ANALYSIS OF THE EMERGING ARTIFICIAL INTELLIGENCE MARKET AND ITS GROWTH OPPORTUNITIES

by

Oleksandr Sundyeyev

A thesis submitted in partial fulfillment of the requirements for the degree of

MA in Business and Financial Economics

Kyiv School of Economics

2024

Thesis Supervisor: _____ Professor Maksym Obrizan

Approved by _____
Head of the KSE Defense Committee, Professor Olena Besedina

_____

_____

_____

Date _____

ACKNOWLEDGMENTS

The author wishes to show how the artificial intelligence market was established within the last year, elaborate on its technical background and value chain, as well as identify its growth opportunities and factors.

The current artificial intelligence market may be associated with the success of OpenAI's ChatGPT launch and success as it has triggered the emergence of the artificial intelligence market in 2023, which has been growing rapidly since that time.

The market can be segmented by its value chain, with more concentration observed at the beginning of the chain and less concentration at the end of the chain. The latter segment of the market, artificial intelligence services, presents most of the current opportunities in the market.

The thesis is further extended with the analysis of software and artificial intelligence developers' proposals at a freelance platform, UpWork, in the US. US region has been chosen as most of the emerging artificial intelligence market is concentrated there, it's structure is well defined and its rules may be later applicable in other countries like Ukraine.

Analyzing and comparing hire proposals of the two, software and artificial intelligence developers, may help in benchmarking and evaluating the predictability of the market for technical talent, as it is a critical resource for companies or investors who consider entering the artificial intelligence market.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABS Annual Business Survey

AI Artificial Intelligence

API Application Programming Interface

BFS Business Formation Statistics

BTOS Business Trends and Outlook Survey

CAGR Compounded Annual Growth Rate

EY Ernst & Young

GPT Generative Pretrained Transformer

GPU Graphics Processing Unit

HHI Herfindahl-Hirschman Index

IaaS Infrastructure as a Service

IP Intellectual Property

IT Information Technology

KPMG Klynveld Peat Marwick Goerdeler

LLM Large Language Model

MLOps Machine Learning Operations

OECD Organization for Economic Cooperation and Development

PWC PricewaterhouseCoopers

UI User Interface

US United States

USD United States Dollar

UX User Experience

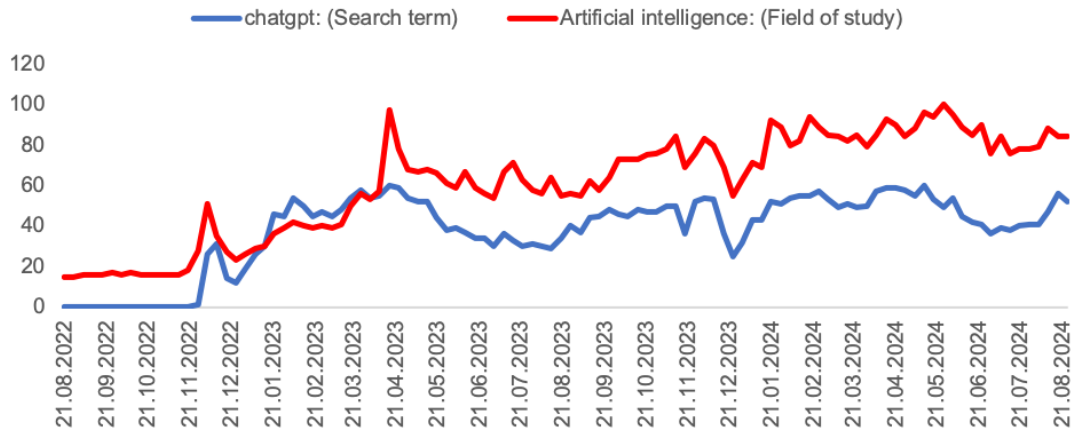
VC Venture Capital

CHAPTER 1. INTRODUCTION

Recent developments in AI technologies have drawn the interest of many news agencies like Reuters, CNN, The Economist, ABC News, Wall Street Journal, as well as of management consulting and audit companies, like McKinsey, PWC, KPMG, EY, Accenture many of which publish regular report on developments of AI and have dedicated respective services. Such developments even created a whole new profession, prompt engineering, which stands for practices of designing proper inputs for AI tools in order to get optimal outcomes of the interaction[1].

Many would agree that the most significant event in the recent years was the launch of Chat GPT, an AI-powered chatbot by OpenAI, reaching 1 million active user in 5 days after launch[2] and remain at the top 1 place in the productivity category on Apple AppStore. Spike in interest in AI is likely associated with the popularity of Chat GPT, at least in the US.

Figure 1. Daily Google searches of "chatgpt" and "Artificial intelligence" in US 2022-2024

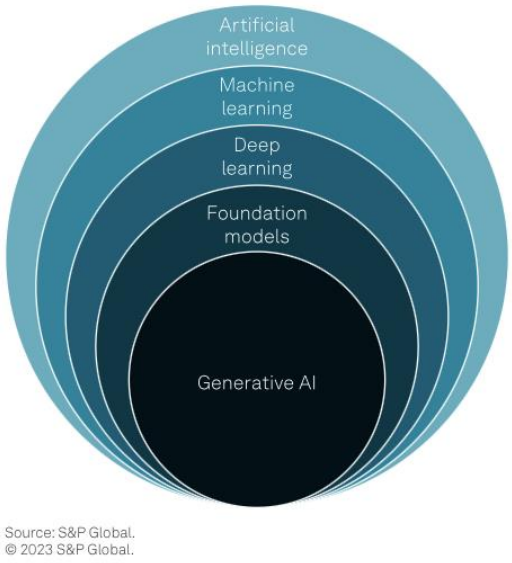


Source: Google trends, own visualization

OECD defines artificial intelligence as follows: "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."[3]

According to S&P Global[4], AI can be subdivided into the following subgroups: artificial intelligence, machine learning, deep learning, foundation models, generative AI. Earlier mentioned Chat GPT is an example of generative AI application based on OpenAI GPT foundation model.

Figure 2. The AI universe



Source: S&P Global, 2023

Rishi Bommasani et al. 2022[5] suggest that AI can be regarded as *general-purpose technology* (Bresnahan and Trajtenberg 1995 via Rishi Bommasani et al. 2022[6]), like steam engine and electricity, that can drive innovation and productivity, increase the standard of living in waves due to persuasiveness and improvement over time and due to the ability to spawn other innovations.

Foundation models, a term coined at Stanford, are usually large language models trained by big tech companies like OpenAI or Google as will be discussed in later chapters. Such models form a corner stone for AI applications providing a basic tool set for generative AI, for which a foundation model is "tuned" for specific applications[7].

According to McKinsey & Co[8], a US management consulting company that actively tracks the progress and economic activity around generative AI, generative AI comprises algorithms used to create content such as text, audio, video, code, etc. The company's 2022 survey has shown that both generative AI adoption and investments have doubled during the year. Examples of well-known generative AI applications can be Chat GPT, GitHub Copilot and Stable Diffusion.

Finally, in its recent article, JP Morgan[9] suggests that recent increase in AI companies' stock value should not be regarded as a "bubble", but viewed as a good investment opportunity with the view to transformation into the so-called "AI 2.0" where most of AI's unrecognized value will be in software and application with the focus on "AI adopters" in industries such as customer service, health care, finance and logistics.

Figure 3. Comparing today's price performance for the Nasdaq 100 and the AI Leaders to the Dot.com

Source: JP Morgan, 2024

JP Morgan mentions a notable case of Klarna, a buy now pay later company, which has introduced an AI assistant now handling two thirds of the company's customer service chats, reducing the costs by doing the equivalent work of 700 agents. This can be a good example of source of value by cost reduction for AI adopters.

As Ukraine prioritizes development in the sphere of AI at the governmental level, the goal of this study is to provide an overview of the current situation in the artificial intelligence market to Ukrainian investors, business decision-makers and other stakeholders concerned, using up-to-date available related research, studies and data to identify trends, patterns and benchmarks of the emerging AI market.

CHAPTER 2. INDUSTRY OVERVIEW AND RELATED STUDIES

2.1 Overview of AI market demand and supply sides

The author would like to start industry overview by providing a clearer view on the background business activities on the supply side as well as on the demand side of the AI adopters. It should be noted that available research and related studies determines AI adopters as businesses that adopt AI to improve their operations or to better serve their consumers as it will be discussed later.

US Census Bureau in its recent report "Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey" (March, 2024)[10] states that there has been a number of recent researches in AI diffusion in businesses using high-frequency Census Bureau's Business Formation Statistics (BFS) providing evidence of recent surge in AI-related applications by businesses. Besides that, AI businesses have played a major role in new business formations in year 2021-2023 in high-tech sectors. Though BFS does not provide information on AI diffusion among the general population of businesses, there is a number of related researches by firms like McKinsey and other researchers on indirect measurement of AI diffusion with the use of online resume and job postings, patenting behavior and large unrepresentative surveys. US Census Burau focuses its paper on the analysis of Business Trends and Outlook Survey (BTOS) for September 2023 to February 2024.

According to the paper, there has been a surge in AI use by businesses from 3.7% to 5.4% with an expected rate of 6.6% by Fall 2024. Large firms, especially in the information sector (18% of firms in information sector), present higher rates of AI use and the relationship between AI use rate and firm age is U-shaped. Most common AI applications in businesses are marketing automation (2.5% overall share, 28.4% share of users), virtual agents or chat bots (1.9% overall share, 21.6% share of users), natural

language processing (1.7% overall share, 19.3% share of users) while the most common reason for non-adoption is inapplicability of AI to a business (80.9%, firm-weighted) with far less common reason being the lack of knowledge on the capabilities of AI (7.3%, firm-weighted) as well as concerns about privacy or security (6.6%, firm-weighted). Many firms undertake organizational changes in order to adopt AI, such as development of new workflows, training staff on AI use, purchasing cloud services and storage, using vendors or consulting agencies to install and integrate AI, etc. AI users show higher performance and employment expansion if compared to other businesses. The paper also provides a snapshot of other statistical information, such as use of AI in businesses by states.

Another study conducted by Kristina McElheran et al., 2021[11], investigated the determinants of AI adoption across companies using Annual Business Survey (ABS) data with n=170,000 from US Census Bureau. In 2021, US nationwide AI use was low yet concentrated in highly important firms, it was an attribute of innovative, venture-funded and early growth firms. One of the key findings of the paper is that founder characteristics are important determinants of AI adoption with younger, more educated and more experienced leader being more likely to adopt rising AI technologies such as machine learning or vision, voice recognition, natural language processing and automated vehicles. The key findings of the paper are as follows.

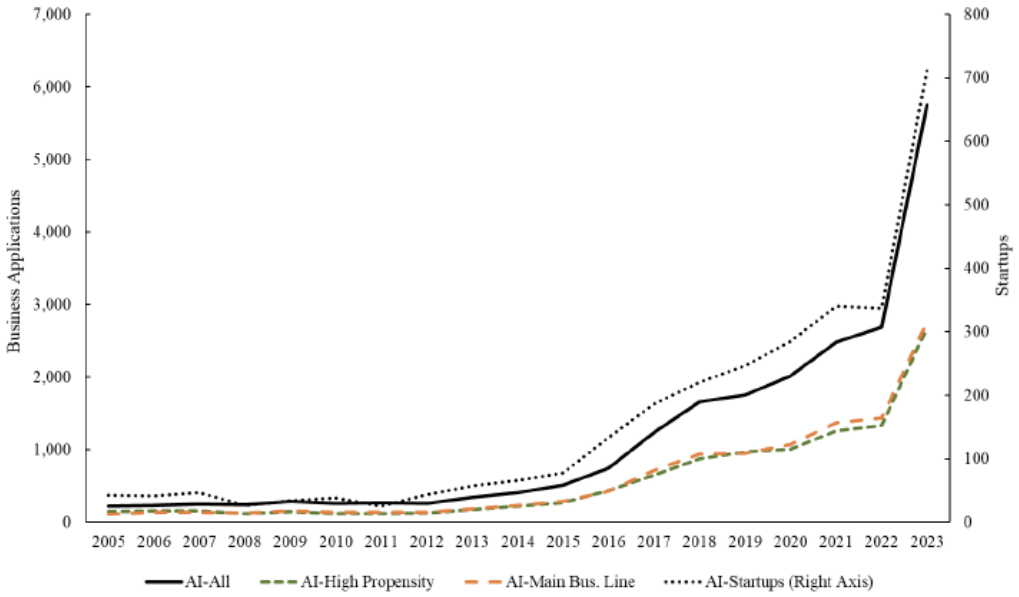
1. AI adoption was low on average and skewed towards large firms: there are limiting factors in adoption such as scale effects (amount of data required to utilize AI), clear hierarchy of technology adoption for proper AI deployment.
2. Observable firm characteristics have limited success in predicting AI adoption: only up to 14% of variation in adoption can be explained by models presented in the paper.
3. Owner's education, prior startup experience, and age are significant predictors of AI adoption.

4. Firms with high capitalization requirements are more likely to adopt AI: VC funding as a characteristic of high-growth innovative startups that may not be properly captured in the available data.
5. Innovativeness matter for AI adoption: firms with process innovation coefficient are 5.8% more likely to adopt AI and firms that identify IP protection as important have 5.7% higher probability of adoption.

More information on the technological hierarchy, AI adoption rates highlighting various characteristics of firms and other information can be found in the paper. Table 7 of the paper presents the results of a multiple linear regression of the determinants of AI adoption.

Another study conducted by US Census Bureau in 2024 "Starting Up AI"[12] presents recent findings drawn from comprehensive administrative data on business applications over years 2004-2023. It was discovered that new AI business applications were stable over years 2004-2012, began to rise after 2012, rose faster after 2016 and took a leap after 2023 (Figure 4).

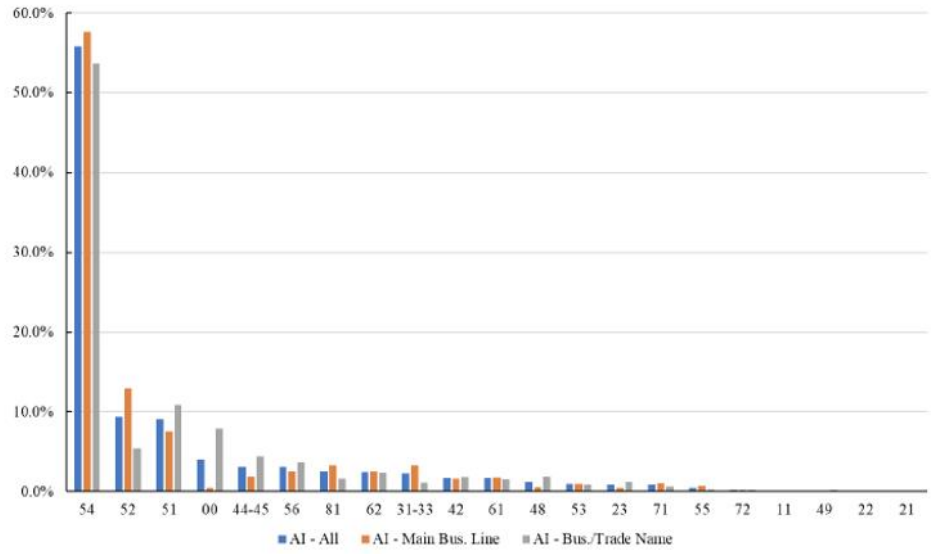
Figure 4. AI business applications and formations by year



Source: "Starting Up AI", US Census, 2024

The largest share of applications (56%) are in professional scientific and technical services, followed by finance and insurance and informatics (9.4% and 9.1% share respectively).

Figure 5. The share (%) of AI business applications by sector



Notes: Includes all AI business applications 2004q3-2023m10. Sectors are in decreasing order of AI business applications (AI-All) share from left to right. Sector code "00" refers to applications that are not classified into any NAICS sector by the industry assignment algorithm used for BFS. AI-Main. Bus Line and AI-Bus./Trade Name shares for NAICS 21 and 22 are suppressed to protect confidentiality.

Source: "Starting Up AI", US Census, 2024

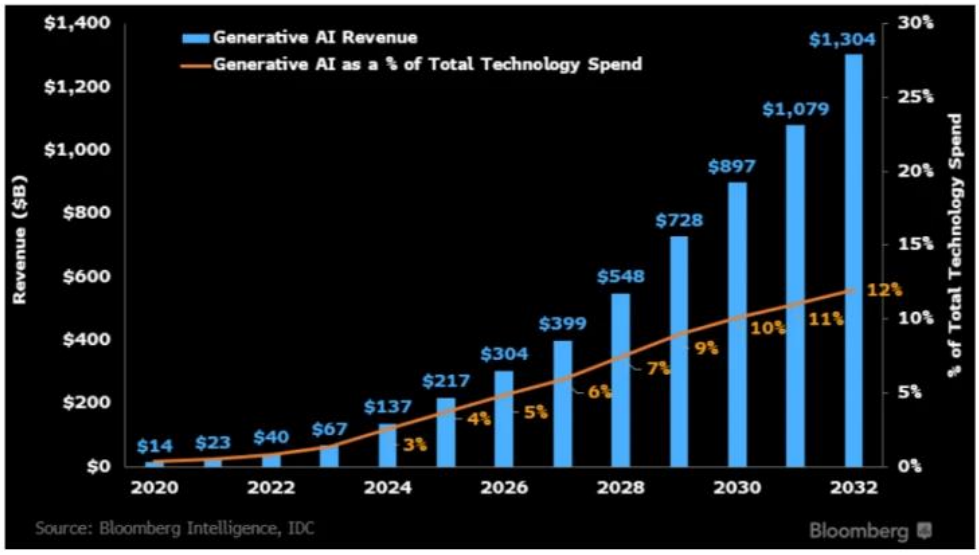
2.2 Generative AI market overview

Following the discussion of AI market demand and supply sides, the author would like to focus on and determine the generative AI market as generative AI is the technology providing direct value to its serviced businesses.

It should be noted that as Sam Altman, CEO of OpenAI, the developer of Chat GPT, recently discussed the possibilities of AI at Stanford[13], he noted that the main opportunities in AI market are concentrated around 1) building AI infrastructure and 2) figuring out what to do with the new technology. This seemingly simplistic overview of the AI opportunities can be clearly reflected on the generative AI market, that has been subject of discussion in business news recently.

In 2023, Bloomberg[14] estimated that generative AI will become a USD $1.3 trillion market by 2023 growing at CAGR of 42% over the next 10 years reaching 12% of total technology spending.

Figure 6. Generative AI revenue



Source: Bloomberg, 2023

Bloomberg[14] estimates that the largest drivers of revenue from generative AI on the software side will be its infrastructure as a service ($247 billion by 2023) that is used for LLM training, digital ads powered by the technology ($192 billion) and generative AI assistant software ($89 billion).

The largest drivers of revenue on the hardware part are estimated to be AI servers ($123 billion), storage ($93 billion), computer vision products ($61 billion) and conventional AI devices ($108 billion).

Bloomberg[14] devotes special place in generative AI market growth to life sciences and education with spending increase at 113% CAGR in generative AI drug

discovery software and at 53% CAGR in generative AI in education spending over the years 2022-2032.

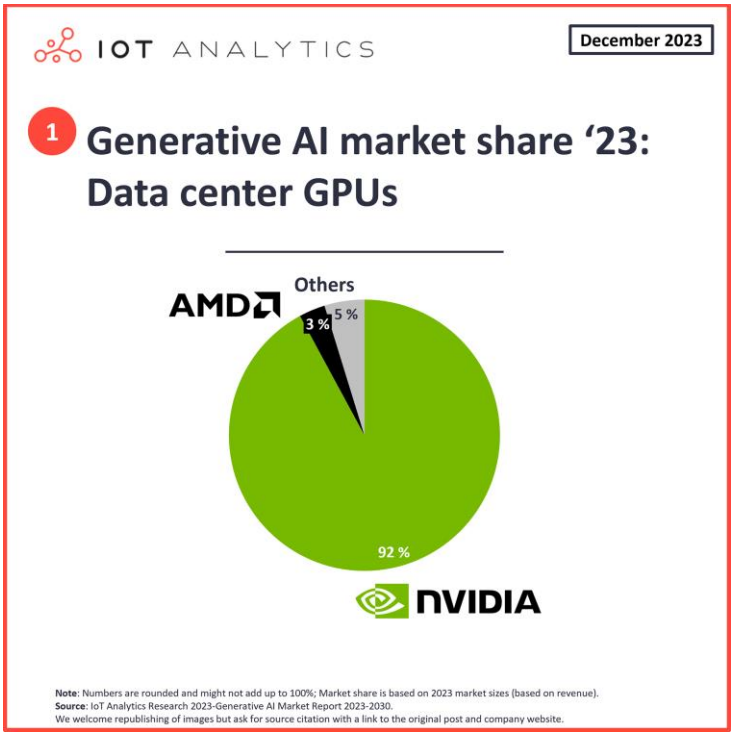
In 2023, North America dominated generative AI market with a 49.78% of market share[15].

2.3 Generative AI market and its three components: data-center GPUs, models and platforms, services

IoT analytics 2023 generative AI market report[16] provides a comprehensive overview of the generative AI market, suggests that year 2023 was when the market was established (the market went from nothing to hot, as formulated in the report). The generative AI market consists of three interconnected markets: 1) data center GPUs, 2) foundation models and platforms, and 3) generative AI services. NVIDIA is an absolute leader of the GPU segment with 92% market share (and the world's most valuable company with market capitalization of $3.34 trillion, $135.58 price per share, Yahoo finance, 2024[17]), with OpenAI and Microsoft covering 69% combined share of foundation models and platforms segment and Accenture being the leader of services segment with 6% market share.

2.3.1 Data center GPUs market

Figure 7. Data center GPU market



Source: IoT analytics, 2023

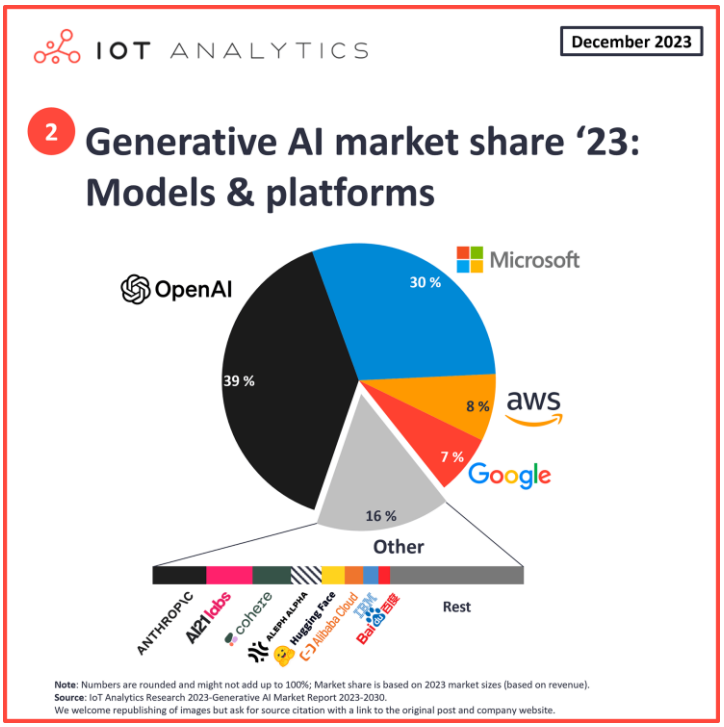
Data center GPUs refer to GPUs that are specifically designed to address high computational needs of AI data centers and are the corner stone of generative AI. GPUs that were originally designed for graphics processing excel at parallel computing which is a fundamental process for deep learning in generative AI.

Data center GPUs market has reached $49 billion in 2023 increasing by 182% from year 2022 mostly driven by NVIDIA alone.

NVIDIA is the absolute leader at the market with a share of 92%. Its quarterly revenue jumped 272% in year 2023.

2.3.2 Generative AI foundation models and platforms

Figure 8. Generative AI models and platforms market



Source: IoT analytics, 2023

The foundation models and platforms are two interconnected areas with foundation models being large pre-trained models that can be adopted to specific tasks without the need of training. Common examples of tasks are language processing, image recognition and decision-making.

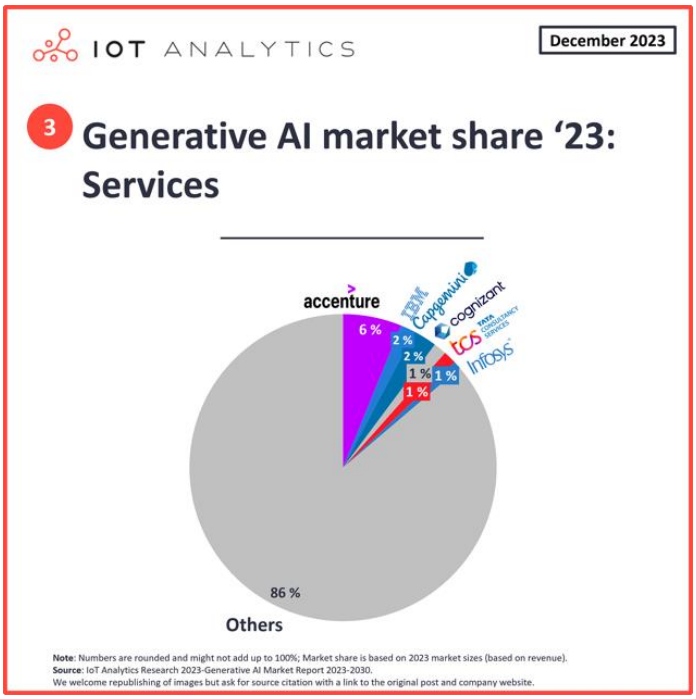
Platforms involve generative AI activities outside of foundation models. IoT analytics has identified six platform types: 1) development, 2) data management & databases, 3) AI IaaS & GPU as-a-service, 3) middleware & integration, 5) MLOps, and 6) UI & UX.

Foundation models market has exploded in 2023 with the release of Chat GPT reaching $3 billion in 2023 and IoT Analytics expects the market to continue to grow as the companies invest billions and report real value from generative AI adoption.

The leading foundation models and platforms companies are OpenAI and Microsoft with shares of 39% and 30% respectively. OpenAI became the market leader with the launch and success of Chat GPT in 2022 while Microsoft, OpenAI's largest shareholder, offers Azure OpenAI at its platform Azure AI, which uses OpenAI's LLM but promises greater data security and custom API if compared to Chat GPT.

2.3.3 Generative AI services market

Figure 9. Generative AI services market



Source: IoT analytics, 2023

Generative AI services market comprises of companies providing services of generative AI consulting, integration and implementation support to organizations

undertaking the implementation of generative AI capabilities. IoT Analytics stresses the high time for entering opportunity into this market due to novelty of the technology as well as lack of expertise in the organizations with that in turn requiring such companies to seek professional services.

Market leader of generative AI services is Accenture, investing $3 billion in data and AI practices, with the announcement made in June 2023 with the goal of doubling its AI talent and improving capabilities thereof. Later in November 2023, Accenture announced its launch of generative AI studios network in North America to provide companies a platform where they could experiment with generative AI.

2.4 Generative AI market concentration analysis

In order to estimate the concentration, the author used HHI, formula of which is defined as follows, where s_n is market share of firm n expressed as percentage[18].

$$HHI = s_1^2 + s_2^2 + s_3^2 \ldots s_n^2$$

HHI takes into account relative size and number of firms in a market. The index approaches zero when the market is occupied with a large number of firms of relatively equal size and reaches a maximum of 10,000 points when it is controlled by a single firm[19].

Markets with HHI between 1,000 and 1,800 points are generally considered to be moderately concentrated and HHI above 1,800 to be highly concentrated[19].

Based on the data from IoT analytics, using available share estimates of firms named explicitly and ignoring the category of others as it cannot provide the exact share estimates of firm separately as required for HHI calculation, generative AI market can be evaluated to be around 8473 HHI points for data center GPUs segment, 2534 HHI

points for models and platforms segment and at most 43 HHI points for generative AI services market. This calculation is based on the assumption that firms with market share <1% may be included in the category of others for the three generative AI segments decreasing their respective HHI point estimates.

It may be concluded from HHI calculation, that both data center GPUs as well as models and platforms segments are highly concentrated, while concentration of generative AI services is low. Composition of data center GPUs as well as of models and platforms segments of the generative AI market can be expected to remain relatively unchanged, while if it concerns generative AI services segment of the generative AI market, we can expect new entrants.

2.5 Other studies related to AI

There is a number of reputable institutions and research centers that are tracking impact of AI on societies and weren't mentioned earlier in this research, such as Stanford Center for Research on Foundation Models (CRFM) with its publicly available database of foundation models; Stanford Human-Centered Artificial Intelligence Index with its annual AI Index Report covering AI research & development, technical performance, responsible use, economy, science and medicine, education, policy and governance, diversity and public opinion; Georgetown University Center for Security and Emerging Technology with its publicly available databases tracking patenting, publishing, employment, research, law, policy and government activities in various countries, as well supply chain of semiconductors; Epoch AI tracking notable as well as large-scale AI models on training compute, parameters, data-set sizes, compute cost, time to train, as well as trends in compute, data, hardware, algorithms, investments, biology and also publish research on economics of AI automation. Most of the data bases mentioned provide online visualization tools offering insights without the need of own analysis.

CHAPTER 3. DATA AND METHODOLOGY

With the view to the goals laid down in the introductory part of the thesis, the author has identified the data source suitable for further analysis, UpWork 2024, a freelancing platform based in the US.

Search for sources of data was limited to the US as it is the most active and well-defined AI market as of 2024. Data search, collection, processing, analysis and visualizations, as presented below and in the following chapter of the thesis, were performed by the author individually.

In order to analyze the current state as well predictability of core input for AI development, developers' labor and services, listings of the AI developers were chosen for further analysis. Such collection of listings was further complemented with listings of AI developers' more familiar counterpart, software developers. This choice of data is not only useful for exploration of current state and predictability of the AI developers' labor and services, but is also useful in establishing benchmarks when such costs of labor and services are compared against their established and familiar counterparts, software developers, in the framework of using the same methods of statistical analysis for both data sets of listings.

While cost of AI developers' labor and services is not the only data of interest in the analysis of growth opportunities, such as the opportunity for good hire in the case of the emerging AI market, access to other data may be complicated due special permission requirements, such as requirement of inclusion into the Federal Research Program for access to US Census data, or simply high cost of data if considering commercial data providers.

With the view to the above mentioned, the author used Instant Data Scraper, 2024, to collect two samples of the AI and software developers' listings of services, that

of UpWork 2024 freelancers namely. The resulting two data sets have contained observations of 157 AI developers' listings and of 240 software developers' listings including variables of their hourly rates (USD), total amount earned on the platform (USD), amount of hourly jobs as well as fixed price jobs performed, and total hours worked.

The author has identified the key variable of interest, developer's hourly rate (USD), for further analysis of its predictability and stability, as such cost can be a critical part of costs for a company or investor actively pursuing the AI market, generative AI in particular.

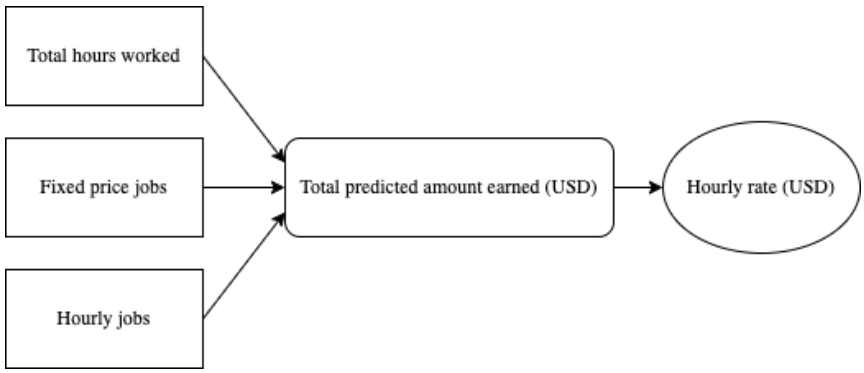
Preliminary analysis of hourly rates of AI and software developers have shown that they are close to being normally distributed making them suitable for further analysis using regression methods. Nevertheless, these variables contained observable amount of outliers in both sets of observations, requiring further processing for better prediction capabilities of the regression models that were applied in later analysis (Figures 11,12). All observations with 150 USD per hour and above for software developers and 120 USD per hour and above for AI developer have been removed from the data sets.

After initial data cleaning, exploration and processing specified in the previous steps, the data was further investigated in order to identify potential endogeneity and collinearity problems with the view to the plan of employing regression methods of statistical analysis. Correlation matrices have shown that there are observable levels of strictly positive correlations among most of the variables, while the variable of hourly rates is generally less correlated with the rest of the variables (Figures 13, 14). This finding suggested that method of two-stage least squares may be appropriate for further analysis with the view to the goal of benchmarking AI developers' hourly rates against such of its software counterpart.

Two-stage least squares method implies the employment of two subsequent least squares models in which the first stage least squares model identifies relevant instrumental variables based on the criteria of their respective p-values as well as F-statistic for the estimation of overall model fit. In case of establishment of the instruments' relevancy, the second least squares model may be employed controlling for the variable as predicted by its instruments. The use of predicted control variable allows the investigation of causal effect between the explanatory and response variables by removing endogeneity from the predicted control variable.

Based on the findings of the correlations observed in the two sets, as well as with the view to the need of comparison of the two sets of hourly rates for the purpose of establishment of their predictability and benchmarking, which in turn implies that the use of identical models is more desirable, investigation of the determinants of hourly rates begins with verification of total hours worked, amount of fixed price jobs and hourly jobs as instruments for total amount earned (USD). Upon the establishment of the instruments' relevance in the first stage least squares model, the predicted values of total amount earned (USD) may be used in the second stage model in order to investigate whether the developer's work experience help them charge higher hourly rate, or the relationship between the developer's experience is not so well defined, and the developers are charging their rates based on current demand in the market (Figure 10) or other. In other words, such models may describe flexibility of pricing at the job market and therefore may be useful for optimal hiring decisions.

Figure 10. Two stage least squares model for the analysis of the developers' hourly rates, overview diagram



The proposed two-stage least squares model can be described using the following equations.

First stage least squares model equation:

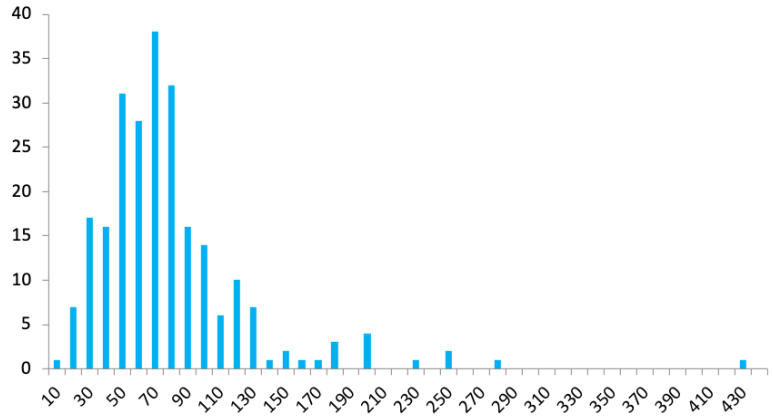
$$\widehat{total\ amount\ earned} = \beta_0 + \beta_1 * total\ hours\ worked + \beta_2 * fixed\ price\ jobs + \beta_3 * hourly\ jobs + \epsilon$$

Second stage least squares model equation:

$$\widehat{hourly\ rate} = \beta_0 + \beta_1 * \widehat{total\ amount\ earned} + \epsilon$$

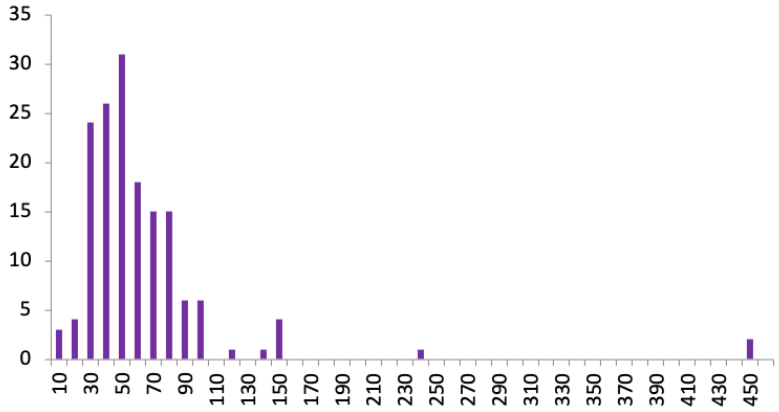
Sample of processed data used in the analysis can be referred to in Appendices.

Figure 11. Distribution of US based software developers' hourly rates, USD, before processing for further analysis



Source: UpWork 2024, own analysis and visualization

Figure 12. Distribution of US based AI developers' hourly rates, USD, before processing for further analysis

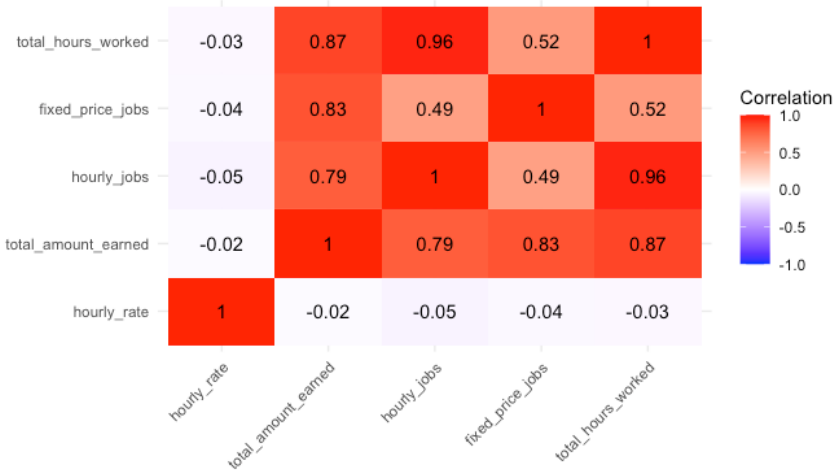


Source: UpWork 2024, own analysis and visualization

Figure 13. Correlation heatmap, software developers' listings

Source: UpWork 2024, own analysis and visualization

Figure 14. Correlation heatmap, AI developers' listings



Source: UpWork 2024, own analysis and visualization

CHAPTER 4. ANALYSIS AND RESULTS

Based on the author's analysis, the inclusion of all instruments into the first stage model for the prediction of total amount earned (USD) by software and AI developers on UpWork platform in order to try to predict the developers' hourly rates (USD), has been proven to be a reasonable approach as to the choice of statistical modeling of the data.

The author offers a valuable insight that may be draw from the first stage least squares models. Given the overall first stage model fit and significance of the coefficients in both cases of software and AI developers' data, effective market rates may be observed in models' outputs, namely the coefficients of the total hours worked, which may be interpreted as effective market hourly rates.

In case of software developers, the first stage least squares model has shown good model fit with adjusted R-squared of 0,9 and F-statistic of 638,9 (Table 1). Nevertheless, not all instruments used have been proven to be relevant to the response variable in question (hourly jobs at p-value = 0,67). Therefore, the author reduced the amount of the instrumental variables, using only fixed price jobs and total hours worked as candidates for instruments of the total amount earned variable. The latter approach yielded a better model fit, with adjusted R-squared value of 0,9, F-statistic of 961,9, with both fixed price jobs and total hours worked being relevant instruments to total amount worked as determined by their respective p-values < 0,05 (Table 2). The relationships between the relevant instruments and the total amount earned from the first stage least squares model, software developers' data, can be seen in Figures 15 and 16.

For AI developers, inclusion of all variables as candidates for instruments to the total amount earned variable (USD) in the first stage least squares model has shown even better results with all variables being statistically significant at p-value <0,05, that is constituting relevant instrumental variables, with adjusted R-squared 0,98 and F-statistic

2983 (Table 3). The relationships between the relevant instruments and the total amount earned from the first stage least squares model, AI developers' data, can be seen in Figures 17, 18 and 19.

After the choice and evaluation of relevance of the instruments, the author then created a variable of total predicted amount earned (USD) using first stage least squares for both software and AI developers sets. While the original variable of total amount earned wasn't correlated with the hourly rate variable, it was highly positively correlated with the rest of the variables, meaning that using all available variables for prediction of hourly rates wouldn't prove or disprove their causal effect on hourly rates (USD).

Second stage least squares model has shown that there is no statistically significant relationship between the total predicted amount earned (USD) and the desired hourly rates in both cases of software (Table 4) and AI (Table 5) developers. F-statistic for software developers' model is 2,8 and AI developers' model is 0,1, insignificant at 0,05 p-value threshold, suggesting poor predicting capabilities of the models.

On one hand, the results of the models may indicate that previous proven earnings of a developer, software or AI, do not determine the price of their future offers. On the other hand, if we consider the scatter plots of total predicted amount earned (USD) vs. hourly rate (USD) for both software (Figure 20) and AI (Figure 21) developers, we can see that most of the listings, whether cheap or expensive, are made by developers with low total amount earned as a result of their work on UpWork platform. It may suggest that new developers posting their offers on UpWork either do not know the real price of their work and choose their rates randomly or have other, reasonable, basis for their hourly rate calculations.

Table 1. First stage least squares model, software developers' listings

Dependent variable — total amount earned (USD)	
Hourly jobs	-109,23
Fixed price jobs	524,33 *
Total hours worked	49 ***
Adjusted R-squared	0,9
F-statistic	639 ***

Note: " " if p-value < 1, "." if p-value < 0,1, "*" if p-value < 0,05, "**" if p-value < 0,01, "***" if p-value < 0,001.

UpWork 2024, own analysis

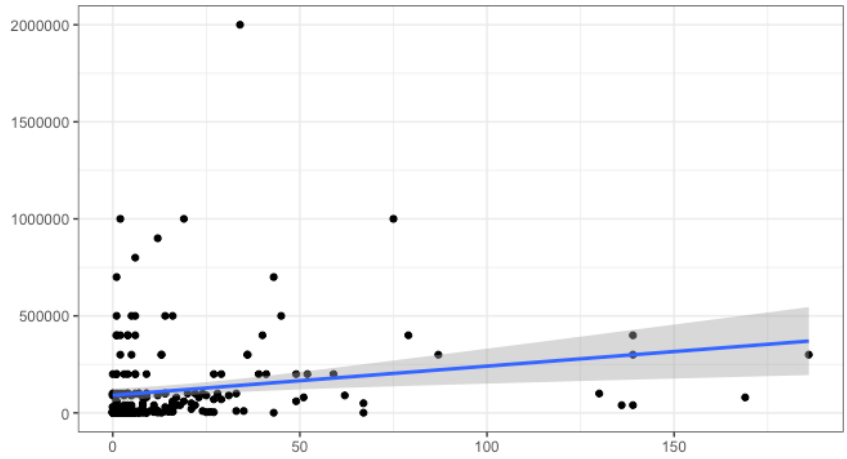
Table 2. First stage least squares model with irrelevant instrumental variable excluded, software developers' listings

Dependent variable — total amount earned (USD)	
Fixed price jobs	450,14 **
Total hours worked	48,57 ***
Adjusted R-squared	0,9
F-statistic	962 ***

Note: " " if p-value < 1, "." if p-value < 0,1, "*" if p-value < 0,05, "**" if p-value < 0,01, "***" if p-value < 0,001.

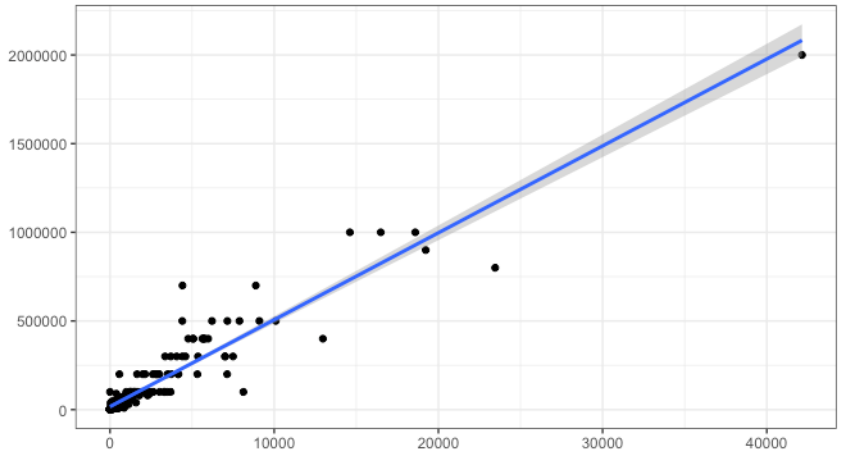
UpWork 2024, own analysis

Figure 15. Total amount earned, USD, vs. fixed priced jobs, software developers' listings



UpWork 2024, own analysis

Figure 16. Total amount earned, USD, vs. total hours worked, software developers' listings



UpWork 2024, own analysis

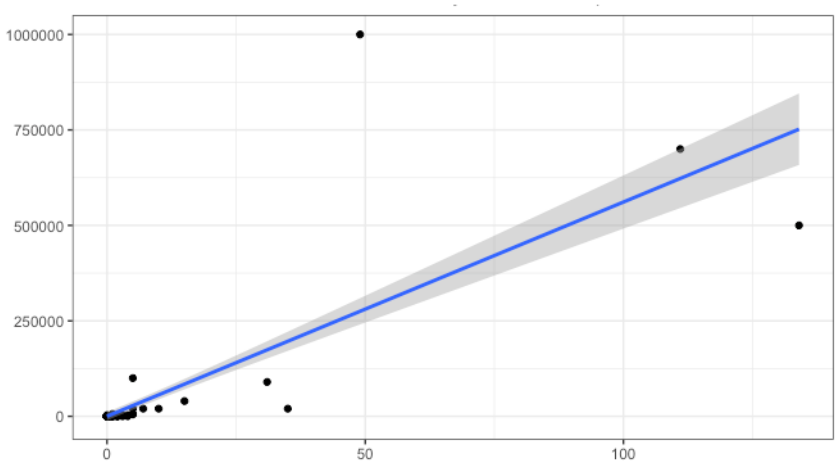
Table 3. First stage least squares model, AI developers' listings

Dependent variable — total amount earned (USD)	
Hourly jobs	-3990,89 ***
Fixed price jobs	4021,7 ***
Total hours worked	60,29 ***
Adjusted R-squared	0,98
F-statistic	2983 ***

Note: " " if p-value < 1, "." if p-value < 0,1, "*" if p-value < 0,05, "**" if p-value < 0,01, "***" if p-value < 0,001.

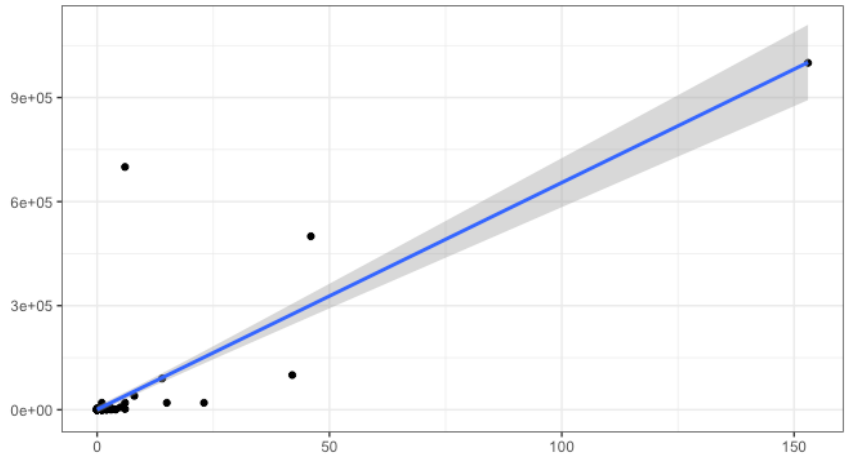
UpWork 2024, own analysis

Figure 17. Total amount earned, USD, vs. hourly jobs, AI developers' listings



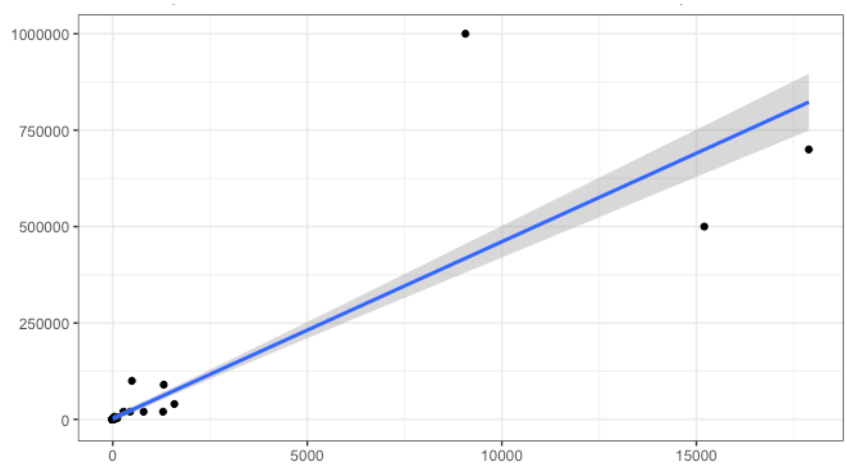
UpWork 2024, own analysis

Figure 18. Total amount earned, USD, vs. fixed price jobs, AI developers' listings



UpWork 2024, own analysis

Figure 19. Total amount earned, USD, vs. total hours worked, AI developers' listings



UpWork 2024, own analysis

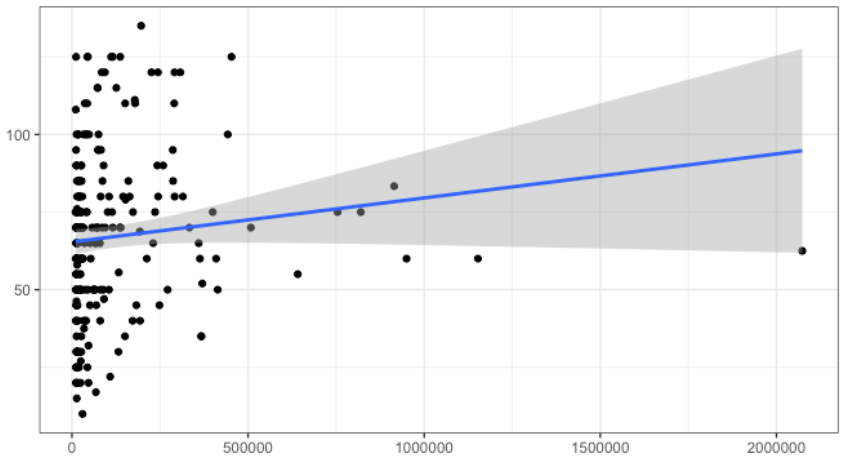
Table 4. Second stage least squares model, software developers' listings

Dependent variable — hourly rate (USD)	
Total predicted amount earned	0,0000014 .
Adjusted R-squared	0,008
F-statistic	2,8 .

Note: " " if p-value < 1, "." if p-value < 0,1, "*" if p-value < 0,05, "**" if p-value < 0,01, "***" if p-value < 0,001.

UpWork 2024, own analysis

Figure 20. Hourly rates, USD, vs. total predicted amount earned, software developers' listings



UpWork 2024, own analysis

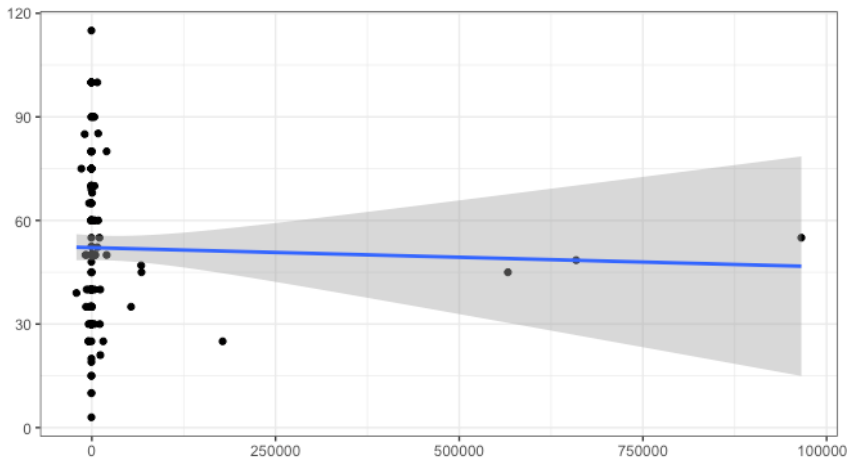
Table 5. Second stage least squares model, AI developers' listings

Dependent variable — hourly rate (USD)	
Total predicted amount earned	-0,0000056
Adjusted R-squared	-0,006
F-statistic	0,1

Note: " " if p-value < 1, "." if p-value < 0,1, "*" if p-value < 0,05, "**" if p-value < 0,01, "***" if p-value < 0,001.

UpWork 2024, own analysis

Figure 21. Hourly rates, USD, vs. total predicted amount earned, USD, AI developers' listings



UpWork 2024, own analysis

CHAPTER 5. CONCLUSIONS

5.1. AI market perspectives

Established in 2023, generative AI market went from nothing to hot with high growth expectations of CAGR 42% over the next 10 years. It is the most concentrated in its GPU segment with HHI score of 8472, less concentrated in models & platforms segment with HHI score of 2534, and not concentrated in generative AI services segment with HHI score of 43. This could indicate that across all three segments, there is room for business opportunities only in the services segment of the generative AI market and timing for market entry in now.

There are certain especially attractive segments of generative AI markets, such as generative AI drug discovery software which is expected to grow at 113% CAGR, and generative AI in education which is expected to grow at 53% CAGR. Certain business services, such as generative AI studios run by Accenture, are worth attention.

Since 2022, the number of new AI companies registering in the US has spiked, which can further signal high demand expectations for such products and services.

Larger firms, especially in the information sector, present higher rates of AI adoption across various spheres of application.

5.2. AI talent sourcing

As mentioned in introductory part of the thesis, AI talent is a critical resource for companies and investors pursuing the goal of entering or expanding into AI market.

Statistical modeling using two-stage least squares allowed to disprove a causal relationship between a proven success track record in terms of previous jobs and income for AI developers when their data is drawn from a single data source. This may either indicate flexibility of the developers to capture value arising from demand for their services at the time of such demand shifts or suggest that there could be other factors in the developers' background affecting hourly rates that are not easily measured or collected.

Nevertheless, even though there is no relationship between the developers' previous earnings and desirable current earnings in terms of hourly rates, it is important for both developers and their employers to know market situation in order to make optimal service provision and hiring decisions.

Author's analysis provides valuable insights into the structure of the developers' rates and earnings. As new talent enters the labor market, given that it may be difficult for inexperienced developers to reasonably rate their services, employers may use rates drawn from current market data in order to formulate their offers to the developers that are desirable for hire in terms of their qualifications but which do know a reasonable price for their services. Such approach may be useful for outsourcing and product companies, as well as for investors, for the purpose of evaluation of hiring effectiveness, which is critical to such businesses.

REFERENCES

1. "What is prompt engineering?" McKinsey & Co, published on March 22, 2024 https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-prompt-engineering

2. ChatGPT Statistics (AUG 2024) – Users Growth Data. DemandSage, published on August 10, 2024 https://www.demandsage.com/chatgpt-statistics/

3. OECD (2024) "Artificial intelligence, data and competition" - Background Note, presented on June 12-14 2024 https://one.oecd.org/document/DAF/COMP(2024)2/en/pdf

4. "Foundation Models Powering Generative AI: The Fundamentals", S&P Global, published on November 29, 2023 https://www.spglobal.com/en/research-insights/special-reports/foundation-models-powering-generative-ai-the-fundamentals

5. Bommasani, Rishi & Hudson, Drew & Adeli, Ehsan & Altman, Russ & Arora, Simran & Arx, Sydney & Bernstein, Michael & Bohg, Jeannette & Bosselut, Antoine & Brunskill, Emma & Brynjolfsson, Erik & Buch, Shyamal & Card, Dallas & Castellon, Rodrigo & Chatterji, Niladri & Chen, Annie & Creel, Kathleen & Davis, Jared & Demszky, Dora & Liang, Percy. (2021). On the Opportunities and Risks of Foundation Models.

6. Bresnahan, Timothy & Trajtenberg, Manuel. (1995). General Purpose Technologies 'Engines of Growth'?. Journal of Econometrics. 65. 83-108. 10.1016/0304-4076(94)01598-T.

7. "What are Generative AI models?" YouTube, uploaded by IBM Technology, March 22, 2023 https://www.youtube.com/watch?v=hfIUstzHs9A

8. "What is generative AI?" McKinsey & Co, published on April 2, 2024 https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai

9. "How to invest in AI's next phase." J.P. Morgan, published on May 23, 2024 https://www.jpmorgan.com/insights/investing/investment-trends/how-to-invest-in-ais-next-phase

10. Kathryn Bonney et. al "Tracking Firm Use of AI in Real Time: A Snapshot from the Business Trends and Outlook Survey", U.S. Census Bureau, published in March 2024 https://www.census.gov/hfp/btos/downloads/CES-WP-24-16.pdf

11. Kristina McElheran et. al "AI Adoption in America: Who, What, and Where", published on January 15, 2021 https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2022/03/McElheran-Kristina-et-al._AI-Adoption-in-America.pdf

12. Emin Dinlersoz et. al "Starting Up AI", U.S. Census Burearu, published in March 2024 https://www2.census.gov/library/working-papers/2024/adrm/ces/CES-WP-24-09.pdf

13. "The Possibilities of AI [Entire Talk] - Sam Altman (OpenAI)" YouTube, uploaded by Stanford eCorner, May 1, 2024 https://youtu.be/GLKoDkbS1Cg?si=Co1dT5BpM7Zkw6Xq

14. "Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds", Bloomberg, published on June 1, 2023 https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

15. Generative AI Market Size, Share & Industry Analysis, By Model (Generative Adversarial Networks or GANs and Transformer-based Models), By Industry vs Application, and Regional Forecast, 2024 – 2032, published on August 12, 2024 https://www.fortunebusinessinsights.com/generative-ai-market-107837

16. "The leading generative AI companies, IoT Analytics", published on December 14, 2023 https://iot-analytics.com/leading-generative-ai-companies/

17. "Trillion-dollar companies: 6 most valuable tech giants", Yahoo!finance, published on August 20, 2024 https://finance.yahoo.com/news/trillion-dollar-companies-222307746.html

18. "Herfindahl-Hirschman Index (HHI)", Corporate Finance Institute, accessed on August 20, 2024 https://corporatefinanceinstitute.com/resources/valuation/herfindahl-hirschman-index-hhi/#:~:text=The%20Herfindahl%2DHirschman%20Index%20is%20used%20to%20monitor%20the%20potential,to%20veto%20an%20M%26A%20transaction.

19. "Herfindahl-Hirschman Index", U.S. Department of Justice, Antitrust Division, updated on January 17, 2024 https://www.justice.gov/atr/herfindahl-hirschman-index#:~:text=The%20HHI%20is%20calculated%20by,%2B%20202%20%3D%202%2C600).

APPENDIX 1

Sample of processed UpWork listings of software developers' proposals, 2024

hourly_rate	total_amount_earned	hourly_jobs	fixed_price_jobs	total_hours_worked
65,00	60000,00	14,00	1,00	1152,00
50,00	5000,00	6,00	26,00	33,00
180,00	300000,00	105,00	32,00	2770,00
200,00	10000,00	12,00	65,00	24,00
35,00	300000,00	72,00	36,00	7007,00
55,56	80000,00	219,00	169,00	953,00
75,00	1000000,00	59,00	19,00	16489,00
65,00	200000,00	39,00	39,00	4165,00
250,00	200000,00	30,00	30,00	720,00

APPENDIX 2

Sample of processed UpWork listings of AI developers' proposals, 2024

hourly_rate	total_amount_earned	hourly_jobs	fixed_price_jobs	total_hours_worked
75,00	0,00	0,00	0,00	0,00
30,00	50,00	0,00	1,00	0,00
45,00	100,00	1,00	1,00	5,00
50,00	0,00	0,00	1,00	0,00
35,00	300,00	3,00	1,00	9,00
35,00	0,00	0,00	0,00	0,00
60,00	0,00	0,00	0,00	0,00
50,00	0,00	2,00	0,00	0,00
55,00	1000000,00	49,00	153,00	9064,00