0.806	0.610	**31.2**
Claude-3-opus	0.899	0.101	0.877	0.696	152.2

TABLE 4 | Evaluation questions for problems and solutions detected and generated by AI and selected and refined by humans.

Evaluation domain	Evaluation dimensions	Questions
AI-detected problems	Customer impact	Q1: How urgent is it to address this problem to prevent customer dissatisfaction or loss?
	Business impact	Q2: How critical is it to resolve this problem to avoid negative business impacts?
	Customer needs and preferences	Q3: How significant is this problem in addressing customer needs and preferences?
	Resources allocation	Q4: How feasible is it to allocate the necessary resources to solve this problem?
	Cost vs. benefit	Q5: How does the cost of addressing this problem compare to the expected benefits?
	timeframe	Q6: How realistic is the timeframe for addressing this problem?
AI-generated solutions	Breakthrough Potential	Q1: Do you agree that this is a breakthrough innovative idea for product modification/improvement?
	Usefulness	Q2: Do you agree that this idea will be useful, i.e., have a practical utility, for product modification/improvement in practice?
	Feasibility	Q3: Do you agree that this idea will be feasible, i.e., capable of being done, in practice?
	Adoption	Q4: Would you add or recommend this idea to future product development or improvement work?

AI-augmented solution generation and (4) human expert solution selection. The implementation of the HAIAPI framework offered three tangible benefits to Simply Headsets, particularly in optimising their product development efficiency and strategic decision making. *First*, by leveraging AI-augmented problem articulation, Simply Headsets can systematically analyse vast amounts of customer feedback to identify product issues, which might have been previously overlooked due to manual efficiencies. This data-driven approach to identifying pain points in product performance, design and user experience led to targeted product improvements and innovations aligning with customer needs. *Second*, the framework's AI-assisted solution generation improves Simply Headsets' ideation process by rapidly producing innovative and feasible product design alternatives. The company can use this GPT-based approach to generate solutions tailored to specific customer concerns, such as improved battery life or enhanced microphone quality; thus, reducing the risk of costly trial-and-error methods for new product designs. Third, the structured human–AI collaboration embedded in the HAIAPI framework fosters a balance between AI-generated insights and human expert judgements in Simply Headsets. The company can make sure that the product solutions generated by GPT are in line with its business objectives and market demands, such as a focus on the use of sustainable materials, leading to higher customer satisfaction and increased market competitiveness.

6.1 | Principles for Human–AI Collaboration in Product Innovation for Small Businesses

Based on our insights from Simply Headsets, we identified three main principles of human–AI collaboration offering practical further guidance for small businesses aiming to integrate AI into their product innovation processes, summarised in Table 5. These principles highlight the need for a balanced and structured approach to human–AI collaboration that enables small businesses to effectively and efficiently integrate AI while mitigating risk and maximising its potential benefits.

First, AI and human expertise must work in tandem to maximise efficiency, drive innovation and ensure feasibility of AI-generated ideas. Human–AI collaboration should be built on complementarity, where AI provides computational efficiency while human expertise ensures contextual understanding (Bouschery et al. 2023). Small businesses often lack the financial and technical resources to experiment with extensive AI-driven product development. AI provides an opportunity to process vast amounts of data quickly, but without human oversight, generated insights can be misaligned with real customer needs or business priorities. Simply Headsets demonstrated that while AI identified frequent complaints about headset weight, human experts recognised that improving battery life would yield a better advantage over currently available headsets on the market. Furthermore, AI proposed groundbreaking solutions including futuristic headset features, such as voice-activated noise cancellation, but human decision-makers prioritised refining microphone quality as a more cost-effective and immediate solution. The human–AI complementarity at Simply Headsets enabled efficient resource allocation, focusing on high-impact and actionable improvements—crucial for preserving the company's

TABLE 5 | Principles for human–AI collaboration in product innovation.

#	Principle (what?)	Rationale (why?)	Enabling actions (how?)	Examples of Implementation
1	Human–AI complementarity for strategic decision-making	AI efficiently analyses data but lacks strategic judgement, requiring human oversight to align insights with business goals.	Use AI to process customer feedback while human experts validate, interpret, and prioritise improvements based on feasibility and impact.	Simply Headsets used AI to detect common complaints but relied on human experts to prioritise battery life improvements over minor headset weight issues.
2	Refinement and validation of AI-generated Insights	AI outputs need continuous refinement to enhance accuracy, ensuring that small businesses avoid costly misinterpretations.	Establish feedback loops to iteratively refine AI outputs with human validation and cross-functional collaboration.	Simply Headsets refined AI prompts multiple times to distinguish overlapping product issues in customer reviews.
3	Transparency in AI processes for trust and adoption	Ensuring AI outputs are interpretable builds trust, enabling businesses to use AI effectively in decision-making.	Implement traceability mechanisms that link AI findings to original data sources, enhancing accountability and reducing errors.	Product experts at Simply Headsets ensured AI-suggested issues could be traced back to original customer reviews for validation.

limited resources. In conclusion, this principle is crucial for small businesses to avoid investing in impractical product improvements or innovations and to ensure that AI-driven insights lead to tangible business value.

Second, AI systems require iterative refinement to enhance accuracy and relevance while mitigating the risks of AI 'hallucinations'—fabricated insights that may mislead decision-making (Rudin 2019). Small businesses, unlike larger firms with dedicated AI teams, cannot afford costly trial-and-error approaches, making precision essential. In Simply Headsets, AI produced vague or misleading problem statements, such as broadly categorising customer dissatisfaction with connectivity, in its early iterations. However, through an iterative refinement process that included prompt adjustments and human validation, Simply Headsets improved the outputs generated, enabling the identification of specific issues—such as Bluetooth pairing failures and interference with multiple devices—leading to more targeted and effective solutions. Moreover, AI sometimes misidentified non-existent issues, such as headset discomfort, which upon review, was found to be a rare occurrence. By systematically cross-referencing AI findings with human analysis, Simply Headsets avoided unnecessary product modifications and directed development resources towards solving actual, high-priority customer issues. In conclusion, for small businesses, this principle highlights the importance of refining AI outputs at multiple stages of the product innovation process to maintain their reliability, actionability, and alignment with real market needs.

Third, transparency in AI processes is critical for building trust and ensuring accountability in decision-making (Dolata et al. 2022). Small businesses often lack the in-house technical expertise to fully comprehend AI-driven insights, increasing the risk of misinterpretation. Without transparency, AI-generated product problem and solution recommendations can lead to misinformed decisions, either by being blindly trusted or outright dismissed due to uncertainty. Ensuring traceability of AI findings builds confidence in its outputs and enhances decision-making credibility. Simply Headsets encountered this challenge when AI prioritised minor complaints due to their high frequency in customer reviews. By implementing traceability mechanisms, human experts could verify AI conclusions against original data sources, ensuring that only the most relevant issues were acted upon. This transparency fostered greater trust in AI recommendations and encouraged its broader adoption across the organisation. In conclusion, this principle emphasises that for small businesses, embedding clear and interpretable AI processes is crucial for maximising AI's value while preventing errors that could result in misinformed decisions or resistance to AI integration.

6.2 | Limitations and Future Research

Despite the promising results, our research has several limitations. The scalability and adaptability of the HAIAPI framework to different business environments and domains require further experiments and validation through additional case studies. While our case study on the headset industry demonstrates the framework's potential, additional case studies across diverse sectors are necessary to validate its broader applicability. Another limitation is the reliance on manual prompt engineering when utilising GPTs, which relies on domain experts' expertise and reduces the framework's level of automation. Future research could explore automated prompt engineering (APE) techniques to mitigate this challenge and enhance the framework's generalisability across different use cases. These enhancements could further streamline the innovation process, making AI-driven insights more accessible and practical for a wide range of industries.

Importantly, the foundational tenets of the HAIAPI framework—(1) iterative human–AI collaboration, (2) contextualised prompting and (3) systematic validation—are subject to adaptation as AI capabilities advance. *First*, iterative human–AI collaboration ensures that generative AI augments human creativity and decision-making rather than replacing it entirely. As AI technologies advance, this collaborative interaction might shift from structured prompting to more intuitive, conversational engagements, allowing users to interact naturally without extensive prompt engineering (Amershi et al. 2019). *Second*, contextualised prompting enables AI-generated outputs to closely align with specific organisational contexts, tasks and objectives. Currently, detailed prompts are meticulously crafted; however, with future AI models gaining enhanced contextual awareness and sophisticated conversational capabilities, less rigid prompting methods may emerge (Radford et al. 2019). For example, organisations might increasingly rely on general guidelines, allowing AI to infer specific requirements from broader instructions, thus simplifying the interaction process. *Third*, systematic validation mechanisms address risks such as misinformation and inaccuracies in AI-generated outputs. Advances in model transparency and interpretability could significantly streamline validation processes (Rudin 2019). For instance, more transparent AI models may enable automated or semi-automated validation, reducing manual checking and allowing practitioners to focus primarily on strategic innovation aspects rather than routine AI output verification.

Although these tenets form a robust conceptual foundation, practitioners must remain agile, proactively adapting their application to evolving AI capabilities. Continuous learning, iterative experimentation and openness to technological updates will be essential. The sustained effectiveness of the HAIAPI framework relies on its flexibility to integrate emerging technological affordances while consistently preserving its core intent: fostering responsible, contextually relevant and impactful AI-driven innovation in small business environments.

6.3 | Conclusion

User-generated data, such as customer reviews and social media posts, offer a valuable yet underutilised resource for businesses seeking to enhance their innovation processes. However, challenges such as noise, sparsity and ambiguity in online texts pose significant obstacles for traditional NLP tools. In this paper, we addressed these challenges by leveraging advanced AI techniques, including ML and GPTs, to develop a comprehensive framework for extracting actionable insights from large-scale online reviews. Our framework aims to facilitate open innovation by enabling businesses, especially small enterprises, to systematically incorporate user-generated insights into product development.

Our contributions to the field of AI in innovation are threefold: (1) we present a structured, scalable approach for automating the extraction and generation of innovation-relevant insights from large-scale textual data; (2) we demonstrate the potential of AI in overcoming resource constraints typical of small businesses, enhancing both the quantity and quality of innovation outputs and (3) we underscore the role of human–AI collaboration in navigating complex problem-solving tasks, providing a balanced framework that integrates AI-driven insights with human expertise. We advance the state of the art by highlighting how AI can be strategically deployed across various stages of the innovation process—specifically problem identification, solution generation and evaluation.

Acknowledgements

This research was supported by the Australian Research Council, Discovery Project scheme (DP230102657). We gratefully acknowledge their financial support, which made this study possible. Open access publishing facilitated by The University of Melbourne, as part of the Wiley - The University of Melbourne agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

Endnotes

[1] A small business is defined as having fewer than 20 employees and an annual turnover of less than $10 million. *Source:* Australian Taxation Office (ATO). (2021). *Definitions—Small business entity.* Retrieved from https://www.ato.gov.au/.

[2] https://www.autodesk.com/au/collections/product-design-manufacturing/overview?term=1-YEAR&tab=subscription.

[3] Refer to Appendix A: Table A1 for details on the research method.

[4] Refer to Tables B1 and B2 for the resource toolkit and checklist questions and GPT prompts for implementing the HAIAPI framework, respectively.

[5] Refer to Table B1 for GPT prompts relating to Step 2.

[6] Refer to Table B1 for GPT prompts relating to Step 3.

[7] Refer to Table B1 for GPT prompts relating to Step 4.

[8] Refer to Table C1 for samples from the most relevant AI-detected headset problems selected by the product experts.

[9] Refer to Table B1 for GPT prompts relating to Step 6.

[10] Refer to Table C2 for a sample from an AI-generated solution.

[11] Refer to Table B1 for GPT prompts relating to Step 7.

[12] Refer to Table C3 for samples from further refinement of groundbreaking AI-generated solutions.

References

Aldunate, Á., S. Maldonado, C. Vairetti, and G. Armelini. 2022. "Understanding Customer Satisfaction via Deep Learning and Natural Language Processing." *Expert Systems With Applications* 209: 118309. https://doi.org/10.1016/j.eswa.2022.118309.

Alegre, J., J. Pla-Barber, R. Chiva, and C. Villar. 2012. "Organisational Learning Capability, Product Innovation Performance and Export Intensity." *Technology Analysis & Strategic Management* 24, no. 5: 511–526. https://doi.org/10.1080/09537325.2012.674672.

Amershi, S., D. Weld, M. Vorvoreanu, et al. 2019. "Guidelines for Human-AI interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, Glasgow Scotland, UK."

Arthur, D., and S. Vassilvitskii. 2007. "K-Means++: The Advantages of Careful Seeding."

Berry, G., and K. M. Shabana. 2020. "Adding a Strategic Lens to Feasibility Analysis." *New England Journal of Entrepreneurship* 23, no. 2: 67–78. https://doi.org/10.1108/NEJE-08-2019-0036.

BMW Group. 2023. *How Is AI Revolutionising Production*. BMW Group. https://www.bmwgroup.com/en/news/general/2023/aiqx.html.

Bouschery, S. G., V. Blazevic, and F. T. Piller. 2023. "Augmenting Human Innovation Teams With Artificial Intelligence: Exploring Transformer-Based Language Models." *Journal of Product Innovation Management* 40, no. 2: 139–153. https://doi.org/10.1111/jpim.12656.

Brougham, D., and J. Haar. 2018. "Smart Technology, Artificial Intelligence, Robotics, and Algorithms (STARA): Employees' Perceptions of Our Future Workplace." *Journal of Management & Organization* 24, no. 2: 239–257. https://doi.org/10.1017/jmo.2016.55.

Brown, T. 2008. "Design Thinking." *Harvard Business Review* 86, no. 6: 84–92. https://readings.design/PDF/Tim%20Brown,%20Design%20Thinking.pdf.

Chan, C. M., S. Y. Teoh, A. Yeow, and G. Pan. 2019. "Agility in Responding to Disruptive Digital Innovation: Case Study of an SME." *Information Systems Journal* 29, no. 2: 436–455. https://doi.org/10.1111/isj.12215.

Chiesa, V., F. Frattini, V. Lazzarotti, and R. Manzini. 2009. "Performance Measurement of Research and Development Activities." *European Journal of Innovation Management* 12, no. 1: 25–61. https://doi.org/10.1108/14601060910928166.

Creery, J. 2024. "Nike Is Too Slow to Innovate, Says JD Sports Boss." *Bloomberg News*. https://www.bloomberg.com/news/articles/2024-03-28/nike-is-too-slow-to-innovate-says-jd-sports-boss?embedded-checkout=true.

Design Council. 2015. "The Double Diamond: A Universally Accepted Depiction of the Design Process." https://www.designcouncil.org.uk/our-resources/the-double-diamond/.

Dolata, M., S. Feuerriegel, and G. Schwabe. 2022. "A Sociotechnical View of Algorithmic Fairness." *Information Systems Journal* 32, no. 4: 754–818. https://doi.org/10.1111/isj.12370.

Dorst, K. 2011. "The Core of 'Design Thinking' and Its Application." *Design Studies* 32, no. 6: 521–532. https://doi.org/10.1016/j.destud.2011.07.006.

Duan, Y., J. S. Edwards, and Y. K. Dwivedi. 2019. "Artificial Intelligence for Decision Making in the Era of Big Data–Evolution, Challenges and Research Agenda." *International Journal of Information Management* 48: 63–71. https://doi.org/10.1016/j.ijinfomgt.2019.01.021.

Ehringfeld, I., P. Steinhoff, S. Störk, A. Whittal, and O. R. Herber. 2023. "The Voting Scheme Determines the Winner: Considering Possible Options for Designing a Delphi Study." *International Journal of Qualitative Methods* 22: 16094069221148871. https://doi.org/10.1177/16094069221148871.

Forsman, H. 2011. "Innovation Capacity and Innovation Development in Small Enterprises: A Comparison Between the Manufacturing and Service Sectors." *Research Policy* 40, no. 5: 739–750. https://doi.org/10.1016/j.respol.2011.02.003.

Ghezzi, A., and A. Cavallo. 2020. "Agile Business Model Innovation in Digital Entrepreneurship: Lean Startup Approaches." *Journal of Business Research* 110: 519–537. https://doi.org/10.1016/j.jbusres.2018.06.013.

Heidenreich, S., and T. Kraemer. 2016. "Innovations—Doomed to Fail? Investigating Strategies to Overcome Passive Innovation Resistance." *Journal of Product Innovation Management* 33, no. 3: 277–297. https://doi.org/10.1111/jpim.12273.

Hou, Y., J. Li, Z. He, A. Yan, X. Chen, and J. McAuley. 2024. "Bridging Language and Items for Retrieval and Recommendation." arXiv Preprint arXiv:2403.03952. https://arxiv.org/abs/2403.03952.

Kanbach, D. K., L. Heiduk, G. Blueher, M. Schreiter, and A. Lahmann. 2024. "The GenAI Is Out of the Bottle: Generative Artificial Intelligence From a Business Model Innovation Perspective." *Review of Managerial Science* 18, no. 4: 1189–1220. https://doi.org/10.1007/s11846-023-00696-z.

Klein, H. K., and M. D. Myers. 1999. "A Set of Principles for Conducting and Evaluating Interpretive Field Studies in Information Systems." *MIS Quarterly* 23, no. 1: 67–93. https://doi.org/10.2307/249410.

Kohli, R., and N. P. Melville. 2019. "Digital Innovation: A Review and Synthesis." *Information Systems Journal* 29, no. 1: 200–223. https://doi.org/10.1111/isj.12193.

Mandviwalla, M., and R. Flanagan. 2021. "Small Business Digital Transformation in the Context of the Pandemic." *European Journal of Information Systems* 30, no. 4: 359–375. https://doi.org/10.1080/0960085X.2021.1891004.

Martorelli, M., and A. Gloria. 2023. "Strategies and Generative Design Towards the Development of Innovative Products." In *Springer Handbook of Additive Manufacturing*, 269–286. Springer. https://doi.org/10.1007/978-3-031-20752-5_17.

Mihalcea, R., C. Corley, and C. Strapparava. 2006. "Corpus-Based and Knowledge-Based Measures of Text Semantic Similarity." AAAI.

Noori, B. 2021. "Classification of Customer Reviews Using Machine Learning Algorithms." *Applied Artificial Intelligence* 35, no. 8: 567–588. https://doi.org/10.1080/08839514.2021.1922843.

Novak, L. 1996. "Market and Feasibility Studies: A How-to Guide." University of Oregon. https://pages.uoregon.edu/rgp/PPPM613/downloads/How%20to%20do%20a%20Market%20Analysis.pdf.

Orrù, G., A. Piarulli, C. Conversano, and A. Gemignani. 2023. "Human-Like Problem-Solving Abilities in Large Language Models Using ChatGPT." *Frontiers in Artificial Intelligence* 6: 1199350. https://doi.org/10.3389/frai.2023.1199350.

Patnaik, D., and R. Becker. 1999. "Needfinding: The Why and How of Uncovering People's Needs." *Design Management Journal* 10, no. 2: 37–43. https://doi.org/10.1111/j.1948-7169.1999.tb00250.x.

Prosser, D. 2022. "How Proven Skincare Brought Science to the Art of Beauty." Forbes Media LLC. https://www.forbes.com/sites/dianatsai/2021/10/25/this-is-no-ordinary-unicorn-this-is-a-good-unicorn-how-notco-is-saving-the-planet-by-making-plants-taste-like-meat/.

Radford, A., J. Wu, R. Child, D. Luan, D. Amodei, and I. Sutskever. 2019. "Language Models Are Unsupervised Multitask Learners." OpenAI Blog. https://cdn.openai.com/better-language-models/language_models_are_unsupervised_multitask_learners.pdf.

Ransbotham, S., P. Gerbert, M. Reeves, D. Kiron, and M. Spira. 2018. "Artificial Intelligence in Business Gets Real." https://sloanreview.mit.edu/projects/artificial-intelligence-in-business-gets-real/.

Rudin, C. 2019. "Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead." *Nature Machine Intelligence* 1, no. 5: 206–215. https://doi.org/10.1038/s42256-019-0048-x.

Schwaeke, J., A. Peters, D. K. Kanbach, S. Kraus, and P. Jones. 2024. "The New Normal: The Status Quo of AI Adoption in SMEs." *Journal of Small Business Management* 63, no. 3: 1–35. https://doi.org/10.1080/00472778.2024.2379999.

Thormundsson, B. 2024. "Adoption of AI in Organizations 2022-2023, by Region." https://www.statista.com/statistics/1472617/ai-adoption-companies-by-region/.

Verganti, R., L. Vendraminelli, and M. Iansiti. 2020. "Innovation and Design in the Age of Artificial Intelligence." *Journal of Product Innovation Management* 37, no. 3: 212–227. https://doi.org/10.1111/jpim.12523.

Walsham, G. 2006. "Doing Interpretive Research." *European Journal of Information Systems* 15, no. 3: 320–330. https://doi.org/10.1057/palgrave.ejis.3000589.

Yin, C., C. Jiang, H. K. Jain, Y. Liu, and B. Chen. 2023. "Capturing Product/Service Improvement Ideas From Social Media Based on Lead User Theory." *Journal of Product Innovation Management* 40, no. 5: 630–656. https://doi.org/10.1111/jpim.12676.

Yin, R. K. 2009. *Case Study Research: Design and Methods*. Vol. 5. Sage.

Zachlod, C., O. Samuel, A. Ochsner, and S. Werthmüller. 2022. "Analytics of Social Media Data–State of Characteristics and Application." *Journal of Business Research* 144: 1064–1076.

Zhang, M.-L., and Z.-H. Zhou. 2013. "A Review on Multi-Label Learning Algorithms." *IEEE Transactions on Knowledge and Data Engineering* 26, no. 8: 1819–1837.

Appendix A

Research Method

This study aims to contribute to a better understanding of human–AI collaboration in the context of product innovation within small firms. Given the scarcity of practical research on this topic, our study is exploratory and employs qualitative methods. Our primary data are derived from a single revelatory case study (Yin 2009) and our analysis follows an interpretive approach (Walsham 2006). Simply Headsets serves as a revelatory case due to its successful implementation of a human–AI assisted innovation framework to drive its new product development initiative.

The primary data sources for this case study included: (a) 10 h of intensive collaborative work with product and design expert teams from Simply Headsets to implement the HAIAPI framework, (b) 5 h of focus groups with one group of product experts and another group of design experts from Simply Headsets and (c) 2 h of in-depth interview with the founder of Simply Headsets who also actively participated throughout the different stages of this study, including problem identification, data collection, insights development and further refinement. The primary data sources were supplemented with other relevant secondary sources such as internal company documents and presentations, news reports and press releases. For our data analysis, we adopted an interpretive approach (Klein and Myers 1999).

We began by developing narratives to describe the human–AI collaboration journey at Simply Headsets. This narrative approach helped us understand the key decisions related to the implementation of the HAIAPI framework at Simply Headsets. These narratives, in turn, facilitated a deeper understanding of the practices that enabled the small firm to successfully implement human–AI collaboration for product innovation purposes. Initially, our analysis focused on identifying general leadership principles for AI adoption. We then considered not only deliberate management actions but also organisational circumstances, routines and culture that facilitated the implementation. We iterated between emerging insights, data and literature multiple times to continuously deepen our understanding. This iterative process initially resulted in a set of eight principles that we later combined into six principles, and then further synthesised into three main principles.

TABLE A1 | Primary data sources details.

Data collection	Participants
Collaboration Work for Implementation of HAIAPI Framework	Project Team 1: Product Expert A, Product Expert B, Product Expert C Project Team 2: Design Expert D, Design Expert E, Design Expert F
Focus Group 1 (HAIAPI Framework Operationalisation and Implementation)	Product Expert A, Product Expert B, Product Expert C
Focus Group 2 (HAIAPI Framework Evaluation)	Design Expert D, Design Expert E, Design Expert F
Interview	Founder/manager

Appendix B

Resource Toolkit

TABLE B1 | Resource toolkit for implementing the HAIAPI framework.

Innovation space	HAIAPI framework/resources: Framework stages	Framework steps	Resources	Checklist questions
Product Problem	Stage 1: AI-augmented product problem	Step 1: Data access, collection and preprocessing	Estimated Budget = $147 Personnel: • Entry-level programmer ($49 per hour for 3 h) to create Python script for data pre-processing. Technology: • Free access to Kaggle (https://www.kaggle.com/datasets), UCI Machine Learning Repository (https://archive.ics.uci.edu/datasets) or Papers with Code (https://paperswithcode.com/datasets) to collect relevant product review datasets; • Free access to NLTK (https://www.nltk.org) or SpaCy (https://spacy.io) to filter and segment the collected data.	Budget: • Have you allocated sufficient funds for accessing the dataset (if it is not free) and for hiring or assigning a programmer to preprocess data? Personnel: • Have you hired or assigned a programmer, who is proficient in Python and NLTK or SpaCy, to develop and test scripts for data preprocessing? Technology: • Have you identified specific data source(s) you will use for data collection? • Is the data volume sufficient for meaningful analysis? • Have you installed and configured Python and NLTK or SpaCy? Process: • Have you created test cases for Python scripts to validate accurate data preprocessing? • Have you tested Python scripts for data preprocessing on a small sample dataset?
		Step 2: Data labelling	Estimated Budget = $2376 Personnel: • Entry-level programmer ($49 per hour for 6 h) to create Python script and prompt GPT-4 Turbo to label large dataset; • Three product experts ($55 per hour, each working 7.5 h) to label sample dataset and help refine GPT-4 Turbo prompts. Technology: • GPT-4 Turbo (https://openai.com/chatgpt) (April 2024 model, $845) to assist with data labelling.	Budget: • Have you allocated sufficient funds for accessing GPT-4 Turbo's API and for hiring or assigning a programmer and three product experts to label data? Personnel: • Have you hired or assigned a programmer, who is proficient in Python and experienced with GPT-4 Turbo's API, to develop and test scripts and prompts for large dataset labelling? • Have you identified and hired or assigned three suitable product experts to manually label sample dataset and refine GPT-4 Turbo's prompts? Technology: • Have you installed and configured Python (if you haven't already)? • Have you configured access to GPT-4 Turbo's API? Process: • Have you created test cases for Python scripts to validate accurate data labelling? • Have you tested Python scripts for data labelling using a small sample dataset? • Are the data labelling criteria clearly defined to ensure consistency and accuracy? • Have you verified that GPT-4 Turbo's prompts are optimised for data labelling?

(Continues)

TABLE B1 | (Continued)

	HAIAPI framework/resources			
Innovation space	Framework stages	Framework steps	Resources	Checklist questions
		Step 3: Atomic problem extraction	Estimated Budget = $227 Personnel: • Entry-level programmer ($49 per hour for 2 h) to create Python script and prompts for GPT-4 Turbo to extract atomic problems from large dataset. Technology: • GPT-4 Turbo (April 2024 model, $129) to deconstruct identified issues into core components, improving the clarity and focus of each problem.	Budget: • Have you allocated sufficient funds for accessing GPT-4 Turbo's API and for hiring or assigning a programmer to extract atomic problems? Personnel: • Have you assigned a programmer, who is proficient in Python and experienced with GPT-4 Turbo's API, to develop and test scripts and prompts for atomic problem extraction? Technology: • Have you installed and configured Python (if you haven't already)? • Have you configured access to GPT-4 Turbo's API (if you haven't already)? Process: • Have you created test cases for Python scripts to validate accurate atomic problem extraction? • Have you tested Python scripts for atomic problem extraction using a small sample dataset? • Have you verified that GPT-4 Turbo's prompts are optimised for atomic problem extraction?
		Step 4: Atomic problem deduplication and summarisation	Estimated Budget = $307 Personnel: • Entry-level programmer ($49 per hour for 6 h) to create Python script for atomic problems embeddings, clustering and deduplication and prompts for GPT-4 Turbo to summarise problems. Technology: • Free access to Scikit-learn (https://scikit-learn.org/stable/) for clustering and removing redundant issues; • Text-Embedding-3-Small (https://platform.openai.com/docs/guides/embeddings) ($1) to represent issues in semantic vectors; • GPT-4 Turbo (April 2024 model, $12) for summarising recurring product issues in a coherent, actionable format.	Budget: • Have you allocated sufficient funds for accessing GPT-4 Turbo's API and Text-Embedding-3-Small's API and for hiring or assigning a programmer to deduplicate and summarise atomic problems? Personnel: • Have you hired or assigned a programmer, who is proficient in Python and experienced with GPT-4 Turbo's API, Text-Embedding-3-Small's API and Scikit-learn, to develop and test scripts and prompts for atomic problem embeddings, clustering, deduplication and summarising? Technology: • Have you installed and configured Python (if you haven't already)? • Have you installed and configured Scikit-learn? • Have you configured Text-Embedding-3-Small's API access? • Have you configured access to GPT-4 Turbo's API (if you haven't already)? Process: • Have you created test cases for Python scripts to validate accurate atomic problem deduplication and summarisation? • Have you tested Python scripts for atomic problem deduplication and summarisation using a small sample dataset? • Have you verified that GPT-4 Turbo prompts are optimised for atomic problem summarisation?

(Continues)

TABLE B1 | (Continued)

Innovation space	Framework stages (HAIAPI framework/resources)	Framework steps	Resources	Checklist questions
	Stage 2: Human expert product problem selection	Step 5: Problem selection and prioritisation	Estimated Budget = $1650 Personnel: • Three product experts ($55 per hour for 10 h each) for problem selection and prioritisation based on alignment with strategic goals. Technology: • Free access to Google Sheets (https://docs.google.com/spreadsheets) to organise and recording the expert ratings for problem selection and prioritisation.	Budget: • Have you allocated sufficient funds for hiring or assigning three product experts to select and prioritise urgent and feasible problems? Personnel: • Have you identified and hired or assigned three suitable product experts to select and prioritise urgent and feasible problems? Technology: • Have you setup Google Sheets for collaborative work between product experts? Process: • Are the selection and prioritisation criteria clearly defined to ensure consistency and accuracy? • Are the selection and prioritisation criteria standardised and documented to facilitate proper organisation and record-keeping of decisions?
Product Solution	Stage 3: AI-augmented product solutions generation	Step 6: AI-generated knowledge base	Estimated Budget = $325 Personnel: • Entry-level programmer ($49 per hour for 4 h) to create Python script to manage dataset in MySQL (https://www.mysql.com) or MongoDB (https://www.mongodb.com) database and prompt GPT-4 Turbo to extract atomic ideas from large dataset. Technology: • GPT-4 Turbo (April 2024 model, $129) to extract potential product solutions for constructing a knowledge base with AI-generated insights; • Free access to Google Sheets to store, organise and prepare atomic ideas for extraction.	Budget: • Have you allocated sufficient funds for accessing GPT-4 Turbo's API and for hiring or assigning a programmer to extract atomic ideas and construct knowledge base? Personnel: • Have you assigned a programmer, who is proficient in Python and experienced with GPT-4 Turbo's API, to develop and test scripts and prompts for extracting atomic ideas and constructing knowledge base? Technology: • Have you installed and configured Python (if you haven't already)? • Have you setup Google Sheets for storing extracted ideas (for smaller datasets)? • Have you installed and configured MySQL or MongoDB database for storing extracted ideas (in case of larger datasets)? • Have you configured access to GPT-4 Turbo's API (if you haven't already)? Process: • Have you created test cases for Python scripts to validate accurate atomic idea extraction and knowledge base construction? • Have you tested Python scripts for atomic idea extraction and knowledge base construction using a small sample dataset? • Have you verified that GPT-4 Turbo's prompts are optimised for atomic ideas extraction?

(Continues)

TABLE B1 | (Continued)

Innovation space	HAIAPI framework/resources: Framework stages	Framework steps	Resources	Checklist questions
		Step 7: Atomic idea retrieval	**Estimated Budget = $228** **Personnel:** • **Entry-level programmer ($49 per hour for 4h) to create Python script for atomic idea embeddings, problem-idea matching and prompt GPT-4 Turbo to retrieve atomic ideas from the database and generate concrete and groundbreaking ideas.** **Technology:** • **Text-embedding-3-small ($1) to represent extracted ideas in semantic vectors;** • **GPT-4 Turbo (April 2024 model, $31) for identifying and distilling granular ideas for potential product solutions.**	**Budget:** • **Have you allocated sufficient funds for accessing GPT-4 Turbo's API and Text-Embedding-3-Small's API and for hiring or assigning a programmer to retrieve atomic ideas?** **Personnel:** • **Have you hired or assigned a programmer, who is proficient in Python and experienced with GPT-4 Turbo's API and Text-Embedding-3-Small's API, to develop and test scripts and prompts for atomic idea embeddings, problem matching, atomic idea retrieval and generation of concrete and groundbreaking ideas?** **Technology:** • **Have you installed and configured Python (if you haven't already)?** • **Have you configured Text-Embedding-3-Small's API access? (if you haven't already)?** • **Have you configured access to GPT-4 Turbo's API (if you haven't already)?** **Process:** • **Have you created test cases for Python scripts to validate accurate atomic idea retrieval?** • **Have you tested Python scripts for atomic idea retrieval using a small sample dataset?** • **Have you verified that GPT-4 Turbo's prompts are optimised for atomic idea retrieval?**
	Stage 4: Human experts select promising solutions	**Step 8: Solution review and selection**	**Estimated Budget = $1672** **Personnel:** • **Three experts ($55 per hour for 10h each) to determine the most viable solutions, utilising expert insights for a strategic assessment of feasibility, resource allocation and alignment with the business objectives.** **Technology:** • **Trello (https://trello.com) ($7.5 per month) and StormBoard (https://stormboard.com) ($15 per month) for real-time collaboration and visualisation of the review and selection of product solutions.**	**Budget:** • **Have you allocated sufficient funds for hiring or assigning three product experts to review and select viable solutions?** **Personnel:** • **Have you identified and hired or assigned three suitable product experts to review and select viable solutions?** **Technology:** • **Have you setup Trello for collaborative work between product experts and StormBoard for visualisation of their collaborative decisions?** **Process:** • **Are the evaluation criteria clearly defined to ensure consistency and accuracy?** • **Are the evaluation criteria standardised and documented to facilitate proper organisation and record-keeping of decisions?**

Note: Each step requires the use of a laptop or PC meeting the following minimum specifications: a processor of 2 GHz (GHz) or faster, 16 GB of RAM and 256 GB of hard disk space. All hourly rates are estimated in Australian dollars and based on average annual salaries sourced from Talent.com (https://au.talent.com).

TABLE B2 | GPT prompts for implementing the HAIAPI framework.

Steps	GPT prompts
Step 2: Data labelling	You are an expert in text annotation and classification. Your job is to perform a multi-label classification task on user reviews for headset. You will be provided with the definitions of 15 labels related to the headset; a detailed guide for executing the task; and a user review text. The given user review can have from zero to multiple labels, so you should independently determine whether it meets each label and return your answer in JSON format. ### Labels Definition **label_1**: label_1 is about Material Quality. Accept when the review mentions aspects related to durability, texture, weight, or robustness of the materials used, such as 'feels sturdy' or 'flimsy material'. **label_2**: label_2 is about Connectivity. Accept when the review mentions ease of pairing, connection stability, range, or compatibility with devices, such as 'connects easily via Bluetooth' or 'good range'. **label_3**: label_3 is about Endurance. Accept when the review mentions the product's ability to withstand extended use or tough conditions, including terms like 'long-lasting' or 'withstands wear and tear'. **label_4**: label_4 is about Sound Quality. Accept when the review mentions clarity, bass, treble, balance, or distortion in audio output, such as 'clear sound' or 'strong bass'. **label_5**: label_5 is about Noise Cancellation. Accept when the review addresses the effectiveness of noise reduction or isolation, such as 'blocks out noise effectively' or 'poor noise isolation'. **label_6**: label_6 is about Comfort. Accept when the review discusses comfort during use, especially for prolonged wear or ergonomic fit, such as 'comfortable for hours' or 'causes discomfort after a while'. **label_7**: label_7 is about Shape/Appearance. Accept when the review mentions visual design, aesthetics, colour, or shape, such as 'sleek design' or 'looks bulky'. **label_8**: label_8 is about Battery. Accept when the review discusses battery life, charging speed, or performance stability, such as 'lasts for 8 h' or 'battery dies fast'. **label_9**: label_9 is about After-sale Support Services. Accept when the review covers customer service, warranty response, or repair experiences, such as 'helpful support team' or 'quick warranty replacement.' **label_10**: label_10 is about Microphone Quality. Accept when the review mentions microphone clarity, sensitivity, or sound pickup, such as 'clear voice on calls' or 'muffled sound.' **label_11**: label_11 is about Controls Design/Function Design. Accept when the review refers to button placement, ease of control, or functionality layout, such as 'easy to operate' or 'buttons are too small'. **label_12**: label_12 is about Portability. Accept when the review mentions ease of transport, weight, or compactness, such as 'easy to carry around' or 'too bulky for travel'. **label_13**: label_13 is about Price. Accept when the review discusses value for money, affordability, or pricing perception, such as 'worth the price' or 'overpriced'. **label_14**: label_14 is about Water Proof/Resistant. Accept when the review mentions water resistance, waterproof features, or durability in wet conditions, such as 'waterproof' or 'not suitable for rain'. **label_15**: label_15 reflects whether the evaluation of the headset is useful to improve or innovate the product. Accept when the review provides constructive feedback or suggestions for product enhancements, such as 'could improve sound quality' or 'better battery would be nice'. ### Annotation Instructions **Step 1. Initialize Annotation Structure:** – Create an empty dictionary 'response' to store the analysis and result for each label. – For each label (label_1, label_2, ..., label_15), create an entry in the dictionary containing 'analysis' and 'result'. **Step 2. Independently Evaluate Each Label Following this Step-by-step Workflow:** – **Comprehend the Label Definition:** Carefully read the definition of label_i. – **Comprehend the User Review:** Carefully read the text to be annotated. – **Record Matching Analysis in 'response[f"label_{i}"]["analysis"]':** You should independently assess whether the review text aligns with the definition of label_i, without being influenced by other labels. For label_i, you should quote specific segments from the review text to explain whether the review matches label_i's definition. If the review aligns with label_i's definition, record 'Conclusion: Accepted' at the end of your analysis. Otherwise, record 'Conclusion: Rejected'. **Record Matching Result in 'response[f"label_{i}"]["result"]':** Based on the 'Conclusion' in the Matching Analysis above, select 'Accepted' or 'Rejected'. **Step 3. Generate Final JSON Result:** – After completing the annotation for all labels, convert the 'response' dictionary into a JSON format, and return this JSON object as the model output.

(Continues)

TABLE B2 | (Continued)

Steps	GPT prompts
Step 3/6: Atomic problem/idea extraction	You will be provided with a user review of a headset product from Amazon. As an expert in the headset industry, your task is to identify, extract, and rephrase the atomic issues and atomic ideas related to the ${Product Aspect} of this product from the review text. Atomic issues refer to problems, defects, or shortcomings related to product design or development, while atomic ideas refer to ideas, suggestions, or solutions that can guide product improvement or innovation. A review may raise multiple issues or ideas about ${Product Aspect} from different perspectives, so here 'atomic' means that the description of each issue or idea should be specific and independent. You should store your response in lists and return them in a JSON format, i.e., "'json "${Product Aspect}": { "atomic issues":["issue_1", "issue_2",...], "atomic ideas": ["idea_1", "idea_2",...]}"' Each atomic issue and atomic idea should be faithful to the original review text, retaining detailed descriptions or examples. However, if the given review does not contain specific issues or ideas, the corresponding list should be set to ["none"].
Step 4: Problem summarisation	As an expert in the headset industry, you will receive multiple headset issues reported by users in XML format. Your task is to analyse these texts, and then summarise the 3 to 5 most critical issues that are particularly related to the ${Product Aspect} and would provide the greatest value for product improvement. You should return your summarised issues in XML format, as follows: "'xml <summarized_issues> <issue_1>${The detailed, exhaustive, and clear description of issue-1, along with the specific examples and wording from original content to support this issue}. This issue is based on: ${issue_id1}, ${issue_id2}, ...</issue_1> <issue_2>${The detailed, exhaustive, and clear description of issue-2, along with the specific examples and wording from original content to support this issue}. This issue is based on: ${issue_id1}, ${issue_id2}, ...</issue_2> ...... </summarized_issues>"'
Step 7: Atomic idea retrieval	You are a creative and experienced headset developer, designing improvement plans and innovative directions for your products based on user-reported issues. You will be provided with a description of user-reported issues and several user suggestions that may be related to these issues. Your task is to think up ideas for the headsets' improvement and innovation. The ideas you propose should include the following two categories: 1. Concrete, specific, and feasible ideas: These should be highly detailed and actionable, providing clear and direct guidance for resolving existing issues and improving user experiences. 2. 2. Groundbreaking, visionary, and potentially revolutionary ideas: You can give free rein to your creativity and imagination, thinking beyond existing practices and patterns in the headset industry.

Appendix C

Samples of AI-Generated Output

TABLE C1 | Samples from the most relevant AI-detected headset problems selected by the product experts.

Product aspects	Problem description (AI)	Justification for selection (product experts)
Aftersales support services	Users express frustration over the inability to obtain necessary replacement parts or appropriate repair services. Specific examples include Corsair only selling original ear pads, not compatible with older models (issue_13089) and the unavailability of replacement earbuds leading to partial refunds instead of product replacements (issue_13182). This issue is based on: issue_13089, issue_13182.	Addressing this issue is critical to maintaining customer satisfaction and brand loyalty. Implementing a comprehensive aftersales support system with a wide range of compatible parts can be achieved through partnerships with suppliers and improving inventory management.
Battery	Several reports indicated that the charging cases are draining quickly or failing to hold a charge, significantly impacting the usability of the earbuds. For example, cases drained significantly with brief usage, stopped functioning after being placed on a wireless charger, or failed to retain a charge leading to poor earbud performance. This issue is based on: issue_11773, issue_12364, issue_12595, issue_12886.	Resolving battery issues is essential for ensuring product reliability and customer trust. This can be feasibly implemented by enhancing quality control procedures and using higher quality battery components.
Connectivity	Several users face challenges with the Bluetooth connection's initial setup and maintenance, such as pairing issues, need for frequent re-pairing and specific position requirements for functional connectivity. The setup process is often confusing due to poor manual translations and inadequate instructions. This issue is based on: issue_1387, issue_2180, issue_2305, issue_2404, issue_2588, issue_2594, issue_3085, issue_3201	Reliable connectivity is vital for seamless user experience. Simplifying the connection process through software updates and providing clear, well-translated instructions can address this issue effectively.

TABLE C2 | Sample from an AI-generated solution.

Selected Problem ID	Product aspects	Sample of AI-generated solutions
#3	Aftersales support services	Concrete, Specific and Feasible Ideas: 1. Revamp the Online Return Process: – Develop a streamlined, intuitive online return portal that reduces the number of steps to initiate a return. This portal should guide users through a step-by-step process, clearly laying out each stage of the return, from initiation to shipping instructions. – Include a video tutorial on the website explaining how to navigate the return process. 2. Implement a Transparent Return Policy: – Clearly display the return policy on all sales platforms, packaging, and receipts, specifying the return period, conditions for return, and the state in which items need to be returned. – Standardise the return policy across all platforms to avoid discrepancies and ensure the policy is easily understandable without legal jargon. 3. Enhanced Customer Support: – Establish a dedicated helpline and email support solely for returns and exchanges that is staffed by well-trained representatives. This should help expedite resolutions and reduce confusion. – Implement a Chabot specifically designed to handle returns and warranty questions, equipped with the capability to route complex queries to human operators. 4. Flexible Returns Policy: – Remove the requirement for returns to be in original packaging but ensure that the product is still in a returnable condition. – Allow for a more extended return period of 45–60 days, particularly for online purchases, to accommodate delays in shipping and decision-making. 5. Proactive Communication: – Send automated reminders via email or SMS about the pending end of the return period. Include a direct link or clear instructions on how to initiate a return. – Follow up with customers post-purchase to ensure satisfaction and to inform them about support avenues available if they are not satisfied. Groundbreaking, Visionary and Potentially Revolutionary Ideas: 1. Augmented Reality (AR) Return Assistance: – Develop an AR app that customers can use to scan their product, visually illustrating how to repack the item if necessary, and initiating a return. This app could guide them through each step of the process in real time. – Use AR to assess the condition of the returned product, speeding up the inspection process and reducing the time to process refunds or exchanges. 2. Blockchain for Transparency: – Utilise blockchain technology to maintain transparent and immutable records of purchases, returns, and exchanges. This could help track the lifecycle of a product and its return history, aiding in warranty verifications and reducing fraud. 3. AI-Powered Predictive Troubleshooting: – Before allowing a return, engage customers with an AI-driven troubleshooting system that can predict and solve potential product issues based on previous user data. This could reduce unnecessary returns and enhance product satisfaction. 4. Subscription Model for Upgrades: – Introduce a subscription service where customers can get the latest models of headsets delivered to them as soon as they are released. This model would include a hassle-free trade-in option for the old models that bypasses the traditional return process. 5. Dynamic Return Window: – Implement a dynamic return window that varies with the product's usage patterns detected through smart technology. For less used products, extend the return period, offering more flexibility and potentially enhancing customer trust and satisfaction. Implementing these ideas would not only address the existing issues but also revolutionise the return process, making it a competitive advantage in the marketplace.

TABLE C3 | Samples from further refinement of groundbreaking AI-generated solutions.

Product aspect	Breakthrough ideas/solutions description	Product expert further refinement of AI-generated solutions
Battery	1. Self-Charging Eco-Friendly Case: Develop a charging case that harnesses solar power or kinetic energy (from movement) to continuously or incrementally recharge itself, reducing dependency on electrical charging and aligning with sustainable energy trends. 2. Modular Battery System: Design a revolutionary modular charging case where batteries can be individually replaced or upgraded, allowing for longer overall life and customization based on user needs (such as longer-lasting batteries for travellers). 3. AI-Driven Power Allocation: Integrate AI technologies to analyse user's listening habits and adjust power consumption dynamically. For example, if the system recognises a pattern of short bursts of usage throughout the day, it could manage energy distribution in a way that extends the battery life. 4. Multi-Device Wireless Charging Hub: Develop a futuristic multi-device charging platform where the charging case itself acts as a power bank for other devices (phones, tablets, other headsets) and can simultaneously charge multiple devices, including the earbuds. 5. Smart Temperature Regulation: Innovate a new thermal management system within the charging case that automatically adjusts internal temperatures based on external conditions to prevent overheating and optimise battery efficiency and safety.	1. Self-Charging Eco-Friendly Case: A self-charging case harnessing solar or kinetic energy is exceptionally promising, particularly for travellers. This concept aligns with sustainable trends, but practical implementation requires extensive R&D, pilot testing, and addressing potential inefficiencies in energy conversion. 2. Modular Battery System: Modular systems enhance user convenience and product longevity. Implementing this would involve developing interchangeable components and ensuring seamless integration with existing products. Marketing should emphasise customization and sustainability. 3. AI-Driven Power Allocation: AI can significantly extend battery life by learning from user patterns. Refinement involves robust AI algorithms, user-friendly interfaces, and real-time feedback mechanisms to optimise power use dynamically. 4. Multi-Device Wireless Charging Hub: Creating a versatile charging hub enhances user convenience. Focus on compatibility across various devices, safety standards, and efficient energy distribution. 5. Smart Temperature Regulation: Advanced thermal management can prevent overheating and improve battery efficiency. Implementing sensors and automated controls to adjust temperature based on usage and environment is crucial.