VALIDATION AND DEVELOPMENT OF THE BUSINESS IDEA OF AN AI-POWERED TRANSCRIPTION PRODUCT

by

Mark Motliuk
Roman Mychka

A thesis submitted in partial fulfillment of the
requirements for the degree of
BA in Business Economics, Social Sciences Department

Kyiv School of Economics

2025

Thesis Supervisor: Olena Galytska, Lecturer, Faculty of Social Sciences

Thesis Consultant: Mykhaylo Vidyakin, Lecturer, Faculty of Social Sciences

ABSTRACT

This thesis documents an entrepreneurial journey leveraging OpenAI's Whisper technology across three distinct ventures. It documents the iterative development, validation, and pivoting process, originating from a real-time transcription concept for individuals with hearing loss, transitioning to offline meeting transcription, and culminating in an AI-powered study aid generator for university students. This paper explores how forthcoming advanced AI capabilities can be translated into viable business models within competitive markets.

A systematic approach combined rigorous market analysis via Porter's Five Forces with lean startup principles. Each concept underwent validation through direct user engagement, prototype testing, technical feasibility assessments, and feedback collection. Implementation drew upon Whisper, specialized diarization models like Pyannote, and large language models, adapting the technology stack to each venture's specific requirements. Financial modeling and legal structure analysis informed viability assessments.

Early ventures faced critical roadblocks. The initial accessibility application did not leave up to the expectations due to limited market demand and monetization challenges within the target demographic, despite technical promise. A subsequent pivot to corporate meeting transcription proved unsustainable against entrenched competitors offering superior value propositions. The final shift towards an educational tool — generating notes, quizzes, and flashcards from lectures – achieved significant preliminary validation. Engagement metrics and initial purchase rates confirmed problem-solution fit, demonstrating clear user interest and willingness to pay, though indicating the necessity for scaling beyond a single institution for financial viability.

This journey shows that mere technical sophistication is not enough for the product to be viable. An innovation, to be successful, demands rigorous testing against real-world markets — competitive landscape, customer willingness-to-pay, and economic feasibility. While the final education-focused venture demonstrates potential, its long-term success depends on scaling and expansion. Ultimately, entrepreneurial resilience lies in responsive adaptation, not rigid commitment to initial concepts.

TABLE OF CONTENTS

Introduction

In September 2021, OpenAI introduced Whisper, a series of open-source AI models designed for audio transcription ("Introducing Whisper"). These models demonstrated superior reliability and portability compared to existing alternatives, achieving 86.28% accuracy on Ukrainian audio versus 72.63% for Facebook's open model wav2vec2, fine-tuned for Ukrainian (Smoliakov). Whisper's open-source nature, combined with its ability to run on devices as modest as mobile phones, created numerous opportunities for innovative applications and served as an inspiration and a starting point for this thesis (Kitaiti).

This paper documents the journey through three distinct business ventures, each leveraging Whisper's capabilities to address different market needs. The paper begins with the idea of real-time transcription services for individuals with hearing loss, a product concept aiming to improve accessibility through accurate real-time speech-to-text conversion and speaker identification. Following market validation challenges, a pivot to explore offline meeting transcription services is warranted, targeting organizations seeking efficient documentation of discussions. Struggling with customer acquisition, we pivoted towards the education sector. Our new approach involved developing a system that could automatically generate notes, quizzes, and flashcards from lecture recordings to help students learn more effectively.

For each venture, a systematic approach to market analysis is employed using Porter's Five Forces framework, examining competitor rivalry, the threat of substitution, barriers to entry, and the bargaining power of both suppliers and buyers. Throughout the process, each concept is validated through direct user engagement, prototype testing, and feedback collection.

The documentation of this entrepreneurial journey illustrates the iterative nature of business model development, highlighting how initial concepts evolve through market research, customer feedback, and practical implementation. By detailing successes and setbacks across multiple pivots, this paper provides insights into the practical application of lean startup methodology in technology-driven ventures.

[The Original Idea]

Real-Time Transcription for Individuals with Hearing Loss

THE IDEA

Inspired by advancements in AI technology, the initial product vision was to develop an application for people with partial or complete hearing loss. In this paper, they are referred to as individuals with hearing loss. The solution provided real-time transcription so the users could understand what others were saying by reading a transcript of the conversations on their smartphones. The key feature was diarization — labeling the speech of different speakers. A combination of transcription and diarization allowed individuals with hearing loss to follow discussions with multiple individuals speaking simultaneously. Finally, the users could type their answers to respond, which would then be instantly converted into speech.

DESK MARKET RESEARCH

Porter's Five Forces framework is used to conduct market research. This framework was selected since it helps to analyze the external competitive environment rather than the company's internal characteristics. Therefore, it is suitable for cases where establishing a business to enter the market is considered. Also, this approach may help to identify gaps in what is currently offered by rivals. These gaps can be used to launch the new product and gain a competitive advantage over the existing competitors. Therefore, Porter's Five Forces framework is used to identify possible winning strategies in the market of assistive transcription applications for individuals with hearing loss.

Competitive Rivalry

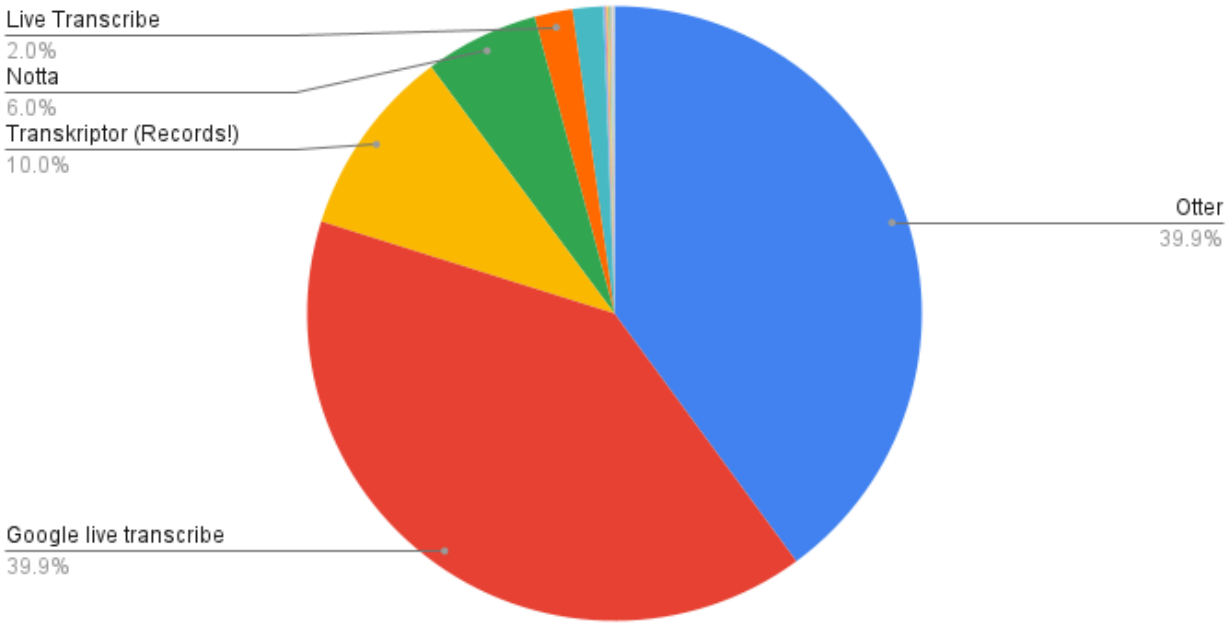
In order to access quantitative data, AppMagic and Sensor Tower are used ("AppMagic Main Page"; "About | Sensor Tower"). These online tools provide additional information about mobile applications published on the App Store and Google Play ("App Store"; "Google Play Main Page"). The data regarding the number of installations, their distribution by country, user ratings, and estimated revenue are available on these websites for free, while more detailed information requires a premium account. Consequently, these sources are utilized for research on competitors in the market.

This research evaluates the shares of players in the mobile applications market by the number of installations. It is worth noting that the figures from the AppMagic and Sensor Tower are not entirely accurate. These tools use data from open sources and their custom estimation models, leading to potential discrepancies ("Key Facts About AppMagic"; "Responsibly Sourced Data"). Also, the number of downloads displayed does not exclude multiple downloads of the same application by the same user owning multiple devices. Nevertheless, these sources highlight the relative positioning of the rivals in the market and provide enough accuracy for a high-level market understanding. In order to mitigate potential biases, only the number of downloads was taken into consideration to assess the market size and shares of competitors since it can be verified for some solutions on their pages in the digital stores.

The number of installations is more than 10 million downloads for transcription applications designed for individuals with hearing loss. However, including solutions that provide transcription for general purposes results in almost 25 million installations. As the target audience might choose solutions not necessarily designed for individuals with hearing loss, all the applications suitable for the problem are considered. Therefore, the market size for the product is significant — as indicated by approximately 25 million downloads (Appendix A).

The shared market of two segments is dominated by *Google Live Transcribe* and *Otter.ai*, with approximately 10 million downloads of each application. The following major products are general-purpose transcription tools *Transkriptor* and *Notta*, with 2.5 million and 1.5 million installations, respectively. The last solutions with downloads greater than 100 thousand are *Live Transcribe* and *Ava* — the applications for individuals with hearing loss as their direct target audience. The rest of the rivals are much smaller. Consequently, the market is dominated by a few players holding significant market power: the share of the biggest six competitors is more than 99%, while the Herfindahl-Hirschman Index is almost 3332 (Appendix A; Figure 1; Bromberg).

Figure 1: Market Distribution of Mobile Applications by Percentage of Downloads (Real-Time Transcription for Individuals with Hearing Loss).



Data from Appendix A.

In order to identify the recommended characteristics of the Minimum Viable Product, the overview of the functionality of existing competitors in the segment of transcription applications designed for individuals with hearing loss should be accomplished. The primary need of people installing these solutions is the ability to read what is being spoken. The accuracy of transcription is the central focus for the market players: some provide users with basic transcription for free but charge for access to higher accuracy. Therefore, the quality of the core function should be good enough to satisfy the customers.

The applications of the target segment also have additional features. Individuals with hearing loss use their eyes more as an alternative way of receiving information, motivating developers to enable the option to make the font size of the transcription bigger. Moreover, some solutions like *Google Live Transcribe*, *Live Transcribe*, *Wushi: Live Transcribe Voice*, *Deaf Transcribe:Speech to Text* help not only to be informed about what is said in the conversation but also to communicate by typing the response in the application and showing it to the speaker. Moreover, *Ava : Transcriptions and Captions* allows users to transform the answer into sound. Moreover, the product identifies and labels words spoken by multiple people if the app is installed on their smartphones. Some of the

aforementioned competitors have made transcription available without internet coverage. Although the rivals offer various functionalities, only one implemented diarization (Appendix B).

Analyzing revenue streams on the market may provide important insights regarding the optimal pricing strategy. The prevailing type of monetization of transcription applications for individuals with hearing loss is a monthly or yearly subscription. Nevertheless, free options like *Google Live Transcribe* and *Captify: Live Caption Deaf+HoH* exist. The prices for monthly payments are in the range between $ 14.99 and $ 0.99. They depend on the variety and quality of the features offered. Also, some products give potential customers full access for a limited time to encourage the acquisition of the solution. Overall, the business model of the majority of the rivals involves charging up to 15 dollars per month (Appendix B).

Threat of Substitutes

The main substitutes are hearing aids and cochlear implants. The most common solution is a hearing aid. This medical device processes the sound of the environment and selectively amplifies it based on the type of hearing loss so that the individual can hear it better. The prices in Ukraine start from 4500 hryvnias and depend on the characteristics of the device: design, integration with the smartphone, ability to customize, and other features ("Slykhovi Aparaty"). Cochlear implants are more suitable for individuals with severe hearing loss since they stimulate the auditory nerve directly ("Hearing Aids vs Cochlear Implants"). A significant disadvantage is that they require surgery to be installed. Moreover, the prices for them without additional expenses start from 20 000 dollars ("Pro Kokhlearnu Implantatsiiu"). However, some organizations report providing these devices for free ("Ukrainian Association"). Also, although sign language is not widely spread across the whole population, it is often used by individuals with hearing loss to communicate with each other. Therefore, there is a high threat of substitutes for buyers in this market.

Threat of Entry

In order to assess the threat of new entrants in the market, the number of solutions released in 2024 is analyzed. For this year, only three applications were published in digital stores. Each segment has exactly one newcomer for this period. The limited number of new entrants suggests potential

barriers to entry, such as high development costs or intense competition from established players. This situation indicates that the market may not be attractive for new businesses (Appendix C).

Supplier Bargaining Power

Supplier power for this initial venture hinges critically on the providers of specialized AI technologies, cloud hosting, and payment processing. The core function — low-latency speech-to-text conversion with speaker diarization — dictates a heavy reliance on sophisticated AI models. While OpenAI's Whisper provides a strong foundation, deploying and scaling it from scratch demands prohibitive upfront investment (Walton). Consequently, utilizing hosted Whisper APIs via services like Azure or Replicate becomes necessary. The stringent low-latency requirement for real-time use significantly narrows the field of viable providers, amplifying their bargaining power.

The challenge intensifies with speaker diarization. Pyannote stands out as a leading model, yet its niche status means no providers offer readily hosted inference (Bredin). Self-hosting through serverless AI platforms like RunPod, which allows pay-per-second usage and custom model versions (crucial for potential Ukrainian language fine-tuning), emerges as the sole practical path ("RunPod Main Page"). The scarcity of affordable serverless AI inference providers like RunPod gives these suppliers substantial leverage.

In contrast, the market for generic compute infrastructure is highly fragmented. Numerous providers, from giants like Amazon Web Services, Azure, and Google Cloud to cost-effective players like Hetzner (the chosen provider for its affordability), compete fiercely on price and features. This competitive landscape grants suppliers of basic compute minimal bargaining power. Similarly, the payment processing sector is crowded with established entities and fintech startups offering comparable services. Switching between providers like Stripe or LiqPay is feasible, keeping their individual influence low.

Considering these factors, overall supplier power is assessed as significant. While compute and payment providers wield little influence, the venture's core functionality creates unavoidable dependencies on a limited pool of specialized AI service providers like Azure and RunPod, conferring notable power upon them due to the lack of ready alternatives for the required low-latency transcription and niche diarization capabilities.

Buyer Bargaining Power

The analysis of the power of buyers in the market is focused on Ukraine. The reason is that the Minimum Viable Product is initially planned to be launched in this country. In order to assess the number of potential customers in this market, the data from the Department of Economic and Social Affairs of the United Nations and the World Health Organization reports are used. According to the report, the estimated population of Ukraine at the beginning of 2024 is 37.441 million (United Nations (2024)). The percentage of individuals with hearing loss worldwide is approximately 5 percent ("Deafness and Hearing Loss"). The resulting number is close to 1.872 million people who might use the product. The non-governmental organization Public Movement "Social Unity" provides similar figures: 41 million of the population in the country and 2 million individuals with hearing loss ("Digital Solutions"). Consequently, the number of potential buyers in the market is significant.

The large number of people in the market does not necessarily guarantee high revenues. The switching costs between apps are close to zero. Users can simply install a new app, pay the cost of a subscription, and start using it. Also, a sample of potential users from Ukraine was interviewed. They reported spending no more than 100 hryvnias per month. Also, the majority of respondents mentioned trying several products before selecting the one with the best accuracy. The feedback reinforces the conclusion that switching costs are low, resulting in a very high power of buyers.

ACTIONS TAKEN

The analysis of the market helped us identify key gaps. Firstly, *Ava : Transcriptions and Captions* implemented diarization. However, the feature could be used only when all the conversation participants installed the application on their smartphones. Resolving this limitation could give us a competitive advantage over the existing players in the market. Secondly, only the aforementioned rival implemented a text-to-audio. It allows users to type in text, which is then converted into audio, allowing individuals with hearing loss to communicate more easily. Therefore, the next step was to validate the idea by testing the technical feasibility and determining whether the problem existed in the Ukrainian market.

To validate technical feasibility, a prototype was rapidly developed using Python. This demonstration application processed real-time microphone input directly on the device. It utilized locally executed instances of the small Whisper model for transcription and Pyannote for speaker diarization.

As those models do not natively work with streaming audio, a significant technical hurdle involved devising an algorithm to segment the continuous audio stream and reconstruct a coherent transcript from the chunks. The functional prototype incorporating this solution was showcased on the website zvuk.ai. This initial demonstration attracted interest from the administrative team at the Kyiv School of Economics, and the project was showcased at a charity event and a psychology conference. The project also secured recognition by winning the Falling Walls Lab Kyiv competition, culminating in a presentation at the final stage in Germany. Preliminary experiments also explored the application's potential network functionality. These tests assessed latency constraints associated with utilizing larger, more powerful Whisper models hosted remotely. However, development beyond these basic network trials was deliberately paused, prioritizing the need for validation and feedback from actual end-users before committing further technical resources.

The analysis of the market led us to a consideration of alternative channels for the acquisition of users. It was concluded that attracting users with paid advertisements could be expensive and complicated. Therefore, it was decided to partner with institutions that helped individuals with hearing loss. Firstly, they could help determine whether there was a demand for the solution among the target audience. They were good examples of the concentration of potential customers. Also, these organizations could sponsor our product. Moreover, their projects could promote the solution to a broader audience. Overall, it was necessary to establish strong relationships with people who worked with individuals with hearing loss.

Communication with the representatives of the aforementioned institutions helped determine the actual demand for a solution in the Ukrainian market. We started by contacting the Ukrainian Society of the Deaf ("Ukrainian Society of the Deaf Main Page"). They positively reacted to the labeling of speakers in the prototype and provided some suggestions regarding its overall appearance. However, they mentioned that the institution is focused on helping individuals with hearing loss for free. Also, they were more interested in developing a mobile application to connect users with sign-language translators to provide free help for individuals with hearing loss in communication with others. The other potential partner was the public organization Vidchui ("Vidchui Main Page"). They shared that there were several more similar products with similar value propositions. Furthermore, they were not willing to assist with customer interviews.

LESSONS LEARNED

The insights discovered during the market research and the process of validating the idea resulted in a change in the direction of product development. Firstly, the research highlighted that a significant number of potential users did not necessarily result in high revenues. Additionally, there was a high threat of substitution from the market of cochlear implants and hearing aids. Moreover, in the process of validating the idea, the feedback from stakeholders from the Ukrainian Society of the Deaf and Vidchui indicated that there was no significant demand for the development of the described product. Consequently, a decision was made to change the direction of the product.

[Pivot 1]

Offline Meetings Transcription Service

THE IDEA

The second idea was to develop a solution for transcribing offline meetings. The key feature was diarization — labeling the speech of different speakers in a single audio file. Implementing this functionality enabled users to upload recordings of conversations and receive transcripts indicating which parts were spoken by each of the participants. In the process of development, the Minimum Viable Product was enhanced with additional features. Firstly, multilingual recognition was introduced to receive accurate results in cases where multiple languages were used in conversations. Also, timestamps were added to each phrase in transcripts so that the customers could verify whether the transcription was accurate. Overall, the feedback from the first customers helped to refine and improve the product.

IMPLEMENTATION

During one of the Strategic Management course lectures, we discussed this idea with the lecturer, Mykhaylo Vidyakin. Mykhaylo mentioned potential clients within their network and offered to introduce us to them. The existence of such demand instantly validated the idea. Therefore, we could proceed to the next step — the Minimum Viable Product validation.

We established working relationships with the first potential client. The work started with us being introduced to employees of the City Council of Kryvyi Rih. They were looking for a solution to delegate the creation of transcriptions of the meetings in their institutions. They usually needed the transcriptions to recall the discussion and plan the next steps in the projects. Therefore, the transcription accuracy of the product was expected to be very high. Also, they mentioned that they had up to 5 hours of audio to transcribe daily. Moreover, the expected monthly budget of customers was 20 dollars. As a result, the first iteration prototype needed to transcribe audio of meetings with high quality and on a minimal budget.

The initial development of the product started. The decision was taken to work closely with the aforementioned organization in order to create a Minimum Viable Product and then sell it to

companies and public institutions with similar tasks. It offered the ability to create a transcription by uploading an audio file. At first, it supported only the Ukrainian language. The price of transcription was 15 hryvnias per hour of audio. The resulting transcription was stored for each audio file so that users could access it again later. Consequently, the organization received access to the prototype.

After a week of active use, the City Council of Kryvyi Rih employees were interviewed on their experience with the prototype. They shared an important insight: people often use Russian in discussions. Therefore, the prototype that supported Ukrainian did not process some phrases correctly. Also, the prototype did not support the format of the audio files they used. Thus, they needed to convert the files before uploading them to get transcription. Moreover, we could track their activity on the website and observe that they were, in fact, using it. Overall, we confirmed the need for the product and discovered some weaknesses in the current version.

The next version of the product addressed the insights learned from the interview. Support for the needed audio formats was introduced. Also, Russian and English languages became available. In order to solve the problem of transcription of different languages in one conversation, a new transcription mode was added. Before transcription, the language of each part of the recording was defined. The price of this new transcription mode was 36 hryvnias per hour of audio. Furthermore, a new feature was implemented highlighting parts of a transcription where the model was unsure, allowing the user to double-check potentially wrong phrases. They could view timecodes for each sentence in the transcription so that they could check it in the original audio. The new version was announced to the customers shortly.

The feedback from the City Council of Kryvyi Rih employees made us research the competitors. The features introduced covered the problems, and the clients continued to use the product. They even paid 100 hryvnias to continue using it after the end of the trial period. However, after two more weeks, they stopped using the tool. Therefore, the interview with the users was planned. The reason behind this was the discovery of a similar product that offered more features and better accuracy for free. Consequently, we decided to analyze the market more deeply before planning the next steps.

DESK MARKET RESEARCH

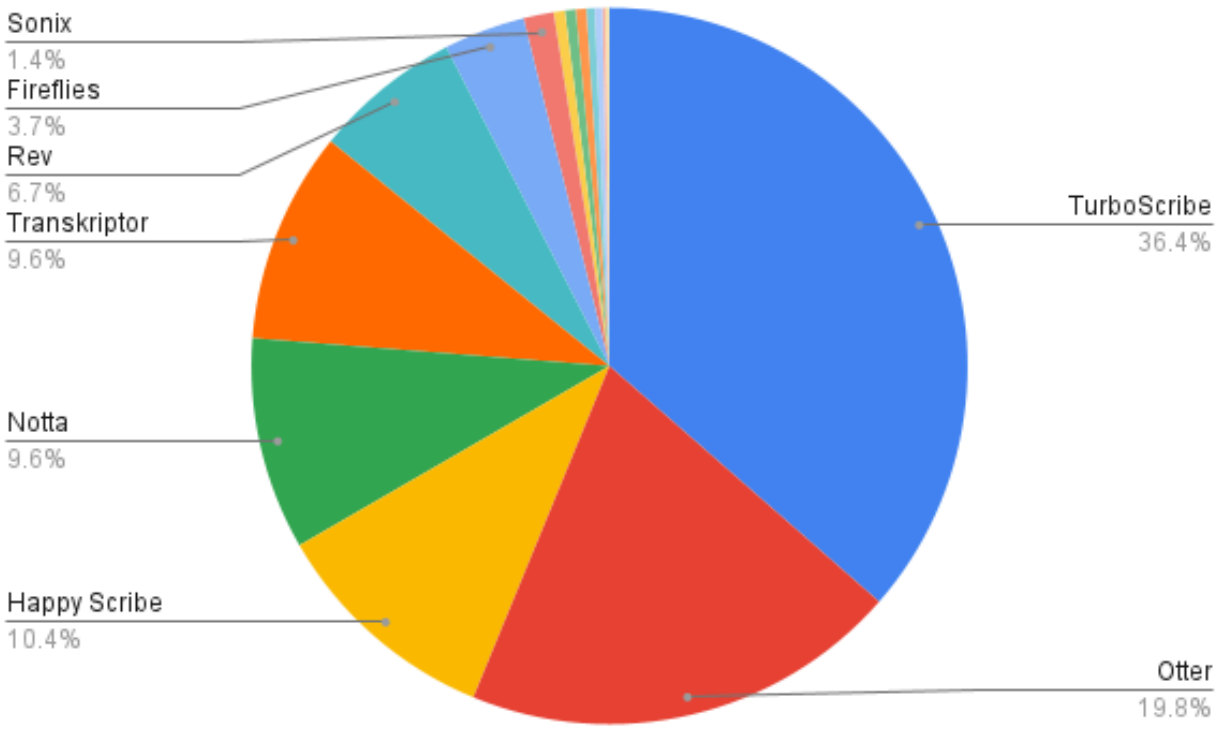
Competitive Rivalry

In order to estimate the market size of the online transcription applications, data from Similarweb is utilized. Similar to Appmagic and Sensor Tower, the product uses custom predictive models to estimate the number of visits to multiple websites ("Similarweb's Data Accuracy"). Additionally, it provides information about the devices, countries, and time spent by the website's visitors. These figures may be inaccurate, but they should show the general trend. The selected method of market estimation involves aggregating the number of average monthly unique visitors of each market player over a selected period of time. Despite potential inaccuracies, this metric is chosen since it reflects the average number of active users in the market.

The data utilized covers the period between November 2024 and February 2025. However, *Jamie* and *Fathom* are omitted since they need to be installed on a personal computer. Consequently, the information about the traffic on their websites is not representative. According to this approach, the estimated market size is approximately 4 million average monthly unique visitors for the segment of AI-powered meeting assistants and 7.738 million for general-purpose transcription software. Therefore, the total size of the market is close to 11.754 million (Appendix D; Appendix E).

The transcription software market is concentrated around a few major competitors. *TurboScribe* attracts more than 4 million unique visitors on average monthly. This corresponds to the largest share of the market — approximately 36%. The next major player is *Otter*, with a share of almost 20%. *Happy Scribe*, *Notta*, and *Transkriptor* each hold an estimated 10% market share. They are followed by *Rev*, attracting approximately 784 thousand unique visitors on average monthly. The remaining portion of the market is distributed among ten smaller rivals. The Herfindahl-Hirschman Index, which measures the concentration of markets, is about 2071. Although these numbers suggest a moderate market concentration, the competition remains intense (Appendix E; Figure 2).

Figure 2: Market Distribution of Web Applications by Percentage of Average Monthly Unique Visitors (Offline Meetings Transcription).

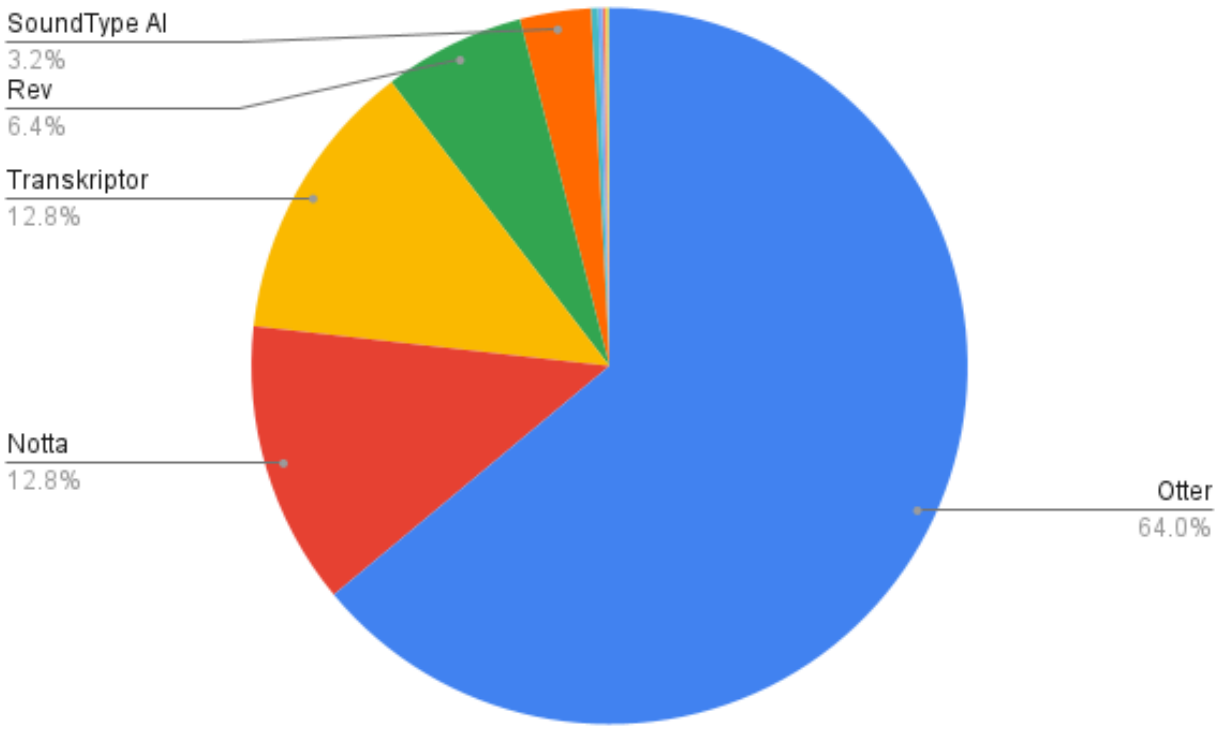


Data from Appendix E.

More than half of the competitors researched have their applications available on the App Store and Google Play. The market size is estimated at approximately 15.63 million installations by aggregating the number of downloads of the products from AppMagic and Sensor Tower. The distribution of installations is different from the distribution of the number of average monthly unique visitors of web applications. Although general-purpose automated transcription software solutions dominate the web segment, they account for approximately 3.53 million downloads. Consequently, the mobile market, estimated at 15.63 million downloads, is dominated by AI-powered meeting assistants (Appendix F).

The mobile segment of the automated transcription software market has a few major players. *Otter* has the largest market share of 64%, corresponding to over 10 million downloads. *Notta* and *Transkriptor* each account for approximately 13% of the market. *Rev* and *SoundType AI* follow them with about 1 million and 500 thousand downloads, respectively. The share of the remaining five products is less than 1%. The value of the Herfindahl-Hirschman Index is approximately 4472. Consequently, the mobile applications segment in the automated transcription software market is highly concentrated among key competitors (Appendix F; Figure 3).

Figure 3: Market Distribution of Mobile Applications by Percentage of Downloads (Offline Meetings Transcription).



Data from Appendix F.

In order to assess the functionality required for the launch of the new product, the features offered by the existing competitors should be analyzed. They all allow users to generate transcripts from uploaded audio or video files. However, some solutions also enable users to provide hyperlinks or directly connect to online meetings via such platforms as Zoom and Google Meet. All the competitors except *TurboScibe* and *SpeechText.AI* can be used to create summaries based on generated transcripts. Moreover, 11 out of 18 products have integrated AI chatbots that can answer questions related to the contents of the provided files. Market players incorporate advanced technologies to process input files. For example, 13 of them support diarization — labeling the speech of different speakers. This exact number of solutions recognize the Ukrainian language. However, only 10 can process multiple languages within a single audio file. In conclusion, the new entrants should support various input methods, advanced capabilities for processing, and tools that maximize the value received by users (Appendix G).

In order to determine a competitive pricing strategy, an analysis of the market players is conducted. Firstly, all the products utilize a subscription-based monetization model. Each of them

offers several monthly plans for different prices. Thus, their customers can choose the options that are the most suitable for them in terms of usage limits. The price for the cheapest subscription plans ranges from $ 9.99 to $ 29 per month. *Trint* is an outlier that charges at least $ 80 monthly. Also, the products grant users trial periods to encourage them to acquire complete access later. Moreover, all the competitors except *TurboScribe* offer special subscription plans for institutions (Appendix G).

Threat of Substitutes

The solutions in the market analyzed can be substituted. Firstly, Microsoft, Google, and Apple offer the users of their products integrated solutions that transform speech into text. Although they do not support labeling the speech of different speakers and recognition of multiple languages, they can be used to dictate instead of typing. Also, there exists a market for manual transcription services. Despite the high prices and the amount of time needed to provide services, some customers prefer human transcription since it is considered to be more accurate. Consequently, the threat of substitution for automated transcription software products in the market is moderate.

Threat of Entry

In order to assess the threat of entry of new solutions, the number of existing companies in the market that were founded in 2024 is analyzed. However, no competitors have entered the market this year. This fact may indicate that significant investments are needed to develop the product. Also, the acquisition of users may be expensive. Therefore, the threat of new entrants in the automated transcription software market is low (Appendix H).

Supplier Bargaining Power

This pivot to offline meeting transcription retains a core reliance on AI for transcription and diarization, echoing the dependencies outlined in the initial venture. However, the shift away from real-time processing fundamentally alters the supplier dynamic for the AI components.

Removing the strict low-latency constraint broadens the pool of suitable AI inference providers significantly. Beyond the previously considered options like Azure and Replicate, major cloud platforms like Amazon Web Services and Google Cloud offer viable transcription models alongside OpenAI’s direct offerings and numerous more minor players. While Whisper might remain the preferred choice for quality, the increased availability of functional alternatives for offline tasks

diminishes the leverage held by any single AI provider compared to the real-time scenario. The niche requirement for diarization persists, but the flexibility of offline processing may allow for different implementation strategies since absolute real-time accuracy is no longer necessary.

As established previously, the markets for essential support services remain favorable. Generic cloud hosting offers abundant choice, with affordable options like Hetzner readily available, ensuring hosting providers possess limited power. Likewise, the competitive payment processing landscape keeps supplier influence minimal in that domain.

Therefore, while AI remains central, the expanded options for offline transcription dilute the concentrated power observed in the first venture. The overall supplier power for this offline transcription service is assessed as moderate.

Buyer Bargaining Power

Determining the power of the customers in the market requires the analysis of multiple factors. The presence of numerous competitors that offer free trials results in intense competition and low subscription prices. The combination of price competition and a wide choice of alternatives indicates the high power of buyers. Also, switching costs between solutions are relatively low. The primary output of the products is transcripts. Therefore, they can be easily exported if the product is not good enough to satisfy users' needs. Consequently, the power of buyers in the market of automated transcription software is high.

LESSONS LEARNED

The analysis of the market and the feedback from our first clients made us reconsider the next steps. The interview with the employees of Kryvyi Rih City Council raised concerns regarding the competitive features of the product and the price of the transcription. The market research findings indicated that these concerns were valid since rivals offered the same functionality for lower prices. Also, no significant gaps in the market of transcription software were identified that could be addressed. Therefore, a decision was made to change the direction of the existing product.

[Pivot 2]

Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings

THE IDEA

The idea of the final product lies in assisting university students in their studies. Specifically, the product is about the creation of quizzes, flashcards, and other learning materials from the recordings of the lectures. The solution utilizes the previous product's technology to create accurate transcripts of the lecture recordings. Then, Large Language Models (LLMs) are used to create supplementary learning materials for users, such as notes, quizzes, and flashcards. These materials are supposed to help students with their studies by offering features that decrease the time needed to write notes and increase the effectiveness of memorization and rehearsal of the contents of the lessons. The product is first tested and refined with the students of Kyiv School of Economics, with plans to scale it to a broader audience.

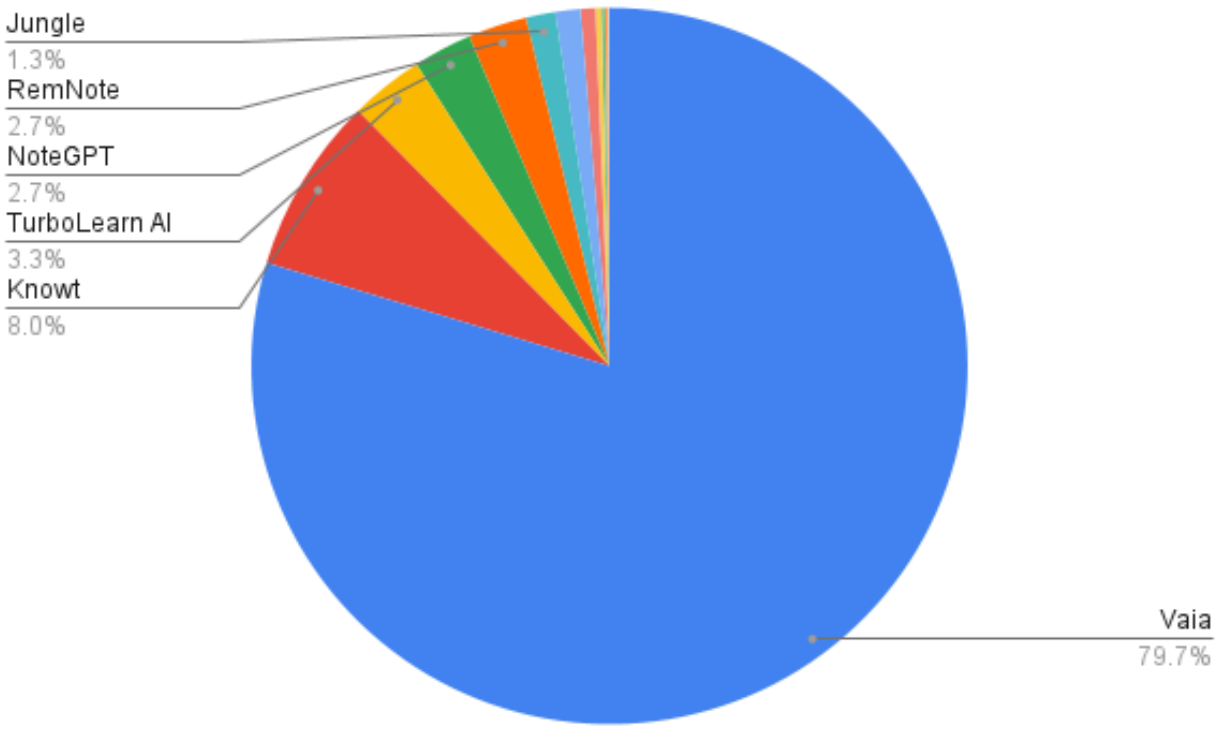
DESK MARKET RESEARCH

Competitive Rivalry

Two approaches are employed in order to estimate the size of the market of such online applications. The first method involves aggregating the number of users of each product. The majority of the solutions analyzed provide this figure on their websites. This approach results in an estimated market size of approximately 37.663 million users (Appendix I; Appendix J).

According to this approach, the market is dominated by *Vaia*. The product reports having more than 30 million users. This is almost 80% of the estimated size of the market. The next major rivals are *Knowt*, *Turbolearn AI*, and *NoteGPT*. *Knowt* shares the number of 3 million users, while *Turbolearn AI* and *NoteGPT* have 1 million each. The last solutions whose market share is above 1% are Jungle and Raena AI, with 500 thousand and 425 thousand users, respectively. Overall, the market is dominated by these six products: their total share is almost 99%, while the Herfindahl-Hirschman Index is approximately 6437 (Appendix J; Figure 4).

Figure 4: Market Distribution of Web Applications by Percentage of Self-Reported Users (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings).



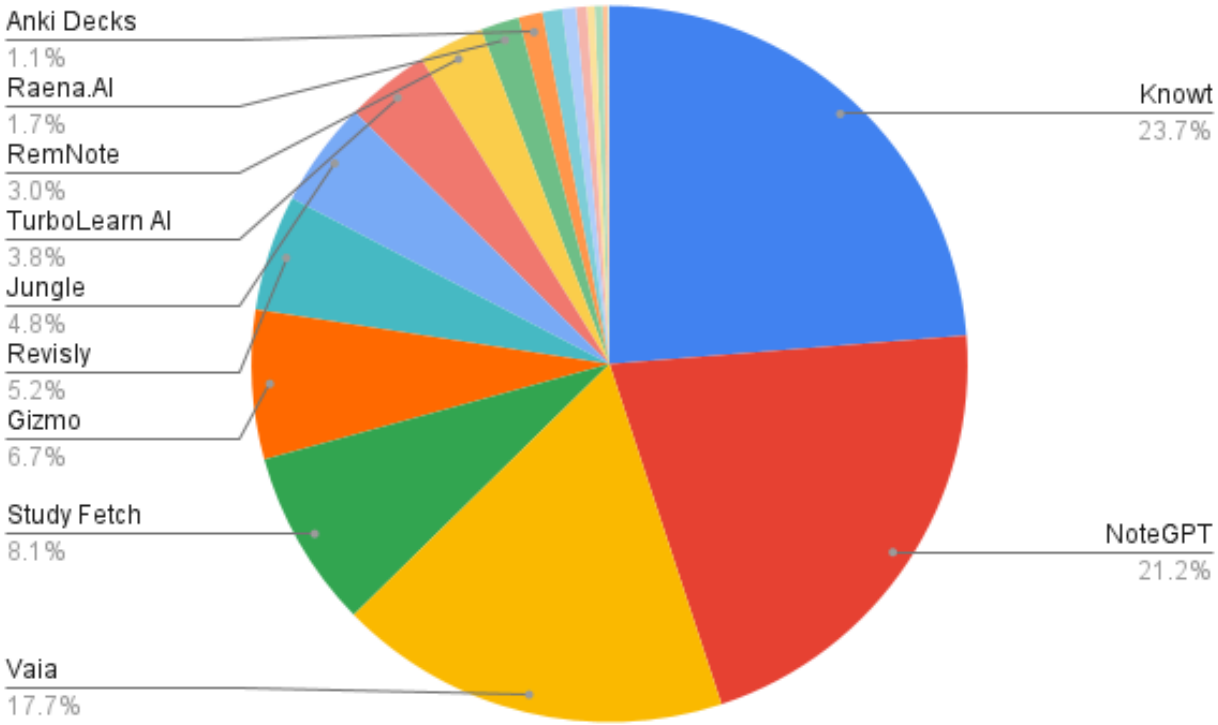
Data from Appendix J.

However, the aforementioned figures may not accurately reflect the current situation in the market since people tend to stop using the products within some period of time. Therefore, an alternative approach is employed. It involves aggregation of the number of average monthly unique visitors to the competitors' websites. In order to estimate it, the data from Similarweb for the period between November 2024 and February 2025 was utilized. The alternative approach results in more than 10 million average monthly active users in the market (Appendix K).

The alternative approach indicates different competitive distribution within the market. The product with the largest share is *Knowt*. Its number of average monthly unique visitors is close to 2.5 million. *NoteGPT* follows closely with about 2.2 million active users per month, representing 21% of the total market. Although *Vaia* dominates the market in terms of reported number of users, its share of average monthly unique visitors is approximately 17.7%. Beyond these three largest products, several other solutions hold notable market shares. *Study Fetch* and *Gizmo* have approximately 826 thousand and 691 thousand average monthly uses, respectively. The value of the metric for *Revisly* is smaller — about 535 thousand monthly unique visitors on average. The shares of the remaining competitors

are below 5%. The value of the Herfindahl-Hirschman Index is approximately 1500. In conclusion, this approach exhibits less concentration among the major market players, contrary to the findings of the first method (Appendix K; Figure 5).

Figure 5: Market Distribution of Web Applications by Percentage of Average Monthly Unique Visitors (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings).

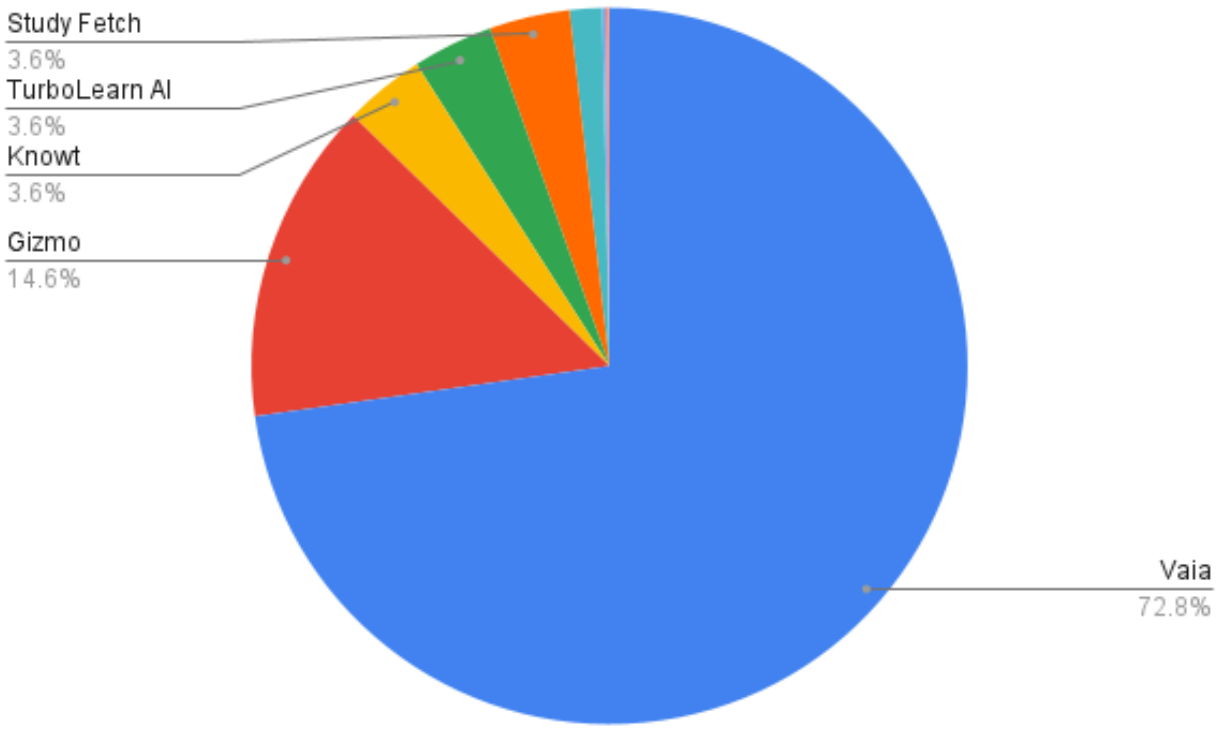


Data from Appendix K

In order to assess the market of mobile applications in the selected niche, data from AppMagic and Sensor Tower is used. The market size is estimated by aggregating the number of downloads of each product. These calculations result in the figure of approximately 13.775 million installations (Appendix L).

The distribution of the number of downloads for mobile applications is similar to the distribution of the number of users reported by the market players. The largest product is *Vaia*. It has approximately 10 million installations and 73% of the market. The share of *Gizmo* is significantly less — about 14.6%. The next competitors are *Knowt*, *Study Fetch*, and *TurboLearn AI*. Each holds more than 500 thousand downloads, which accounts for 3.6% of the market. The Herfindahl-Hirschman Index of approximately 5523 suggests a high concentration in the market. Overall, the market of mobile applications in the selected niche is dominated by *Vaia* and *Gizmo* (Appendix L; Figure 6).

Figure 6: Market Distribution of Mobile Applications by Percentage of Downloads (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings).



Data from Appendix L

The comparison of features of the market players begins with the types of files and languages supported. Fifteen products of the twenty analyzed can process materials in Ukrainian to generate notes, quizzes, flashcards, or other learning materials. Regarding input formats, all the competitors support various formats of text files. However, four do not allow users to upload videos or provide YouTube hyperlinks. Nevertheless, half of the solutions observed enable their customers to upload pictures of their notes for content creation. Consequently, the support of different languages and various formats of input files is a common feature in this market (Appendix M).

The most important aspect of the selected niche is the resulting learning materials. The players in this market offer a wide variety of them. The most common options are the generation of flashcards and quizzes, which help the users to learn efficiently. The generation of structured notes for the uploaded materials is spread less: only eleven of twenty solutions provide this functionality. Also, some products support more unusual types of content. For example, *Study Fetch*, *TurboLearn AI*, and *Raena.AI* generate podcasts in which AI virtual agents discuss uploaded materials. Moreover, *Study Fetch* and *Raena.AI* allow users to generate short video summaries to help people with short attention

spans. Furthermore, *NoteGPT* and *memrizz* can be used to create mind maps to visualize the concepts from the uploaded materials. Overall, although the majority of the competitors offer a generation of quizzes and flashcards, some of them provide additional options in order to differentiate from the rivals (Appendix M).

Competitors in the niche provide additional features in order to facilitate the process of learning and encourage long-term product engagement. Firstly, 17 of the 20 solutions analyzed include an AI chatbot as a personal assistant. It is used to help users with the functionality of the product and answer questions regarding the content of the uploaded materials. Moreover, competitors utilize gamification on the people to motivate students to study more. For example, *Gizmo* awards virtual points for learning new material. It allows users to compete with their friends and customize their profiles. Such mechanics not only encourage people to study regularly but also make their engagement with the product longer. Furthermore, half of the solutions provide an option to export created flashcards for use in other applications. While this feature is convenient for users, it may also increase their probability of switching to alternatives. The functionalities of the market players analyzed are not limited to those mentioned above: the rivals continuously introduce new features to increase learning efficiency, engage their users, and differentiate themselves from the others (Appendix M).

In order to determine the optimal pricing strategy, it is crucial to analyze the monetization models of the existing competitors. The prevailing pricing model among the market players is subscription-based: 19 out of 20 solutions have adopted it. The prices for individual subscription plans vary between $ 5 and $ 20 per month, depending on the number of features offered. Additionally, pricing varies depending on the usage limits for specific functionalities. For example, a $ 8.99 subscription allows users of *Raena AI* to generate up to 30 flashcards per day, whereas a $ 16.99 plan offers unlimited usage. Also, some competitors like *Study Fetch*, *Turbolearn AI*, and *Study Snail* offer specialized plans for institutions. Nevertheless, their prices are not publicly disclosed. Adopting a subscription-based monetization model with tiered pricing based on usage limits is a viable strategy for new products in this niche (Appendix M).

The players in the market employ additional strategies to increase the number of subscriptions. For example, all the products offer their new users temporary access to all the features for free to encourage them to subscribe. Also, half of the competitors offer users the opportunity to promote the

product to their friends by offering discounts. The aforementioned methods can be utilized to compete in the market (Appendix M).

Threat of Substitutes

There is a relatively high threat of substitution for the solutions of the niche analyzed. Firstly, traditional flashcard-based learning tools that do not implement AI-generated content remain popular. For example, according to Similarweb, Quizlet attracts almost 27 million unique monthly users ("Quizlet Main Page"). Moreover, general-purpose AI models like ChatGPT, Claude, and Gemini pose a significant threat of substitution. They may also be used to create flashcards, quizzes, and summaries from the text files. Moreover, users can extract transcripts from YouTube videos and use them as input for AI tools to generate personalized study materials. Therefore, the products in this market must differentiate significantly to avoid competition from substitutes.

Threat of Entry

The next important aspect is to analyze the potential for new entrants in the market. It is relatively easy to enter the observed niche for new products. At least five new solutions were launched in 2024. However, the fact that the share of the newcomers is approximately 3.7% of the total average monthly unique visitors in the market suggests that it is challenging to compete with the incumbents. Therefore, although entering the market is not difficult, acquiring and retaining users may require strong differentiation and an outstanding marketing strategy (Appendix N).

Supplier Bargaining Power

This venture introduces a critical new dependency: Large Language Models (LLMs) for generating educational materials from transcribed lectures. Accurate transcription remains foundational, drawing on the AI processing capabilities similar to the offline transcription service (Venture 2), but the dominant supplier influence now shifts towards LLM providers.

The market for powerful LLMs is currently dominated by a handful of major corporations – OpenAI, Google, Anthropic, Microsoft – and specialized AI firms (Fernandez). This concentration inherently grants these providers significant influence. Building and maintaining the infrastructure for self-hosting state-of-the-art LLMs presents substantial technical and financial barriers. However, a crucial mitigating factor exists: abstraction tools like LangChain dramatically simplify the process of

integrating and switching between different LLM APIs (“Providers”). This drastically reduces vendor lock-in and switching costs for the core generation capability, tempering the power of individual LLM providers.

Robust cloud infrastructure for hosting, data storage (handling audio/video uploads), and processing remains essential, echoing the needs of the previous ventures. While migrating entire infrastructures between major providers (Amazon Web Services, Google Cloud, Azure) is complex and costly, creating potential lock-in, the continued availability of cost-effective alternatives like Hetzner provides leverage against excessive pricing or unfavorable terms from the giants. Payment processing dependencies remain unchanged, characterized by numerous options and low supplier power, as detailed in the initial analysis.

In summary, the reliance on core AI (both transcription and LLMs) and cloud technologies from relatively concentrated providers exert notable supplier pressure. This pressure is balanced by the increasing ease of switching between LLM providers thanks to abstraction tools, the competitive nature of non-specialized cloud services, and the low power of payment processors. Consequently, the overall supplier power for this educational generation service is assessed as moderate.

Buyer Bargaining Power

The bargaining power of buyers in this niche in Ukraine is relatively high. Similar to the research on the market of transcription applications for individuals with hearing loss, the focus is on the Ukrainian market. The target audience is students. The article published by Ukrainska Pravda examines the dynamics of university student enrollment in the country for the period between 2016 and 2024 (Krechetova). According to the author, the data for the publication was provided by the state enterprise Inforesurs. The findings indicate that the number of applicants for bachelor’s and master’s degrees is approximately 197 thousand and 135 thousand, respectively. These are the lowest figures within the analyzed period. Consequently, the decreasing number of potential buyers makes their bargaining power stronger. Moreover, the requirement for the availability of lecture recordings further narrows down the number of suitable universities and students.

ACTIONS TAKEN

The findings of the market research partially validate the idea. The competitors observed, such as *Gizmo*, *Knowt*, and *Study Fetch*, offer their customers to utilize AI agents to create notes, quizzes, and flashcards with the study materials. The problem stated in the idea subsection is similar. However, it is essential to inspect whether this issue exists among the students of the Kyiv School of Economics before investing resources in the development of the prototype. Moreover, the additional validation may provide valuable insights regarding the features needed in a future product.

In order to test the existence of the problem among the students of the Kyiv School of Economics, the engagement funnel was utilized. The path of the potential customers began with a short message in the group chat within their messaging application. This text provided a brief overview of the product along with a hyperlink to its landing page. The web page included a more detailed explanation and a call-to-action button inviting people to gain access to the solution. Upon clicking the button, the students were redirected to a questionnaire. The quantitative metric chosen to assess the problem's existence was conversion rate.

The survey at the end of the funnel was included to gather additional information regarding the target audience. The first questions were about the year of the study and academic programs. Then, the respondents were asked about their frequency and reasons for writing notes on the lessons and watching recordings of the lectures. The next part was added to define the most popular subjects among the students and prioritize the most anticipated features of the future product. Also, people were asked about their need for the product on a scale from 1 to 5. Finally, the questionnaire was used to gather the contacts in order to inform the potential customers about the release in the future.

The results of the survey confirmed the existence of the problem among the target audience. The average value of the conversion rate of the call-to-action button was almost 30%. The metric was calculated as the number of clicks on the button in order to gain access to the product divided by the total number of views of the message. The threshold of 20% was set before the test. Since the actual value of the conversion rate was higher than that, the idea was considered to be successfully validated. Additionally, the questionnaire revealed that the target audience expressed the most significant interest in the creation of notes and supplementary quizzes, while flashcards were less popular. Furthermore, the percentage of people who completed the survey was calculated to predict the number of paying customers. At that moment, the value was approximately 15%. Therefore, it was decided to develop and test a Minimum Viable Product (Appendix O).

The audience selected to test the Minimum Viable Product consisted of the students enrolled in three specific subjects. These disciplines were chosen for their theoretical nature. Processing calculation-centered lessons could result in less quality of the learning materials generated and more resources needed to provide them. Also, the decision to develop the initial solution for the limited audience and range of disciplines was made since supporting more subjects could require more resources to generate the learning materials. Therefore, in order to minimize the investments, the product was initially developed and tested on a limited group of students.

The initial product was designed for the aforementioned audience. Lecture recordings of the chosen subjects were first transcribed. These transcripts were then used to generate notes and supplementary quizzes. Flashcards were not added since they were not popular among the survey respondents. Each student in the test group was granted access to the subjects they were enrolled in. As a result, upon registering for the web application, users could see a list of disciplines and corresponding lessons available. After choosing one of them, the users could read generated notes and supplementary quizzes designed to assess their knowledge.

In order to test the solution described, it was decided to measure the willingness of users to pay for the use of it. Firstly, the students enrolled in the selected subjects were notified about the launch of the product. Upon registration, they were granted access to the generated learning materials of the subjects so that they could have enough time to discover the value proposition. After this period, they were notified that the learning materials for new lesson recordings were added to the website. However, upon attempting to access them, the users were redirected to the webpage with information that the free access limit had ended and offered to purchase full access to the course materials. The next page revealed the price and provided users with the option to pay. The design of such a funnel made it possible to collect information about the behavior of users and derive their willingness to pay for the solution.

The results indicated that the developed product solved the addressed problem of the users. Preliminarily, 60% of students who registered on the website chose to make a purchase on the webpage where the price was not disclosed. This figure confirms the willingness to pay for the use of the product. However, only 10% of registered users did the same on the page where the price of 199 hryvnias per course was displayed. Therefore, it appears as if only this percentage of users consider the value proposition sufficient to justify paying this price. It is worth noting that the small size of the

sample and the very limited MVP presented are potential limitations here. Based on these findings, it was decided to conduct additional testing with a larger number of users and continue refining the idea with the intention of broadening the market (Appendix P).

LEGAL CONSIDERATIONS

In Ukraine, three legal forms can be utilized to operate the business described. These alternatives are Sole Proprietor (Fizychna Osoba–Pidpryiemets) under the third group of the simplified taxation system, Limited Liability Company under the third group of the simplified taxation system, and Limited Liability Company under the general taxation system. Formally, all three options are permitted for engaging in business activities with individuals, Ukrainian legal entities, and foreign legal entities. Also, operating under these forms allows the hiring of employees. However, due to significant differences in taxation, financial limitations, and administrative complexity, a more detailed analysis is required to select the optimal variant.

The Limited Liability Company under the general taxation system does not have revenue limits, whereas the maximum revenue of the other options is 1167 Ukrainian minimum wages — more than 9 million hryvnias (Tatchyn). Also, these two forms are required to pay 6% tax on their gross revenue: 5% regular tax and 1% military levy (“Viiskovyi zbir”). In contrast, the Limited Liability Company, under the general taxation system, pays an 18% tax on net profit(Barbashyn and Syroid). Additionally, in both types of Limited Liability Companies, dividends are subject to 10% tax: 5% regular tax and 5% military levy (Diachkina). Moreover, their administrative complexity requires the employment of accountants and legal advisors. However, the ability to create special internal rules allows such entities to attract investments and guarantee the execution of shareholder agreements.

By analyzing the advantages and disadvantages of different legal forms, it can be concluded that using Sole Proprietor is optimal for the early stage of business due to its low administrative complexity and favorable taxational terms. However, the growth of gross revenue and the need to attract investments may require establishing a Limited Liability Company under the general taxation system in the future. Also, businesses in the technology sector have the option to obtain Diia.City residency. One of the main advantages is the option to pay a 9% tax on distributed profit instead of an 18% tax on net profit (“Diia.City Main Page”).

FINANCIAL MODEL

This section details the financial model underpinning the final product pivot: an AI-powered service generating notes, quizzes, and flashcards from lecture recordings. The analysis outlines projected fixed and variable costs, estimates potential revenues, and establishes key financial viability metrics. The cost associated with initial software development is omitted from this model. This simplification reflects the authors' possession of the necessary technical skills to build the product without incurring direct development expenses.

As a technology-intensive service, the primary operational expenditures stem from server hosting and AI processing. These constitute the most substantial ongoing costs, both fixed and variable. To ensure cost efficiency, vendors providing the most advantageous balance of performance and pricing were selected.

It is pertinent to acknowledge that $ 5000 in Azure Cloud credits were secured through the Microsoft for Startups program ("Microsoft for Startups Founders Hub"). To secure the credits, an application with all the necessary information about our project was filled out, and a website was developed as per the requirements of the program. While these credits significantly reduce immediate expenses, the subsequent calculations deliberately utilize standard market rates for hosting and AI services. This approach provides a realistic projection of the venture's financial requirements once the initial credits are exhausted.

Regarding taxation, it is assumed that relevant taxes are incorporated within the cost estimations provided by suppliers. Taxes related to revenue generation are explicitly calculated and discussed within the *Revenues* subsection.

Fixed Costs

The table below outlines the fixed costs of this venture for the first year of operations (Table 1). The costs are taken from the official websites of the services ("Porkbun Main Page"; "Pricing"; "Cloud").

Table 1: Project's fixed costs

Description	**Provider**	**Cost**
The zvuk.ai domain	Porkbun	$ 72.40, yearly
Corporate email	Google Workspace	$ 6.30, monthly
Virtual Private Server	Hetzner (CAX31 instance)	$ 14.09, monthly
Total		**$ 317.08, yearly**

Variable Costs

AI processing of lecture recordings constitutes the project's sole variable cost. Cost estimation hinges on key assumptions: an average 80-minute lecture translates to 15 000 English words, or approximately 11 250 tokens, using OpenAI's standard conversion rate (1 token ≈ 0.75 words) ("Key Concepts"). These assumptions form the basis for calculating the processing cost per individual lecture recording. Prices are presented below (Table 1). The data on prices is taken from the official websites of the services ("Azure AI Speech Pricing"; "Azure OpenAI Service Pricing").

Table 2: Project's variable costs

Description	**Provider**	**Unit cost**	**Units**	**Total cost**
Transcription	Azure AI Speech, Batch, East US	$ 0.18 per hour	1.34 hours	$ 0.2412
Notes generation	GPT-4o 2024-1120	$ 2.50/M input, $ 10.00/M output	11 250 input, 1400 output	$ 0.0421
Title & tags generation	o3 mini 2025-01-31 (high)	$ 1.10/M input, $ 4.40/M output	1700 input, 1000 output	$ 0.0062
Quiz generation	o3 mini 2025-01-31 (high)	$ 1.10/M input, $ 4.40/M output	1800 input, 7000 output	$ 0.0327
Total				**$0.3223**

Extrapolating this, processing a typical 4 ECTS discipline, encompassing roughly 40 lectures, incurs an average total cost of **$ 12.89**.

This total discipline cost represents a shared expense. As multiple students attend the same course, the variable cost of a lecture should be divided by the number of people paying for the access to this course on the platform — resulting in the *cost per course taker*.

Securing a minimum number of paying students per course becomes essential for economic viability, directly reducing the variable cost of processing a lecture per paying customer. The financial implications of this model are detailed further in the *Revenues* analysis.

Revenues

A transactional, per-course pricing model underpins the revenue strategy. This approach directly addresses student needs, as individuals often seek support for specific, challenging courses rather than requiring blanket access. Furthermore, it avoids artificially capping revenue from students taking numerous courses, unlike flat-rate subscriptions. This structure justifies a focused value proposition, enabling a marginally higher price point per unit.

Informal customer interviews set the optimal price per course at approximately $\mathbf{\$\,5}$ (equivalent to around 200 UAH). To ensure economic viability, we establish a minimum demand threshold: a discipline will be offered if at least four students express interest in paying for it. This safeguard ensures that initial fixed and variable costs are adequately covered.

The demand estimation for the first year of operation is intentionally conservative. Kyiv School of Economics currently enrolls approximately 1100 students across BA, MA, and MBA programs (“Kyivska shkola ekonomiky [Kyiv School of Economics]”), of which 606 new enrollments recorded as of September 2024 (“Universytet Kyivska shkola ekonomiky”). Based on this pattern, total enrollment is projected to reach around 1800 students by the end of 2025. Given our targeted marketing through KSE’s internal channels, total product awareness among students is a realistic assumption.

It was calculated that on a student curriculum, an average discipline contributes around 3.6 ECTS credits. Activities that count toward ECTS credits but not disciplines take up around 40 credits (practice each summer semester and a capstone project). Hence, the average KSE student undertakes approximately 55.55 disciplines throughout their 4-year study.

{932.042 24}

Assuming that in the first year 20% of all students will pay for the platform's access to 25% of their subjects (and all of those disciplines have at least four students interested), 1250 transactions would be made in the first year, totalling **$ 6250.00** in revenue.

Assuming those transactions would be distributed across 80 disciplines, The incurred costs constitute $ 317.08 (fixed) + $ 928.43 (variable) = **$ 1245.52**

As mentioned previously, the FOP tax includes 6% plus unified social tax, which is calculated as 22% of the minimum wage ("3 Hrupa"). In 2025, minimum wage is 8000 ("Minimalna Zarplata"), hence the unified social tax is 1760 UAH, or approximately $ 44. So after subtracting the FOP tax **$ 4660.21** are left.

Finally, applying the assumed revenue sharing with the university of 20%, the total profit for the first year should approximate to **$ 3728.16**.

As the predicted monthly profit of $ 310.68 (12 427.22 UAH) is just around 55% higher than the minimum wage in Ukraine, this preliminary financial analysis highlights the need for scale beyond the university of Kyiv School of Economics, possibly targeting not only students pursuing formal education, but learners from all walks of life.

NEXT STEPS

After the validation of the Minimum Viable Product, high-level objectives were established for further development. In order to develop the business efficiently and effectively, the Diamond-and-Square framework is utilized (Eisenmann). It consists of eight key elements divided into two sections: the Square represents external factors of the success of the business, and the Diamond highlights the fundamental components of the business model. The latter elements are Customer Value Proposition, Go-to-Market Strategy, Profit Formula, Operations and Technology. Since they are the main determinants of the success of the business, the plan is divided into four phases, each dedicated to the development of one of these elements.

The first phase is about determining the Customer Value Proposition and its validation. Work on this stage should be started only when the problem is considered to exist. Qualitative research methods like surveys and interviews can then be employed to gather relevant information. This knowledge is used to design and develop the Minimum Viable Product that addresses the problem. The product is tested with users to determine whether the Customer Value Proposition solves their

needs. Some steps of this phase can be repeated multiple times before the Customer Value Proposition is validated.

The second phase is focused on the marketing of the product. Firstly, to define the ideal buyer persona, the insights from qualitative research in the first phase and interviews with active users of the Minimum Viable Product are analyzed. Then, research should be conducted to identify marketing channels for reaching the target audience. Tests are conducted to assess the effectiveness of these channels. The decision on which of them should be pursued is based on scalability and values of key metrics like Customer Acquisition Costs and Return on Investment. Since the options can be unprofitable, the aforementioned testing process can be iterative. In conclusion, businesses find ways to attract new customers profitably.

Ensuring the profitability of the product is the objective of the third phase. In order to achieve sustainable growth, Customer Lifetime Value (LTV) should be at least three times higher than Customer Acquisition Costs (CAC). Customer Lifetime Value is "the average revenue a single customer is predicted to generate over the duration of their account," while Customer Acquisition Costs are the average costs of attracting a customer ("LTV:CAC Ratio"). This phase does not have a predetermined set of steps but rather involves conducting different experiments to optimize key financial metrics. Firstly, founders should consider changing the monetization model of the product. In some cases, adopting a transactional model can be more profitable than charging a monthly subscription. Another area that requires multiple tests is the optimization of product prices. Also, the customer acquisition costs of marketing channels can be optimized by refining advertising messages or creatives. The result of these steps is a profitable business that is ready for further growth.

The last major phase focuses on the optimization of the business operations. The main objective is to ensure that attracting a significant number of new customers does not have adverse effects on the quality of the product and user satisfaction. Similarly to profitability optimization, this phase does not follow a predetermined set of steps. However, there are some recommendations, such as creating and monitoring metrics for areas of business to identify inefficiencies. Automating repetitive tasks can also improve processes. Moreover, enhancing technical infrastructure is a standard method to meet increased demand. Consequently, the phase is needed to make further scaling successful.

In order to create a structure for the next steps for the product, the Objectives and Key Results framework is utilized(Grove). It is selected since it is focused on results and allows flexibility in

choosing the steps needed to achieve them. The team members define objectives for a predetermined period of time and establish key results to measure progress. Key initiatives are also outlined as potential steps needed to achieve the objectives. In this case, the objectives are defined for the period of six months. Although they reflect the plan derived from the Diamond-and-Square framework, the last part focuses on the operations and technology of business, which is beyond the scope of the selected timeframe.

- Objective 1: Validate the Customer Value Proposition on a significant number of users
  - Key Results:
    - At least 10% of users are converting to purchase
    - The results are validated on at least 300 registered users
  - Key Initiatives:
    - Conduct customer interviews
    - Implement new features based on insights from the interviews
    - Support more courses
    - Attract more students
- Objective 2: Develop Go-to-Market Strategy
  - Key Results:
    - Obtain more than three marketing channels with a Return On Investment > 1
    - Each of the channels
    - Each of the tested channels attracted at least 300 registered users
  - Key Initiatives:
    - Define the ideal buyer persona
    - Brainstorm ideas for new marketing channels
    - Establish partnerships with more universities
    - Create advertising messages and creatives for new campaigns
    - Analyze the performance of campaigns and channels
- Objective 3: Ensure Profitability
  - Key Results:
    - LTV to CAC ratio is at least 3
  - Key Initiatives:

- Brainstorm ideas for increasing the LTV to CAC ratio
- Conduct tests to determine the most profitable monetization model
- Conduct tests on product pricing

The first objective is to validate the Customer Value Proposition on a bigger number of users. In the previous section, the developed product was considered to be validated. However, the conclusion regarding the willingness of users to pay for the use of it should be based on a larger sample of people to be representative. Also, new results can be less than the threshold of 10%. Therefore, customer interviews may be conducted to identify product weaknesses and implement new features. Consequently, the objective is to validate the Minimum Viable Product on a larger sample.

The following two objectives reflect the second and the third phases of the plan described before. The first indicates that profitable marketing channels are needed to attract more customers. In order to achieve this, it is necessary to analyze the data gathered from users of the Minimum Viable Product to identify the characteristics of those who need this solution the most. These insights are used in the following steps to find channels that can reach similar people. The last objective is to achieve profitability by ensuring the LTV to CAC ratio is at least 3. As mentioned before, this process involves conducting multiple experiments. Although conducting tests on the pricing and monetization model of the product may be successful, brainstorming and testing more ideas may be necessary.

Conclusion

This thesis documents an entrepreneurial journey pursuing three distinct ventures, each leveraging Whisper's advanced transcription capabilities. The initial concept — a real-time transcription app with speaker diarization, designed for individuals with hearing loss — faced serious market challenges. Although technically feasible and socially valuable, limited willingness to pay in Ukraine and lukewarm institutional interest rendered monetization impractical. The resulting pivot toward transcribing offline meetings of businesses and governmental organizations initially showed promise, with Kryvyi Rih City Council as an encouraging adopter. However, intense competitive pressure from rivals offering more features at lower prices quickly eliminated opportunities. This competitive landscape compelled a subsequent shift toward the sector of education. The refined product offered lecture-generated notes and quizzes, leveraging transcription with the integration of large-language models to accelerate and deepen students' learning process.

Although modest in the initial scale, significant validation was achieved. High engagement — 30% conversion rate from initial messages, 60% willingness-to-pay at zero-disclosure, and a solid 10% purchase rate at explicit pricing — in preliminary testing of the MVP product affirmed problem-solution fit. Despite this validation, the financial model — as projected within this thesis — highlights limited potential revenues at a single university. Broader market expansion into larger institutions and independent learners is essential to assure meaningful profitability.

Strategically, the analysis throughout all ventures consistently revealed the critical role market forces play in shaping entrepreneurial outcomes. Porter's Five Forces concretely guided assessments of competitive rivalry, substitution danger, significant supplier-driven constraints (notably AI-model providers), the elevated bargaining power of customers across nearly all product contexts, and entry barriers tied strongly to technical complexity. The iterative, evidence-based pivots embodied textbook lean startup principles: concepts evolved rapidly through explicit hypothesis tests, spanning technical prototypes, detailed competitor analysis, stakeholder interviews, rigorous financial modeling, and direct user engagement.

Fundamentally, this study provides granular visibility into the nuanced realities underlying innovative, tech-focused ventures. Opportunities identified through theoretical frameworks require decisive assessment in practical tests against real-world constraints and competitive markets. The core

lessons were clear and consistent. Product creation alone — no matter how technically sophisticated or socially beneficial — is insufficient without carefully confirming customer demand, willingness-to-pay, economic feasibility, and scalable distribution. Successful ventures are not simply built around advanced technologies — they are constructed around tested, validated value propositions addressing clearly defined market needs.

In sum, beyond merely documenting a succession of entrepreneurial experiences, this thesis provides critical insight for innovation ventures leveraging technologies of generative AI. While initial results for the education-focused iteration appear promising, long-term success depends on continued market testing and strategical scaling into broader markets. The iterative journey documented here illustrates that genuine entrepreneurial resilience lies not in stubborn adherence to original ideas, but in disciplined responsiveness to concrete market evidence: pivoting swiftly toward genuine market-validated opportunities.

LIST OF FIGURES

LIST OF TABLES

Works Cited

"About | Sensor Tower." *Sensor Tower*, https://sensortower.com/about. Accessed 4 Nov. 2024.

"App Store." *Apple*, https://www.apple.com/app-store/. Accessed 4 Nov. 2024.

"AppMagic Main Page." *AppMagic*, https://appmagic.rocks/?hl=en. Accessed 4 Nov. 2024.

"Azure AI Speech Pricing." *Microsoft Azure*, https://azure.microsoft.com/en-us/pricing/details/cognitive-services/speech-services/. Accessed 22 Apr. 2025.

"Azure OpenAI Service Pricing." *Microsoft Azure*, https://azure.microsoft.com/en-us/pricing/details/cognitive-services/openai-service/. Accessed 22 Apr. 2025.

Barbashyn, Serhii, and Iryna Syroid. "*Varyanty vedennya IT-biznesu v Ukraini z pil'hovymy 5 % podatkiv* [Options for Running IT Business in Ukraine with Preferential 5% Taxes]." *Barbashyn Law Firm*, 13 Sept. 2023, https://barbashyn.law/statti/varianty-vedennya-it-biznesu-v-ukrayini-z-pilgovymy-5-podatkiv/.

Bredin, Hervé. *Pyannote.Audio Speaker Diarization Toolkit*. pyannote. GitHub, https://github.com/pyannote/pyannote-audio. Accessed 27 Apr. 2025.

Bromberg, Michael. "Herfindahl-Hirschman Index (HHI): Definition, Formula, and Example." *Investopedia*, 12 June 2024, https://www.investopedia.com/terms/h/hhi.asp.

"Cloud." *Hetzner*, https://www.hetzner.com/cloud. Accessed 22 Apr. 2025.

"Deafness and Hearing Loss." *World Health Organization*, 26 Feb. 2025, https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

Diachkina, Anastasiia. "*Zarplaty, depozyty, orenda: khto i z yakykh dokhodiv mae splachuvaty viiskovyi zbir* [Salaries, deposits, rent: who and from which incomes must pay the military levy]." *Ukrainska Pravda*, 27 Jan. 2025, https://epravda.com.ua/finances/viyskoviy-zbir-2025-hto-i-z-yakih-dohodiv-maye-platiti-802447/.

"Digital Solutions." *NGO "Public Movement "Social Unity,"* 7 Nov. 2024, https://se.org.ua/en/cifrovi-rishennya/.

Eisenmann, Thomas R. *Why Startups Fail*. Currency, 2020.

Fernandez, Joaquin. "The Leading Generative AI Companies." *IoT Analytics*, 4 Mar. 2025, https://iot-analytics.com/leading-generative-ai-companies/.

"Google Play Main Page." *Google Play*, https://play.google.com/store/apps?hl=en. Accessed 4 Nov. 2024.

Grove, Andrew S. *High Output Management*. 2nd Vintage Books ed, Vintage, 1995.

"Hearing Aids vs Cochlear Implants." *Hearing Aid UK*, https://www.hearingaid.org.uk/hearing-loss-awareness/hearing-aids-vs-cochlear-implants. Accessed 7 Nov. 2024.

"Introducing Whisper." *OpenAI*, 21 Sept. 2022, https://openai.com/index/whisper/.

"Key Concepts." *OpenAI*, https://platform.openai.com. Accessed 22 Apr. 2025.

"Key Facts About AppMagic." *AppMagic*, https://appmagic.rocks/about. Accessed 4 Nov. 2024.

Kitaiti, Makoto. *Whisper.Cpp*. 2022. ggml. GitHub, https://github.com/ggml-org/whisper.cpp. Accessed 27 Apr. 2025.

Krechetova, Diana. *"Operatsiia 'dyplom'. Chy spratsiuvala antiukhyliantska osvitnia polityka MON? Infohrafika* [Operation 'Diploma': Has the Ministry of Education's Anti-Draft Evasion Policy Worked? Infographic]." *Ukrainska Pravda*, 28 Jan. 2025, https://life.pravda.com.ua/society/antiuhilyantska-politika-mon-shchodo-vishiv-spracyuvala-vidpovid-ye-u-statistici-306112/.

"*Kyivska shkola ekonomiky* [Kyiv School of Economics]." *Osvita.UA*, https://osvita.ua/vnz/guide/37271/. Accessed 22 Apr. 2025.

"LTV:CAC Ratio." *Geckoboard*, https://www.geckoboard.com/best-practice/kpi-examples/ltv-cac-ratio/. Accessed 1 Apr. 2025.

"Microsoft for Startups Founders Hub." *Microsoft*, https://foundershub.startups.microsoft.com/signup. Accessed 27 Apr. 2025.

"*Minimalna Zarplata v Ukrayini* [Minimum Wage in Ukraine]." *Minfin*, https://index.minfin.com.ua/ua/labour/salary/min/. Accessed 22 Apr. 2025.

"Porkbun Main Page." *Porkbun*, https://porkbun.com/tld/ai. Accessed 22 Apr. 2025.

"Pricing." *Google Workspace*, https://workspace.google.com/pricing. Accessed 22 Apr. 2025.

"*Pro Kokhlearnu Implantatsiiu* [About Cochlear Implantation]." *Kind InterSlukh*, https://kind.kiev.ua/uk/page/o-kohlearnoy-ymplantacyy. Accessed 7 Nov. 2024.

"Providers." *LangChain*, https://python.langchain.com/docs/integrations/providers/. Accessed 27 Apr. 2025.

"Quizlet Main Page." *Quizlet*, https://quizlet.com/latest. Accessed 27 Feb. 2025.

"Responsibly Sourced Data." *Sensor Tower*, https://sensortower.com/responsibly-sourced-data. Accessed 4 Nov. 2024.

"RunPod Main Page." *RunPod*, https://www.runpod.io/. Accessed 27 Apr. 2025.

"Similarweb's Data Accuracy." *Similarweb*, https://support.similarweb.com/hc/en-us/articles/360002219177-Similarweb-s-Data-Accuracy. Accessed 27 Feb. 2025.

"*Slykhovi Aparaty* [Hearing Aids]." *Tsentr Slukhu "Yevroton,"* https://evrotonsluh.com.ua/ua/sluhovye-apparaty?sorting=by_price_asc. Accessed 7 Nov. 2024.

Smoliakov, Yehor. *Speech Recognition & Synthesis for Ukrainian.* 1.0.0, Feb. 2025. GitHub, https://doi.org/10.5281/zenodo.14895258.

"Ukrainian Association of Cochlear Implant Users Main Page." *Ukrainian Association of Cochlear Implant Users*, https://sluh-e.com.ua/. Accessed 7 Nov. 2024.

"Ukrainian Society of the Deaf Main Page." *Ukrainian Society of the Deaf*, https://utog.org/index/. Accessed 15 Mar. 2025.

United Nations (2024). *World Population Prospects 2024: Summary of Results*. UN DESA/POP/2024/TR/NO. 9. New York: United Nations. Accessed 27 Feb. 2025.

"*Universytet Kyivska shkola ekonomiky — 2024* [University of Kyiv School of Economics — 2024]." *Vstup.OSVITA.UA*, https://vstup.osvita.ua/y2024/r27/3924/. Accessed 22 Apr. 2025.

"Vidchui Main Page." *Vidchui*, https://vidchui.org/. Accessed 7 Nov. 2024.

Vidyakin, Mykhaylo. *LinkedIn Profile*. LinkedIn, https://www.linkedin.com/in/mvidyakin/. Accessed 27 Feb. 2025.

"*Viiskovyi zbir dlia yurosib 3 hruppy EP* [Military levy for legal entities of the 3rd group under the single tax]." *7eminar*, 20 Mar. 2025, https://7eminar.ua/news/5881-viiskovii-zbir-dlya-yurosib-3-grupi-jep.

Walton, Jarred. "OpenAI Whisper Audio Transcription Benchmarked on 18 GPUs: Up to 3,000 WPM." *Tom's Hardware*, 11 May 2023, https://www.tomshardware.com/news/whisper-audio-transcription-gpus-benchmarked.

"*3 Hrupa* [3rd Group]." *Diia*, https://diia.gov.ua/tax-systems/3-grupa. Accessed 22 Apr. 2025.

Appendix A

Market Distribution of Mobile Applications by Downloads (Real-Time Transcription for Individuals with Hearing Loss)

Available at: https://docs.google.com/spreadsheets/d/1Gm8lqFQWunQlQ_QqHLXCSLsl3ZOxpjo7idZr_mjsUM0/edit?gid=1195374650#gid=1195374650

Appendix B

Comparison of Features and Prices (Real-Time Transcription for Individuals with Hearing Loss)

Available at: https://docs.google.com/spreadsheets/d/1Gm8lqFQWunQlQ_QqHLXCSLsl3ZOxpjo7idZr_mjsUM0/edit?gid=1149162440#gid=1149162440

Appendix C

Market Entry Years (Real-Time Transcription for Individuals with Hearing Loss)

Available at: https://docs.google.com/spreadsheets/d/1Gm8lqFQWunQlQ_QqHLXCSLsl3ZOxpjo7idZr_mjsUM0/edit?gid=1774720204#gid=1774720204

Appendix D

Competitors (Offline Meetings Transcription)

Available at: https://docs.google.com/spreadsheets/d/1zYXK2IpgR0z1n9T0YOj24DbzrTYqZGkYs7nalwfPRko/edit?pli=1&gid=0#gid=0

Appendix E

Market Distribution of Web Applications by Average Monthly Unique Visitors (Offline Meetings Transcription)

Available at: https://docs.google.com/spreadsheets/d/1zYXK2IpgR0z1n9T0YOj24DbzrTYqZGkYs7nalwfPRko/edit?pli=1&gid=1087060276#gid=1087060276

Appendix F

Market Distribution of Mobile Applications by Downloads (Offline Meetings Transcription)

Available at: https://docs.google.com/spreadsheets/d/1zYXK2IpgR0z1n9T0YOj24DbzrTYqZGkYs7nalwfPRko/edit?pli=1&gid=949620670#gid=949620670

Appendix G

Comparison of Features and Prices (Offline Meetings Transcription)

Available at: https://docs.google.com/spreadsheets/d/1zYXK2IpgR0z1n9T0YOj24DbzrTYqZGkYs7nalwfPRko/edit?pli=1&gid=740534118#gid=740534118

Appendix H

Market Entry Years (Offline Meetings Transcription)

Available at: https://docs.google.com/spreadsheets/d/1zYXK2IpgR0z1n9T0YOj24DbzrTYqZGkYs7nalwfPRko/edit?pli=1&gid=1054384629#gid=1054384629

Appendix I

Competitors (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=2074919244#gid=2074919244

Appendix J

Market Distribution of Web Applications by Self-Reported Users (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=2099771440#gid=2099771440

Appendix K

Market Distribution of Web Applications by Average Monthly Unique Visitors (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=346960487#gid=346960487

Appendix L

Market Distribution of Mobile Applications by Downloads (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=968014129#gid=968014129

Appendix M

Comparison of Features and Prices (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=1611430698#gid=1611430698

Appendix N

Market Entry Years (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/12iCGmMZekwxZBRYPWR3mhGb9x85gdm4fDQHoqko14Oc/edit?pli=1&gid=745059047#gid=745059047

Appendix O

Idea Validation Results (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/1eUOHt2sp7RlSKcSxsRrC1tEx-SNhYSc6t-SG-D_2EbQ/edit?gid=0#gid=0

Appendix P

Minimum Viable Product Validation Results (Generation of Notes, Quizzes, and Flashcards based on Lecture Recordings)

Available at: https://docs.google.com/spreadsheets/d/1eUOHt2sp7RlSKcSxsRrC1tEx-SNhYSc6t-SG-D_2EbQ/edit?gid=1742915648#gid=1742915648