# M&A and Big Tech's Competitive Advantage: A patent approach to Big Tech's evolving technological profile

Ilan Strauss and Jangho Yang *

March 2025

**Abstract**

Using a unique dataset covering all patents ultimately owned by Big Tech, including through M&A, we describe the dynamic capabilities acquired and developed by Big Tech. We examine the nature, evolution, and differences in the technological capabilities of the five major Big Tech companies, highlighting the vital role of M&A in this process. Our analysis, combining M&A and patent data, shows that M&A has been crucial in Big Tech developing integrated hardware-software ecosystems. Big Tech's evolving capabilities closely track their competitive potential and market entry strategies. Our analysis of patent data, including through a logit regression, reveals diverse strategies and motivations behind Big Tech's external patent acquisitions.

**Keywords**: Patents, capabilities approach, Big Tech, M&A, dynamic competition.
**JEL Codes**: K21, L41, L44, M13, E14.

---

*Equal Contributors (alphabet author order). Ilan Strauss is Co-Director of the AI Disclosures Project at the Social Science Research Council (300 Cadman Plz W, Brooklyn, NY 11201), and Honorary Senior Fellow at the Institute for Innovation and Public Purpose (IIPP), University College London (UCL), London, UK, WC1E 6BT. Jangho Yang is Assistant Professor of Management Sciences in the Faculty of Engineering, University of Waterloo. We would like to acknowledge grateful funding from The Economic Security Project (ESP) for this project grant and Becky Chao at ESP for grant assistance. We would like to express our deep gratitude to Nigel Swycher for access to Cipher's patent dataset (acquired by LexisNexis) and to Patrick Curry and others at Cipher for working through tedious dataset construction together. A big thank you to our two incredible research assistants at University of Waterloo: Elisha Matharu for data and programming, and Fuyang Deng for graphing and case studies. Thank you to Tim O'Reilly for edits and comments. Thank you to those who provided helpful comments at two workshops: at the UCL Centre for Law, Economics and Society on 10th October 2023, and at the UK CMA on 12th October, 2023, where a more antitrust focused version of this paper was presented. All errors are our own. Contacts: i.strauss@ucl.ac.uk (Corresponding Author) and j634yang@uwaterloo.

# Contents

# 1 Introduction

There is a growing sense that existing antitrust, focused on regulating competition as a static struggle for existing market share within a stable industry (Evans et al. 2008), has allowed for many merger and acquisitions (M&A) to go through that were anti-competitive with hindsight. This has allowed Big Tech to acquire hundreds of smaller, highly innovative companies, each with minimal existing market share, yet instrumental in helping Big Tech create diversified product ecosystems (Doctorow 2023). Our dataset shows at least 995 majority acquisitions from Apple, Amazon, Alphabet, Meta, and Microsoft between 2000 and 2022. The importance of this is reinforced by the technological shift underway towards artificial intelligence and generative language models, which can dethrone incumbents and foster competition, or reinforce Big Tech's dominance, especially if existing patterns of acquisitions are permitted to continue (Sharma 2023).

But measuring "innovation and invention", and defining appropriate market boundaries have led to considerable analytic confusion (UK Competition Appeal Tribunal 2022). Hovenkamp (2008, p. 3) goes so far as to say that: "the consequences of innovation are often radically indeterminate."

In this paper, we try to take an initial step towards measuring innovative potential and dynamic capabilities by using Big Tech's *patent* holdings, focusing on the role of M&A in acquiring them. Our analysis and goals are largely descriptive: using visualizations and individual case studies we show the value of our unique patent data and its ability to provide a potential competitive analysis. Taken together, we find that Big Tech's acquired capabilities from M&A patents are essential technologies that have allowed them to enter new product markets, adapt to constantly changing market conditions (D. J. Teece et al. 1997), and to erect deeper ecosystem moats reliant on complimentary assets (Petit and D. Teece 2020). Yet when evaluated individually, the technologies are highly uncertain and often fail.

In the dynamic capabilities literature which we draw on, patents are not just legal protections – they are an integral part of a firm's broader strategy to secure the returns from its innovative capabilities and manage complementary assets (D. J. Teece 1986; D. J. Teece et al. 1997). Patents, as technologies, are a natural fit for focusing on competitive potentialities, yet still relevant to actual product markets: we find a strong connection between Big Tech's (patent) capabilities concentration and entry in the market for goods and services.

Yet two considerable challenges exist to using patents for a competitive analysis, that we try and overcome. Firstly, existing patent technology categories suffer from several well known shortcomings (Krestel et al. 2021), leading even the USPTO to increasingly use machine learning (ML) to classify patents into relevant technology categories (A. Toole et al. n.d.; Giczy et al. 2022). Secondly, tracking changes in patent ownership is cumbersome, but necessary in order to gain a true picture of patent ownership by a company, especially when the company in question has undertaken considerable M&A activity. This is because in the U.S. especially, recording an assignment (change of ownership) with the patent office is not required.

Our primary contribution is to correct for the above two fairly severe defects in patent data by constructing a comprehensive dataset of all patents owned by Big Tech, including those it acquired through M&A. We also use machine-learning determined technology categories. We do both of these tasks in partnership with Cipher, now acquired by LexisNexis (Cipher Platform). Our key specific findings are that:

1. Quantitatively, M&A has been central to Big Tech developing its capabilities. At least 10.3%, 13.1%, and 10.8% of total patent counts on an unweighted, forward-citation-weighted, and forward-citation-weighted adjusted for patent-age basis, respectively, have been acquired externally by Big Tech through majority stake M&A activity (Figure 1).

2. Patent (forward) citations show that external acquisitions provide the Big Tech firms with access to technology that is often more productive than their own internally developed technology, judged by various measures of forward citations of M&A patents vs. internally developed patents (Figure 1 and Table 3). Individual, highly advanced, technologies (high median citations) in key firms acquired are especially evident in touch screens, voice recognition, advertising technologies, and other software capabilities.

3. Descriptively, there appears to be a tight link between diversification in capabilities (including through M&A) and diversification and entry into product markets, highlighting that a potential competitive analysis on the basis of technologies is reasonable in the aggregate, but often difficult in any individual case or technology. Many of the products entered into by Big Tech after 2010 relied on building out integrated hardware-software ecosystem.

4. Most acquisitions have highly uncertain competitive effects because the technologies acquired are uncertain (nascent). Capabilities in acquired firms tend to be extremely young (judged by low median patent counts), with few proven technologies (judged by few patents with citations).

This makes a quantitative patent analysis of M&A's impact on Big Tech's capabilities difficult on a case by case basis.

5. Big Tech firms show a diverse range of strategies in the capabilities that they have acquired through M&A. Regression findings indicate some Big Tech firms prefer to buy largely young and unproven patent assets (Microsoft), others prefer more established technologies (Amazon), and others engaging in a mix of the two (Meta and Apple). Alphabet makes the greatest use of M&A to develop a broad range of capabilities across technological categories, resulting in the most integrated and diverse suite of ecosystem products.

A notable limitation of our study is that while patents can help illuminate a firm's dynamic capabilities (D. J. Teece et al. 1997), they can also serve as legal assets acquired for defensive purposes or risk mitigation, particularly in technology sectors with complex intellectual property landscapes (Graham et al. 2003). Such assets do not usually translate into real technology transfers. This is exemplified by cases like Google's acquisition of Motorola Mobility, where patent portfolios were strategically acquired to support existing platforms rather than operational capabilities (Cohen et al. 2019). We account for this distinction in our analysis by examining both citation patterns and technology integration trajectories. Additionally, the software and internet-services sector also relies on other forms of IP (copyright, secrecy, or first-mover network effects). This means that major patent acquisitions in our dataset do not always reflect cutting-edge capabilities but are instead often mature, defensive assets.

Our paper provides the first publicly available dataset on all patents owned by Big Tech, adjusting for changes in ownership and assignee status. Our data approach, linking Big Tech's M&A data to patent data, is similar to Gugler et al. (2023), but they focus on the time-series dimension and not on adjusting for changes in ownership.

The policy implications of our findings can be used to highlight the importance of reviewing the assets held by Big Tech and other companies in order to understand their true competitive capabilities, the markets they could (or intend to) enter, along with the potential antitrust implications of mergers and acquisitions into seemingly unrelated vertical markets, but where technological complementarities might in fact make it a natural fit for market entry.

Section 2 starts with a review of the relevant literature and our exploratory visualization method. Section 3 provides an overview of our data, exploring potential motivations for M&A through a patenting lens. Section 4 examines the core figures that explore the relationship between M&A and competition, focusing on the technological patents that have given Big Tech access to new capabilities. Section 5 presents a logit regression analysis, predicting the probability of a patent being externally acquired through M&A to assess whether capability acquisition strategies differ by Big Tech firm. Section 6 discusses the implications of our findings and concludes.

# 2 Literature Review

## 2.1 Dynamic Competition

Competition, especially in digital markets, increasingly centers around a firm's ability to innovate over time, capturing future rents (Cadman 2023). This perspective, often labeled dynamic competition, focuses on disruptive threats emerging from outside traditional market boundaries (Caffarra et al. 2023a,b). Recent scholarship has emphasized the importance of innovation-driven threats to incumbent firms, especially in technology-intensive industries (Sidak et al. 2009; Petit and D. J. Teece 2021). However, regulatory approaches differ across jurisdictions; notably, Europe and the U.S. maintain distinct frameworks for analyzing and addressing dynamic competition concerns (Gifford et al. 2011).

Dynamic competition has been defined by Sidak et al. (2009) as: "a style of competition that relies on innovation to produce new products and processes and concomitant price reductions of substantial magnitude. Such competition improves productivity, the availability of new goods and services and, more generally, consumer welfare." Likewise, Petit and D. J. Teece (2021) sees dynamic competition as: "a situation in which firms compete for future rents. In dynamic competition, firms use innovation to introduce new products, processes and services. Rivalry results in product differentiation, integration, diversification, or platformisation."

Both definitions highlight innovation as the engine driving sustained competitive advantage (Cadman 2023). Their roots trace back to D. J. Teece et al. (1997), who argues that ongoing innovation and the capacity to renew capabilities underpin a firm's competitive advantage. Posner (2006) further emphasizes the cumulative nature of innovation: "most intellectual property builds upon intellectual property," making incremental improvements an integral part of the innovation process. Within this dynamic framework, complementary assets — including patents — work together to produce competitively valuable products, services, and ecosystems (Sidak et al. 2009).

Following D. J. Teece et al. (1997) and D. J. Teece (2007), we interpret M&A actions (the acquisition of firms and their patents) as part of Big Tech's broader dynamic capabilities strategy, which includes: Sensing new opportunities or threats (e.g., nascent competitors) seizing these opportunities through M&A, licensing, or collaborations and reconfiguring acquired patents or integrating them into existing R&D portfolios (Cunningham et al. 2021; Gugler et al. 2023).

## 2.2 Patents as an Indicator of Capabilities

Emerging antitrust approaches are trying to incorporate a dynamic approach to harms from M&A premised on a firm's capabilities (Sidak et al. 2009; McSweeny et al. 2018; Petit and D. J. Teece 2021) – in the UK especially (UK Competition and Markets Authority 2021, 2022; Cadman 2023), but also the U.S. through the new merger guidelines (FTC and DoJ 2023). In this context, the competitive threat facing the monopolist is from new ("potential") product and technological markets in the future (Bryan et al. 2020; Areeda et al. 2023). Antitrust regulatory impact assessment, therefore, requires looking at more uncertain future events, including (non-price) future impacts on innovation and competition (Hovenkamp 2008). A merger may reduce competition through its impact on innovation – either killing the innovation, removing the innovative competitor, and/or removing access to the innovation for others (Areeda et al. 2023, Section 701).

Estimating a firm's capabilities empirically is challenging. R&D spending (D. J. Teece 2010) or 'acquihires', where the motivating factor is the underlying capability of the labour capital assets (Makinen et al. 2012; FTC 2021), is often used. But this might fail to provide a more granular picture of capabilities. Further, data on hires is often difficult to obtain without relying on platforms like CrunchBase or LinkedIn (Tunguz 2025). Another indicator is the number and quality of patents, which typically corresponds to a firm's R&D intensity, innovative capacity, and potential to create valuable new offerings (Levin et al. 1987; Pavitt 1982; Griliches 1998).

Building on these insights, we interpret patents as strategic assets used to secure future economic rents derived from disruptive or nascent innovations (Cohen et al. 2019; Argente et al. 2020). Such rents derive not simply from being bundles of legal rights but also as proxies for a firm's innovation capabilities – as Schumpeterian rents (D. J. Teece 2012; Morton et al. 2013; Gugler et al. 2023). They represent the codified output of R&D activities and signal a firm's intangible capacity to generate new knowledge or processes (Pavitt 1982; Levin et al. 1987; Griliches 1998; Akcigit et al. 2023). In dynamic competition environments, patents can bolster competitiveness in both product and service markets (Argente et al. 2020).

However, patent measures have well-known drawbacks: Firstly, patent filings occur after the innovation process and take time to be granted or cited (Kim et al. 2016), potentially under-representing immediate yet nascent competitive threats, which have not been granted a patent. Secondly, not all innovations — particularly in software — are easy to patent or fall under uniform legal standards (Graham et al. 2003; Saltiel 2019). Thirdly, some patents are obtained merely to deter litigation or block competitors, rather than to commercialize the underlying invention. Fourthly, alternative IP protections (e.g., trade secrets, copyrights) may dominate over patenting, especially within the software sector. Lastly, patents also reflect inherent future market uncertainty: inventions that might never be commercialized successfully (Griliches 1998). In this regard their innovative capacity is always uncertain.

Given this, our empirical approach to assess how M&A has contributed to Big Tech's evolving technological capabilities is three-fold: First, we operationalize a dynamic capabilities approach (D. J. Teece et al. 1997) through visualizing, in aggregate, across a large number of acquired patents, how patents as technologies integrate with other technologies held by Big Tech companies. Second, we use visual analysis to examine how patent technologies acquired by Big Tech integrate with their existing ones, to assess broader motivations across the hundreds of acquisitions made. Third, we use case studies to explore specific motivations for each Big Tech acquirer, or firm acquired, again using our patent data. Finally, using the patent data, we employ firm-level econometric analysis to confirm this differentiated approach to acquisitions by the five traditional Big Tech firms.

## 3 Methodology

In this study, we set out to explore the innovative potential and dynamic capabilities of major technology firms by examining their patent holdings in the context of mergers and acquisitions (M&A). Our primary goal is to introduce a new dataset and provide detailed observations of how five leading companies—Google, Apple, Facebook (Meta), Amazon, and Microsoft—acquire patents through M&A activities. Our analysis is primarily descriptive: we use data visualizations and focused case studies to demonstrate the dataset's usefulness and uncover key patterns in these firms' patent-based strategies.

Our emphasis on a limited but influential group of firms is intentional. Because these companies play a central role in today's innovation ecosystem, we believe it is critical to examine their approaches to patent acquisition in detail. Expanding our scope to include more firms and conducting comprehensive statistical analyses could yield broader insights. However, this paper focuses on an in-depth analysis of how these five Big Tech firms use M&A strategies to enhance their dynamic capabilities. By focusing on specific case studies—such as Google's acquisitions of Keyhole, ZipDash, and Where 2 Technologies—we can more effectively highlight the context and timeline of corporate decisions that have shaped dominant positions in areas like mapping technology.

Furthermore, the nature of these acquisitions often calls for close attention to historical and organizational factors that may not be immediately evident in aggregated data. For instance, Google's "Pac-Man" style acquisitions for building its Maps platform involved a sequence of smaller deals that, when viewed cumulatively, illustrate how the firm's patent portfolio and technological competencies evolved over time. This interconnection between patents and M&A deals underscores why a descriptive method and case-based perspective are essential for understanding the complexity of Big Tech's innovation strategies.

Overall, our descriptive approach aims to reveal the multifaceted relationships between patents, M&A, and the development of dynamic capabilities. By mapping out individual patent portfolios alongside corporate acquisition timelines, we seek to offer a rich, context-driven account of how these major technology firms build, leverage, and transform their innovation resources.

### 3.1 Data

Our key dataset is at the patent level – showing all patents owned by Big Tech – and combines two datasets: a dataset of firms acquired by Big Tech (Refinitiv, Wikipedia, Web Scraping) and a dataset on all patents owned by Big Tech (Cipher now owned by LexisNexis). We collaborated with Cipher to improve the linking of organizations to patents. This is crucial for tracking changes in patent ownership ("assignee"), which can change when companies are bought and sold, but might instead remain the same. This makes linking the patent to the new owner difficult. Our patent data is structured hierarchically, with an ultimate organization linked to a patent owner, the legal assignee of the patent, an original historical owner, and assignee history.

We merge & match patent data with M&A data (by Big Tech company). Matching is done based on of the following variables: the patent's ownership, its assignee, or its historical ownership. 10%

of patents (13,196 patents) match to external assignees/organizations/historical owners; or 227 firms from our M&A dataset are present in our patent dataset (22% match). This underestimates the true extent of externally acquired patents since internally developed patents might rely on them.

**M&A Data, 2000 - 2022.** The M&A dataset is a firm-level merger and acquisitions dataset of 995 majority-stake acquisitions undertaken by the five Big Tech companies from Refinitiv, Wikipedia, and AI webcrawlers, all of which have been verified by humans (with corresponding verification source provided). See Table 1 for the number of acquired firms by each Big Techs. The dataset ranges from 2000 to 2022 for the year completed of the M&A deal. Share repurchases, minority stakes, and property acquisitions are excluded. Acquisitions of subsidiaries are included, for example Google buying Motorola Mobility. Software is the main 'mid' level category of acquisition, with "Internet Software & Services" being a close second for Amazon. The 'macro' industry classification of the acquired firms given by Refinitiv are almost entirely in High Technology.

FTC data shows that most M&A transactions by Big Tech are below the size thresholds required in the Hart-Scott-Rodino (HSR) Act. Around half of the acquisitions not reported to the FTC between 2010-2019 were firms younger than five years old (FTC 2021). Big Tech reported 819 non-HSR reportable transactions over the 10-year period 2010-2019 alone (ibid.). But this includes 87 minority stakes, non-corporate interests, and non-acquisition licensing agreement that we try to exclude from our dataset.

225 firms from our M&A dataset are present in our patent dataset, or a 22.6% match percentage indicating a high share of recognized technology in the acquired companies. Apple and Microsoft have the highest capabilities motivation in their acquisitions, with 26.2% and 28.3% respectively of their acquired firms having matched patents. Alphabet has the lowest at 17.6%.

**Patent Dataset, 2000 - 2022.**The second dataset is a patent-level dataset from Cipher that contains all patent families owned by the five Big Tech companies, totaling 127,298 patent families. See Table 2 for the number of patents externally obtained by each Big Techs. Cipher, a private patent data company recently acquired by LexisNexis and integrated into their PatentSight service, uses publicly available patent data and applies machine learning classifiers to identify useful technologies. The patent data is at the patent family level - this is global and avoids multiple similar patents filed in different jurisdictions removing considerable redundancies. We can filter patents by geography based on the granting jurisdiction. It includes expired, inactive, and pending patents, but excludes design

Table 1. Firms Acquired by Big Tech, 2000 - 2022

| Acquirer | No. of M&A | M&A w. Patents |
|---|---|---|
| Alphabet | 341 | 60 |
| Amazon | 144 | 35 |
| Apple | 122 | 32 |
| Meta | 116 | 21 |
| Microsoft | 272 | 77 |
| All | 995 | 225 |

Note: Based on 995 firms acquired by Big Tech in Refinitiv database, Wikipedia, and AI Crawlers between 2000-2022. Cleaned to include only majority stake, removing share repurchases, real estate deals, and deals frozen or rejected by regulatory authorities (e.g. Meta/Giphy). We include Microsoft's 2009 deal with Yahoo! (amended 2015) since at the time it provided Microsoft with an exclusive 10-year license to Yahoo!'s core search technologies. Final column (M&A w. Patents) showing number of acquired firms in sample which contains patents in the merged datasets.

patents as we want to capture previously acquired patents from past acquisitions - even if expired. All citation data is as of the present (December 2022).

Table 2. Merged Patent-level Dataset, 2000 - 2022

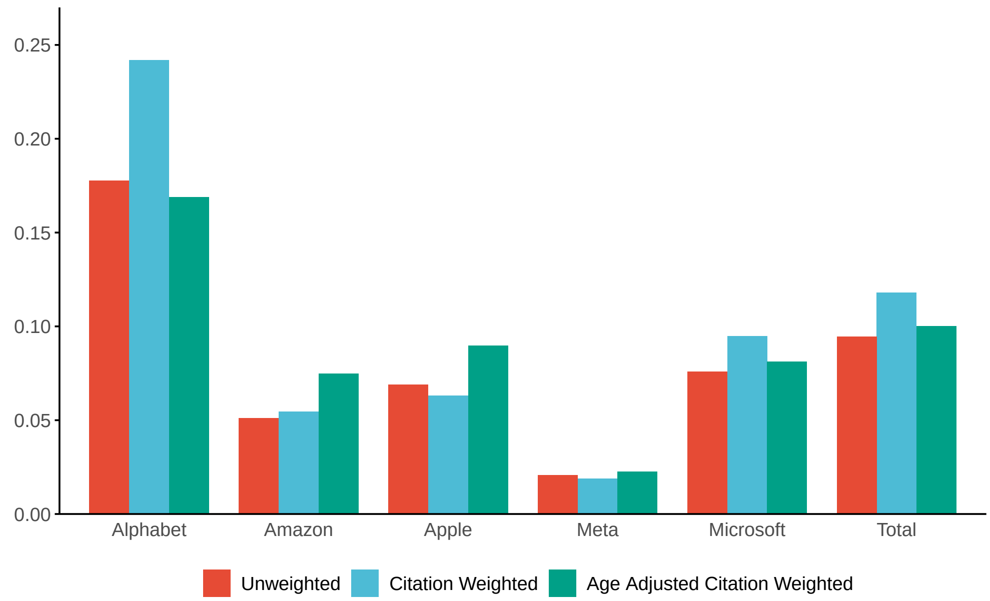
| Big Tech | Total Patents | M&A Patents | M&A Patent (%) |
|---|---|---|---|
| Alphabet | 32,855 | 5,835 | 17.7 |
| Amazon | 15,175 | 775 | 5.1 |
| Apple | 27,489 | 1,896 | 6.5 |
| Meta | 7,582 | 161 | 2.1 |
| Microsoft | 44,197 | 3,417 | 7.7 |
| Total | 127,226 | 12,084 | 9.5 |

Note: Patents acquired by Big Tech. Removing a jointly owned patent between Meta and Microsoft.

Figure 1 shows that at least 10.3%, 13.1%, and 10.8% of patents unweighted, citation weighted, and citation weighted adjusted for age, respectively, are acquired externally by Big Tech. Note that Alphabet has the highest share of patents arising from M&A, despite having the lowest capabilities motivation in its acquisitions (17.6% in Table 1).

This is certainly a lower bound given that internally developed patents may depend significantly on externally acquired ones (including through complementary human capital). In addition, our matching

Figure 1. Share of Patents from M&A: Unweighted vs Citation Weighted




Note: 10.3%, 13.1%, and 10.8% of patents unweighted, citation weighted, and citation weighted adjusted for age, respectively, are acquired externally by Big Tech. "Weighted" means that we multiply patents by forward citations. age-adjusted involves dividing this measure by the age of the patent, defined by the year it was granted.

algorithm between M&A and patent data may be imperfect.[1] Finally, qualitatively, many acquired technologies are not always patented, especially when in their infancy or involving software.

Our patent dataset contains two key variables of interest:

- **Forward citations**: Indicating how many times the patent has been cited by other patents (Hall et al. 2005).
- **ML Technology category**: The ML technology classifier is an unsupervised learning based multi-class classifier. The ML technology classifier is a Universal Technology Taxonomy ("UTT"), meaning it classifies the entire patented world through a common lens, based on 10 major technological categories and 122 subcategories (LexisNexis 2023). This contrasts with the 300,000 CPC codes, which are used for classifying prior art when examining a patent application. CPC codes are less useful for classifying existing patenting technology in relation to one another since they are provided without consideration for how the patent is used or how the technology evolves relative to others.

[1] False positives are not an issue, as we manually validated all M&A cases, but we cannot rule out false negatives, which may lead to missing actual M&A patents.

### 3.2 Motivations for Acquisitions

Based on the broad industry classification in Refinitiv of the firms acquired (high-technology), acquiring advanced capabilities (including patents and human capital, but also data) is clearly a key motivation. Capabilities are often in their infancy when acquired, however, complicating any quantitative analysis. Table 3 shows that median patents held among all the companies acquired are as low as 2 (Microsoft, Alphabet, and Amazon), and with Apple having the highest median at 4 (see Table 3, column four). In other words, the companies which Big Tech are acquiring tend to be often too young to hold any recognized, registered, technological assets.

Table 3. How Innovative are the Acquired Companies (by Citations & Patents)?

| Big Tech | Median Cite (Non-M&A) | Median Cite (M&A) | Median Patent Total | Max Patent Total | Acquired Firm with Max Patents |
|---|---|---|---|---|---|
| Microsoft | 15 | 20 | 2 | 1882 | Nuance Communications |
| Alphabet | 9 | 21 | 2 | 4992 | Motorola Mobility |
| Apple | 9 | 12 | 4 | 1582 | Intel (Modem chip) |
| Amazon | 5 | 5 | 2 | 547 | Zoox |
| Meta | 5 | 4 | 2 | 79 | WhatsApp |

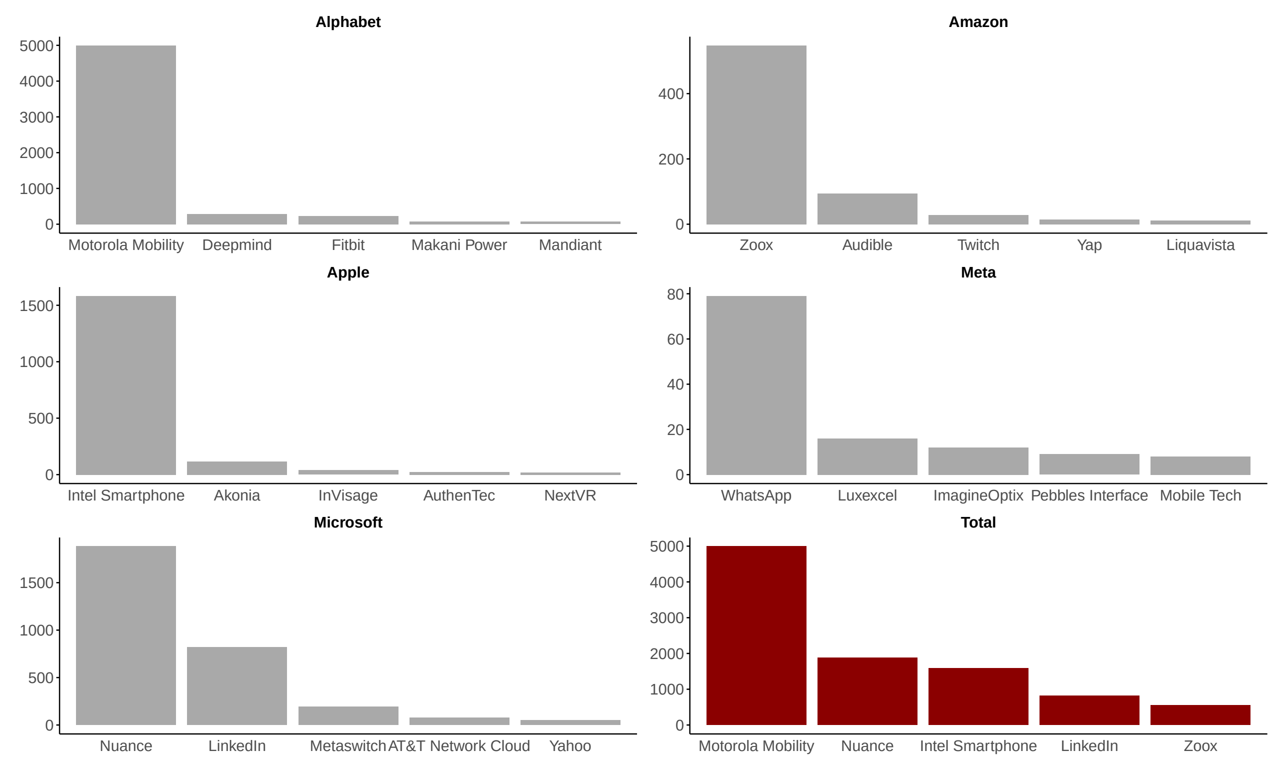
Note: Showing number of M&A firms acquired with patents; median forward citations among pooled acquired patents; median no. of patents acquired considering the total patents acquired in each firm; maximum patents acquired from any one firm; and the firm containing the most amount of patents acquired.

Table 3 highlights the extent to which some Big Tech firms rely on acquisitions for innovation. Median citations among acquired patents are twice as high for Alphabet (21 vs. 9) and one-third higher for both Microsoft (20 vs. 15) and Apple (12 vs. 9). For Amazon and Meta, the citations for internally developed and externally acquired patents are generally low and similar.

A key motivation for Big Tech's acquisitions has been building out new, more expansive, product ecosystems (Jacobides, Cennamo, et al. 2018; Jacobides and Lianos 2021). This reflects vertical (different or unrelated production stages or products) rather than horizontal (same market) acquisitions into complementary hardware, software, cloud, and emerging (sometimes speculative) technologies. This is evident in Figure 2, which shows the five largest acquisitions for each Big Tech company based on total patent count.

It highlights such vertical and technological deepening acquisitions. For Alphabet, this is about

Figure 2. Top Five Acquired Companies with Most Patents by Each Big Tech




Note: Top five companies, by total patents, for each Big Tech company. Including Yahoo! Search and Advertising technologies acquired by Microsoft since it was an exclusive technology license for 10 years (the terms of which were changed in 2015).

shifting from software to hardware and selling goods. For Amazon, acquisitions have supported pure diversification in capabilities in ways which it thinks supports its core retail business, such as Zoox (self-driving cars) and Audible (audiobooks and podcasts), or its web service, such as Twitch (gaming streaming). And this is highlighted at the patent level, as a proxy for capabilities. Integration of AI by Big Tech is evident in large acquisitions for Alphabet (DeepMind), Microsoft (Nuance Communications), Meta (Mobile Technologies) (TechCrunch 2013) and Amazon (Yap) (TechCrunch 2011). Apple's acquisitions often bolstered its iPhone (e.g., InVisage for improved cameras), while Amazon's AI acquisitions focused on voice recognition (Amazon's Echo).

Ecosystem diversification is not always evident in the patent data at the individual firm-level though (Doctorow 2023). Patterns often only emerge at the aggregate technological level, taking into account all firms acquired. It is widely known that Facebook acquired WhatsApp and Instagram, Google acquired DoubleClick and YouTube, and Apple acquired Beats. In fact, most of the beloved products of Big Tech grew out of serial acquisitions. Google's acquisition of Docverse in 2010 helped it build out its online collaborative suite of Office-like products (Google Docs, Google Slides, etc.). Google bought a 1.5-year-old startup called Android in 2005, to expand its core search and ads business beyond the PC platform (Callaham 2022). Today, Android is the most popular mobile operating system (OS) and provides Google, in conjunction with its suite of apps (its "Google Mobile Services"), with considerable leverage over OEMs to make its apps default. Yet no patents are registered under Android at Google. The company Android was too young to have patents at the time of its acquisition. In most of these instances, the acquisitions replaced flailing internal products at Big Tech, thereby reducing competition in the market, as in Maps, or Youtube displacing Google Video, or Beats Music displacing Apple iTunes purchase-to-own model.

The major acquisitions in Figure 2 also highlight a pure patent motivation (for defensive reasons and to enter protected markets, such as with Alphabet buying Motorola Mobility and Apple buying Intel Smartphones). Alphabet acquired patents and phones from Motorola Mobility (for $12.45bn). This is by far Alphabet's largest acquisition to date - with Nest (2014) coming in second at $3.2 billion. Google executives at the time acknowledged patents played a role in the original Motorola deal (Kopytoff 2014). When Google sold Motorola Mobility to Lenovo for $2.91bn in 2014, it kept the majority of patents, for example (which our data tries to track and account for) (Google 2014). Google tried to use the acquisition to better enter into the smartphone market. But this failed and it sold off these assets to Lenovo.

Similarly, Apple's major patent acquisition (for $1 billion) comes from Intel's smartphone modem division, which was motivated by a desire to acquire patents and expertise to pursue chip independence from Qualcomm (Reuters 2022). This remains Apple's second largest acquisition to date (after Beats for $3 billion in 2014). As part of the deal, Apple took over 17,000 wireless technology patents, including protocols for cellular standards, modem architecture, and modem operations (Spiceworks 2022). Apple argues that it is impeded in developing its 5G smartphone modems by two Qualcomm patents (PhoneArena 2022), forcing it to continue to be Qualcomm's largest customer until at least 2026 (Spiceworks 2022).

Another insight into the motivation and nature of acquisitions comes from their over-representation of expired patents in previously acquired companies. Acquisitions account for 10% of patents, yet 20% of expired patents. This may be because the assets and patents are from mature companies with technology necessary for Big Tech to compete against established firms. Alternatively, this could result from risky investments made in young companies often with unproven technologies. Their assets may also compete with internal assets, rendering one set of assets redundant. Although Microsoft holds the highest number of expired patents, this is only because it has the highest number of total patents. In fact, Alphabet has by far the largest proportion of expired patents coming from acquisitions at 40% (3,319 patents), followed by Meta at 24% (78 patents). This is driven by Alphabet's acquisition of patents from Motorola Mobility. Although Google sold Motorola Mobility to Lenovo in 2014, it kept the majority of patents.

The existing literature provides additional insight into the motivation of Big Tech's acquisitions. One-third of the (unreported - non-HSR) majority-stake corporate acquisitions by Big Tech between 2010-2019 were motivated, by acquiring patents (12.5%) and (non human-capital) assets (20.6%), according to firms' self-reporting to the FTC (2021). Though this categorization suffered from definitional issues. Evidence on a "killer acquisition" motivation by Big Tech is mixed (Cunningham et al. 2021; Rinehart 2023). From a capabilities perspective, active integration of acquired technologies by Big Tech – instead of "killing" them – is supported at the patent-level by Gugler et al. (2023). Using a time-series of patent citations in technological classes, the authors find that patent citations increase by 36% for Big Tech's acquisitions (except for Apple's) after 2010 (compared to a control group), in marked contrast to acquisitions before then.

# 4 Acquired Capabilities and Potential Competition: Evidence from Big Tech's Patents

## 4.1 Capabilities Acquisition and Entry in the Market

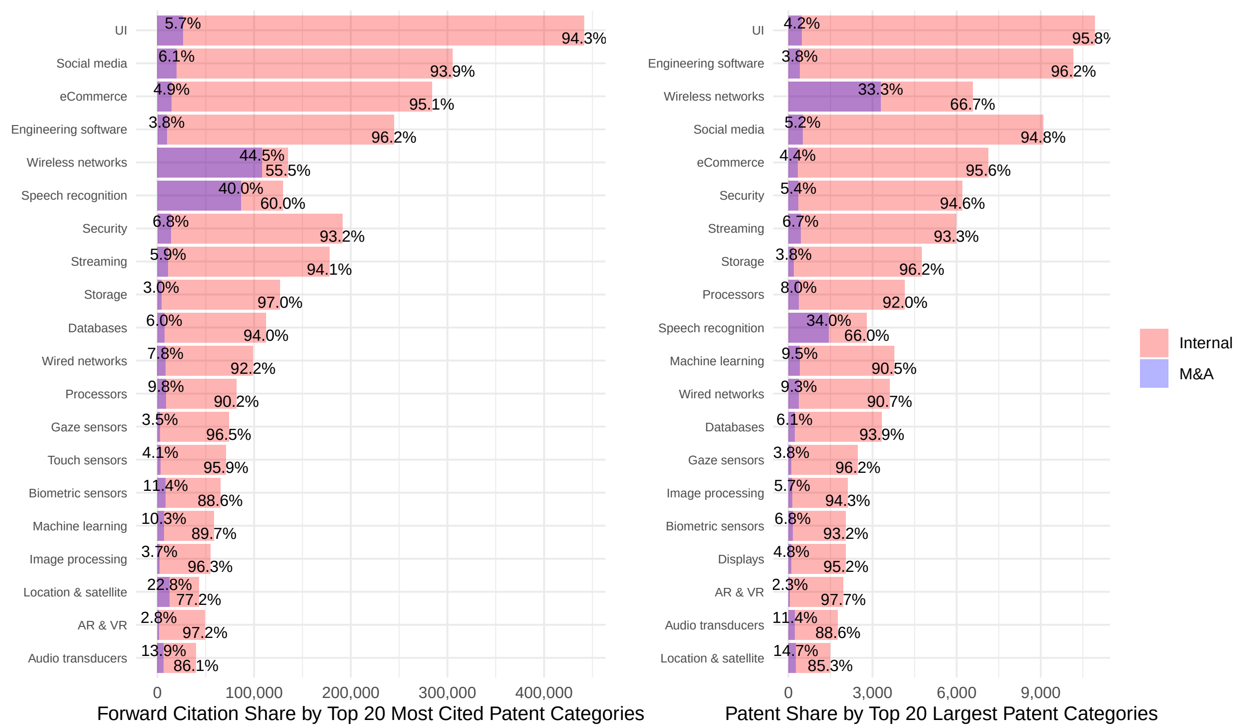
This section presents evidence showing a strong link between capabilities acquisition and entry in the market for goods and services. While individual acquisitions may have limited impact, their combined effect is significant. Excessive concentration in capabilities, especially cross-cutting ones (ML) or complementary combinations (product ecosystem), might be viewed with caution in a competitive framework.

Empirically, what have been the key capabilities acquired by Big Tech? Our patenting approach relies on technological categories. Various machine learning approaches have been proposed, including applying BERT to patents (Srebrovic et al. 2020). The US Patent Office (USPTO) (A. Toole et al. n.d.) uses the machine learning approach outlined in A. A. Toole et al. (2019) and Abood et al. (2018) - also known as "patent landscaping". We use the ML technology classifier from Cipher / LexisNexis Patent Research which is, as noted previously, an unsupervised learning based multi-class classifier, using 10 major technological categories and 122 subcategories, on a global basis. Within this, we focus on key technology categories. These same technology categories have been used by Big Tech who historically have used Cipher's patent dataset.

Figure 3 shows that, across all of Big Tech, user interface (UI), social media, and eCommerce related technologies dominate the total patents held by Big Tech (adjusted for citations), and unadjusted they are UI, engineering software, and wireless networks. The Figure highlights the importance of M&A to building out Big Tech's capabilities in speech recognition (40% of all citation adjusted patents), wireless networks (44.5%), location & satellite (22.8%), for example, Google and Apple Maps (see left-hand graph, adjusted for patent importance based on forward citations). But it has also been important to other hardware, including audio transducers (13.9%) used for speakers, processors (9.8%), and biometric sensors (11.4%). Machine learning patents also feature (10.3%).

Figure 4) highlights significant differences in how each Big Tech company uses M&A to acquire capabilities and compete in new and uncertain markets. Alphabet shows the greatest reliance on M&A for its capability development. This tracks its products fairly closely. For example, 72% of wireless networks and 46% of location & satellite citation adjusted patents come from M&A. The

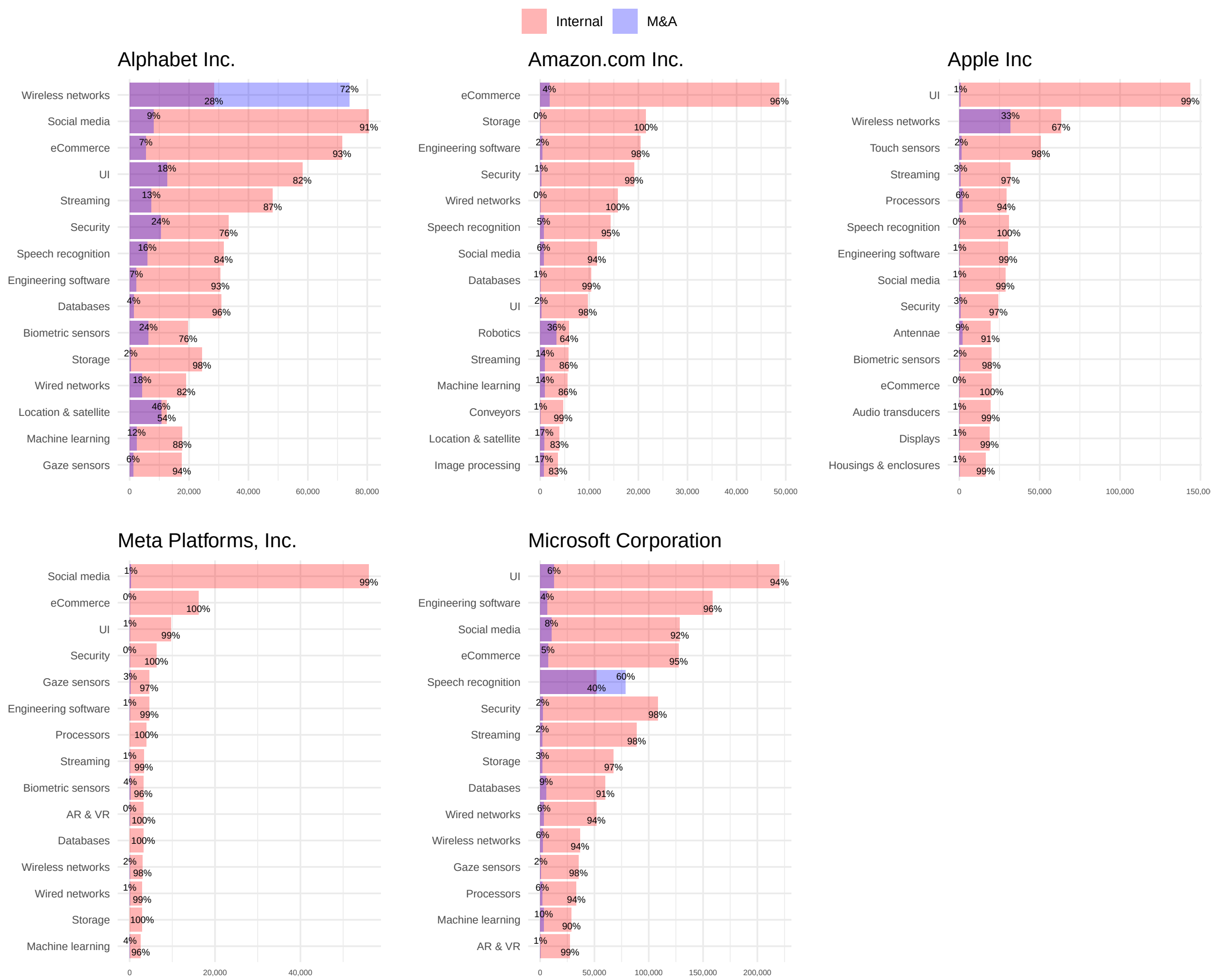
Figure 3. Capabilities Acquired by Major Technology Category (Adjusted for Forward Citations vs. Patent Totals)




Note: Top 20 UTT Technology categories held by Big Tech as a whole, by categories with largest total patent count. Showing share due to M&A and share to internal development.

latter underpins Google Maps. Such patents have been important to Amazon too. Amazon has used M&A to gain robotics capabilities (Zoox) but also streaming (Twitch) and machine learning (14%). As has Alphabet (12%), Microsoft (10%), and Meta (4%). Processors (6%), wireless networks (33%), and antenna (9%) have been important areas where Apple has used M&A to develop its capabilities.

Figure 4. Capabilities Acquired by Major Technology Category by Big Tech (forward citations)



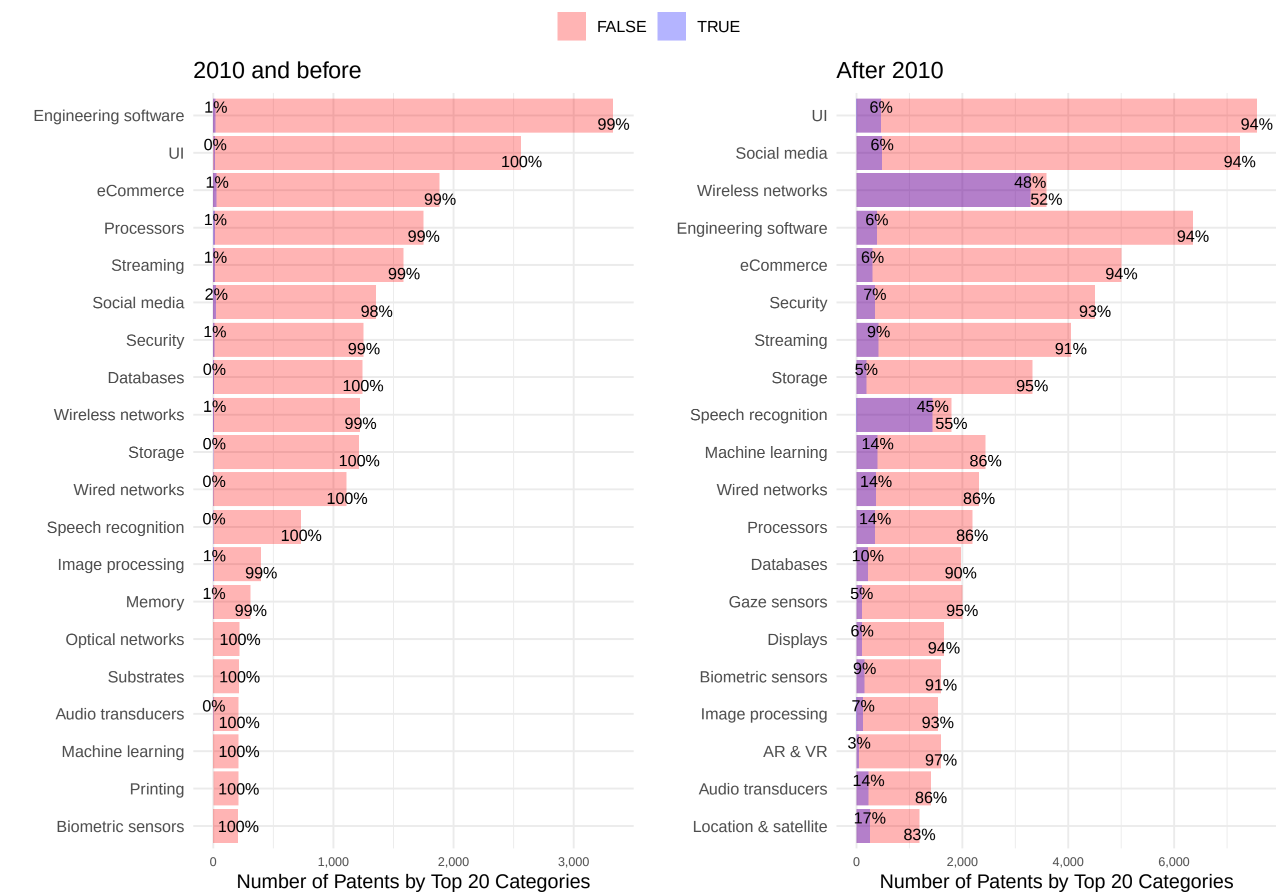

Note: Top UTT Technology categories held by each Big Tech, by categories with largest total patent count by forward citations. Showing share due to M&A and share to internal development.

The above technology categories are the largest patent categories held by Big Tech. These categories are different from the capabilities most acquired through M&A by Big Tech, which have tended to focus on practical hardware technologies to build out the 'things' which their software has filled. Looking at technology categories where at least 20% of forward citations are from M&A patents and total patent citations in that category are above the median of 2210 (to avoid very small categories):

amplifiers (52.5%, total forward citations 12,498), Wireless Networks (44.5%, 243,468), Photovoltaics (41.1%, 2,629), Speech recognition (40%, 216,322), Antennae (29.2%, 42564), ADC & DAC (27.1%, 8,419), Hinges (25.1%, 4,385), Batteries (23.5%, 9,277), Location & Satellite (22.8%, 55,357), and Inductors (21.2%, 13,877).

Next, Figure 5 explores what Big Tech's patents look like in 2010 or earlier compared with after 2010. The closer the correspondence between capabilities and product markets, the greater the potential for acquisitions of capabilities to warrant a competitive threat. Given Big Tech's access and control of existing platforms with large user (and producer) bases, as well as complementary technologies, we would expect to see a greater correspondence as the timeline from technology acquisition to consumer facing product can be dramatically shortened.

Figure 5. Diversification in Capabilities through M&A Leads to Ecosystems




Note: Showing Top 20 Patent Categories owned by Big Tech by total patent count before or equal to 2010 and after 2010. Year is calculated.

Figure 5 shows that, before 2010, patent capabilities were focused on engineering software overwhelmingly (defined here as purpose-built computer code to design and document a product). After

2010 much greater emphasis is put on building out engaging integrated hardware-software product ecosystems, with UI, social media, and wireless network patents dominating the top three. Storage also becomes far more important. User interfaces (UI) facilitate interactions between humans and computers, websites, or applications. Machine learning, gaze sensors (for eye tracking and gaming), displays, AR & VR, and location & satellite patent technology all become far more important post-2010. Machine learning patents show the largest increase after 2010 compared to pre-2010 levels.

M&A also becomes quantitatively central to all capabilities development after 2010 (shaded purple). Before 2010 little importance is attached to M&A in Big Tech's technological development in our dataset. But in practice, its qualitative impact was considerable. Consider, for example, Google's acquisition of DoubleClick, which helped Google dominate the advertising technology stack (New York Times 2020).

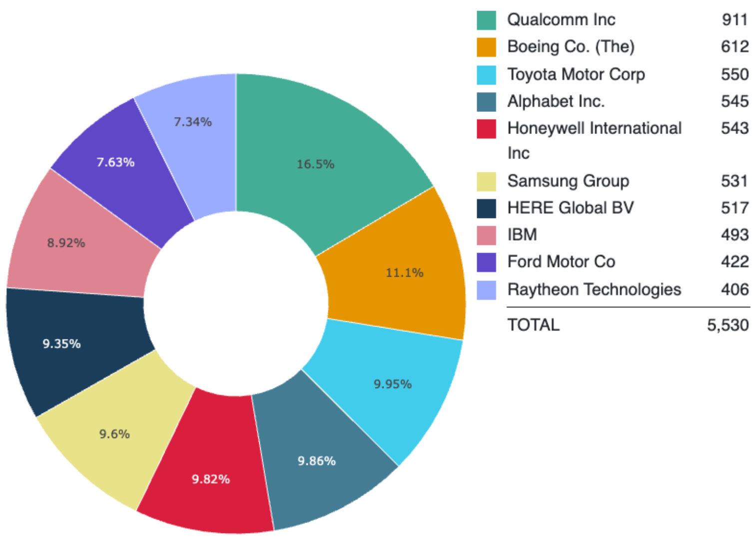
### 4.2 Regulating "Maps" or Location & Satellite Technology?

Below we explore the role of M&A in Google gaining and sustaining an advantage in Mapping capabilities and product markets. Google Maps grew out of three acquisitions in 2004: Keyhole, ZipDash, and especially Where 2 Technologies (Vox 2015; Gilbert et al. 2019). These were all young firms with few if any registered patents. But then Google's acquisition of Israeli mapping company Waze in 2013 - ultimately approved by the FTC, and the UK and Israel competition authorities - gave Google access to further technological dominance, through crowd sourced real-time traffic data.

Using Cipher's UTT ML classifier, we find that Alphabet holds twice as many patent families (758) in 'Location & Satellite' technology as any other Big Tech firm in our dataset (Microsoft with 375, followed by Apple with 347). Though Apple has engaged in dozens of acquisitions to try and compete with Google in Maps, including acquiring Placebase, Poly9, C3 Technologies, WiFiSlam, Locationary, HopStop.com, Embark, BroadMap, Spotsetter, Coherent Navigation, Mapsense, and Indoor.io (Wikipedia 2023).

One could include the other leaders in the market from outside of Big Tech to see their relative shares of "Location & satellite" technology. This analysis could also be refined to include only direct competitors in a given product category. We can apply this same UTT classifier, now also used by LexisNexis (2023), to the competitive landscape in this technological field for patents registered in the U.S. market. We see that Alphabet, Microsoft, Apple appear in the rankings but fall outside the top three. Alphabet is fourth, behind Qualcomm, Boeing, and Toyota. Microsoft is 15th and Apple 17th (not shown). This highlights that using technology alone to classify markets increases the potential size of the market and in turn the potential competition. But not all competitors are as likely to enter a given product market with their technology competencies. However, a complementary acquisition may enable rapid entry if the firm has related pre-existing competencies.

Figure 6. Location and Satellite Patent Ownership (Capabilities) in USA, Top 10 Shares




Note: Showing active patent family owners for USA for location and satellite technologies using a UTT ML patent classifier. This share excludes the next 5,000 patent owners from the market. This includes: Technologies related to satellite communication, geocentric orbit type satellites, remote sensing satellites and global positioning satellite (GPSS) and terrestrial based location technologies. For more information see LexisNexis (2023).

## 5 Do Capability Acquisition Strategies Differ by Big Tech Firm?

Using a logit regression, this section estimates what drives the probability of each Big Tech firm acquiring a patent externally. This builds on the dynamic capabilities literature, with its emphasis on "the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments" (D. J. Teece et al. 1997). The binary outcome variable is whether the patent is "externally" (i.e. via M&A) or internally developed. This treats each patent as largely independent from one another, focusing on each patent as a capability, potentially separable from the firm and capable of combination with other patents - both internal and external to the firm.

We have two main sets of predictors for this binary outcome, which correspond to two approaches to acquiring capabilities:

- *Hypothesis 1*: Big Tech firms acquire high-performing, often proven, assets (capabilities). This is proxied by patents which have high age-adjusted forward citations being more likely to be acquired. (Note that we use log age-adjusted citations as of December 2022 rather than at the time of acquisition.) We would expect the log(patent citations) predictor to be positive and high here.

- *Hypothesis 2*: Big Tech firms acquire assets based on their technological field, either to complement specific assets and/or to enter into completely new fields of production. These assets might have a wide variability in their proven commercial applications and viability. We proxy for this by use of technology field. We would expect the technology categorical predictor to be large and significant for certain categories.

Both are tested in this single regression. The patent-level model predicting the probability of a patent being acquired through M&A (dependent variable) can most simply be expressed as follows:

$$P(\text{External} = 1) \sim \text{Bernoulli}(\pi) \tag{1}$$

$$\text{logit}(\pi) = \ln\left(\frac{\pi}{1-\pi}\right) = \beta_1 \times \text{Citation\_log} + \beta_2 \times \text{Technology\_field}, \tag{2}$$

where $\pi$ represents the probability of `External` being 1. The first equation indicates that the patent response variable `External` follows a Bernoulli distribution with success probability $\pi$. The second equation, representing the logit link function, transforms this probability into a linear combination of the predictors: `Citation_log` and `Technology_field`. We use the `glm` function in R with a `binomial` family and `logit` link function. These predictors are in practice estimated at the patent level, subscript $\beta_i$, and are estimated in separate regressions for each Big Tech firm $\beta_j$ (of which there are five), and so could be written as $\beta_{j[i]}$ for clarity - even though estimated separately. Patent sample sizes vary for each Big Tech firm, as shown in the results Table 4. We are unable to control for the acquired firm from which the patents originate, as their number is disproportionately large compared to the patents.

Patent data is from Cipher/LexisNexis, including the UTT ML technology category used in the regression. The M&A indicator based on a matching from our M&A database combining Refinitiv, Wikipedia, and Webscraping. Results are shown below in Table 4 for hypothesis 1.

Our regression findings show mixed support for hypothesis 1 (Table 4), which is tested by the log(citations) predictor's coefficient. Amazon (0.48) tends to prioritize proven (or more mature) tech-

Table 4. Comparison of GLM Models: Predicting patent being acquired through M&A

| | *DV: Probability of a patent being from M&A* | | | | |
|---|---|---|---|---|---|
| | Alphabet | Amazon | Apple | Meta | Microsoft |
| $log$(Patent Citations) | $-0.003$ | $0.48^{***}$ | $0.22^{***}$ | $0.19^{**}$ | $-0.16^{***}$ |
| | (0.014) | (0.036) | (0.022) | (0.076) | (0.016) |
| Observations ($n$) | 29,273 | 14,670 | 23,605 | 6,618 | 41,297 |

Note: $^{*}$p<0.1; $^{**}$p<0.05; $^{***}$p<0.01

nologies in their acquisitions, to replace existing products or to run them as stand-alone businesses. Microsoft (-0.16) appears more exploratory, focusing less on patent quality and more on strategic or speculative acquisitions, including young firms with fewer citations. Meta (0.19) and Apple (0.22) lie somewhere in-between. Apple's acquisition strategy reflects a balanced approach, valuing patents with higher citation counts (indicative of innovation or impact), while still considering other factors, such as strategic alignment with its ecosystem. The result for Alphabet (-0.003) is inconclusive, showing a minimal effect with a large standard error (0.014).

Our findings for hypothesis 2 (Figure 7), tested in the same regression by the technology categorical variable for each patent, shows relevance across all Big Tech firms to varying degrees.

Alphabet has the highest baseline probability of buying a company across all technology categories (dotted vertical line crossing at -2.4), followed by Microsoft. Unlike the other Big Tech firms, Microsoft and Alphabet also have very strong technological preferences, buying patents in specific technology areas with positive probability, in user equipment for Alphabet and in speech recognition for Microsoft. Apple also has a strong preference for buying capabilities in user equipment. This highlights that for these firms, their dynamic capabilities are best enhanced through acquiring specific technological areas, irrespective of the proven efficacy of those capabilities in isolation from Big Tech.

Figure 7. How Important is Technology Category to Acquisition Motivation?


Note: Vertical dotted line is total (global) average effect across all technology categories showing the average tendency (probability) for a patent to be externally acquired. This is highest for Alphabet followed by Microsoft (though still negative given most patents are internally developed in our dataset). Showing estimates from separate logit regressions. Omitting two technology field estimates from Meta due to the confidence interval being too wide (smaller sample size). Patent data is from Cipher/LexisNexis including UTT ML technology category. M&A indicator based on a matching from our M&A database combining Refinitiv, Wikipedia, and Webscraping.

# 6 Discussion and Conclusion

Our empirical framework, utilizing an extensive patent dataset combined with M&A data, differentiates between Big Tech's acquired and internally developed patents. This approach offers insight into Big Tech's strategic behavior, underscoring the role of acquired technologies in shaping future competitive outcomes. Our study demonstrates the feasibility of analyzing capabilities through patents within a dynamic competition framework. The interplay between capability acquisition, market entry, and market dominance is complex and uncertain, making clear predictions difficult (Argente et al. 2020). We have highlighted the influence of Big Tech's M&A activity in dozens of products, and broader product-market ecosystems, such as digital mapping, where Alphabet has fortified its position through multiple small and one major strategic acquisitions.

Our findings reveal that at least 10.3% of Big Tech's patent portfolios originate from acquisitions, with this percentage rising to 13.1% when weighted by forward citations. This substantial contribution of externally acquired patents underscores how Big Tech relies on M&A to build capabilities across diverse technological domains. The regression analysis further demonstrates that Big Tech firms pursue distinct acquisition strategies: Amazon, Apple and Meta tend to acquire more proven technologies with established citation patterns, while Microsoft is more willing to acquire nascent, unproven technologies.

The temporal analysis shows a significant shift in Big Tech's capability development after 2010, with greater emphasis on integrated hardware-software ecosystems and emerging technologies like machine learning. This transition coincides with Big Tech's expansion beyond their core markets into adjacent product categories, suggesting that capability acquisition through M&A has been instrumental in facilitating this diversification.

From a policy perspective, our findings support a more dynamic approach to antitrust assessment that considers both the immediate and potential future competitive impacts of acquisitions. Rather than evaluating acquisitions solely based on current market share or immediate competitive overlap, regulators should consider how acquired capabilities might strengthen an incumbent's position across multiple markets or enable entry into adjacent markets. This approach aligns with recent developments in merger guidelines, such as the U.S. Department of Justice's recognition that patterns of acquisitions may violate antitrust laws even when individual transactions appear benign (FTC and DoJ 2023, p. 23).

The significant role of M&A in Big Tech's capability development challenges the traditional view that these companies primarily grow through internal innovation. While internal R&D remains important, our data shows that external acquisition of technologies has been crucial in their ability to enter new markets and expand their ecosystems. This finding has implications for how we understand innovation dynamics in the technology sector and suggests that more scrutiny should be applied to acquisitions of innovative companies, particularly when the acquirer has demonstrated a pattern of using such acquisitions to consolidate its position across multiple markets.

A limitation of our study is the difficulty in distinguishing between different motivations for patent acquisitions. Some patents may be acquired for defensive purposes rather than to incorporate the underlying technology into products. Additionally, our data cannot fully capture the value of non-patented innovations or human capital acquired through M&A. Future research could extend this framework by incorporating additional metrics of innovation beyond patents, examining the post-acquisition integration of acquired technologies in more detail, and exploring how non-patented assets and human capital contribute to capability building. Despite these limitations, our study provides a valuable methodological contribution by demonstrating how patent data can be used to assess the competitive implications of M&A activity in technology-intensive industries.

In conclusion, our research highlights the critical role of M&A in Big Tech's capability development and ecosystem expansion. By focusing on the acquisition of technological capabilities rather than just market share, we provide a more nuanced understanding of how these companies have built and maintained their competitive positions. This perspective suggests that competition authorities should pay greater attention to patterns of acquisitions and their potential to shape future competitive dynamics, even when individual transactions appear benign from a traditional antitrust perspective (Areeda et al. 2023, p. 701b).

## References


Abood, Aaron and Dave Feltenberger (2018). "Automated Patent Landscaping." *Artificial Intelligence and Law* 26.2, pp. 103–125.

Akcigit, Ufuk and Sina T Ates (2023). "What Happened to US Business Dynamism?" *Journal of Political Economy* 131.8, pp. 2059–2124.

Areeda, Phillip and Herbert Hovenkamp (2023). *Antitrust Law: An analysis of antitrust principles and their application*. Online Edition.

Argente, David et al. (2020). "Patents to Products: Product innovation and firm dynamics." *CEPR Discussion Paper No. DP14692*.

Bryan, Kevin A and Erik Hovenkamp (2020). "Antitrust Limits on Startup Acquisitions." *Review of Industrial Organization* 56, pp. 615–636.

Cadman, Richard (2023). "Dynamic Competition and Big Tech Mergers Report." *British Institute of International and Comparative Law*. URL: https://www.biicl.org/publications/dynamic-competition-and-big-tech-mergers?cookiesset=1&ts=1704894474.

Caffarra, Cristina, Matthew Elliott, and Andrea Galeotti (June 2023a). *'Ecosystem' Theories of Harm in Digital Mergers: New insights from network economics, part 1*. URL: https://cepr.org/voxeu/columns/ecosystem-theories-harm-digital-mergers-new-insights-network-economics-part-1.

— (June 2023b). *'Ecosystem' theories of harm in digital mergers: New insights from network economics, part 2*. URL: https://cepr.org/voxeu/columns/ecosystem-theories-harm-digital-mergers-new-insights-network-economics-part-2.

Callaham, John (2022). *Google made its best acquisition nearly 17 years ago: Can you guess what it was?* Android Authority. URL: https://www.androidauthority.com/google-android-acquisition-884194/.

Cohen, Lauren, Umit G. Gurun, and Scott Duke Kominers (2019). "Patent Trolls: Evidence from Targeted Firms." *Management Science* 65.12, pp. 5461–5486.

Cunningham, Colleen, Florian Ederer, and Song Ma (2021). "Killer Acquisitions." *Journal of Political Economy* 129.3, pp. 649–702.

Doctorow, Cory (May 2023). *Google's AI Hype Circle*. URL: https://doctorow.medium.com/googles-ai-hype-circle-6158804d1299.

Evans, David S and Keith N Hylton (2008). "The Lawful Acquisition and Exercise of Monopoly Power and its Implications for the Objectives of Antitrust." *Competition Policy International* 4.2, p. 203.

FTC (2021). "Non-HSR Reported Acquisitions by Select Technology Platforms, 2010–2019: An FTC Study." *FTC Study*, p. 246. URL: https://www.ftc.gov/system/files/documents/reports/non-hsr-reported-acquisitions-select-technology-platforms-2010-2019-ftc-study/p201201technologyplatformstudy2021.pdf.

FTC and DoJ (2023). *2023 Merger Guidelines*. URL: https://www.ftc.gov/news-events/news/press-releases/2023/12/federal-trade-commission-justice-department-release-2023-merger-guidelines.

Giczy, Alexander V, Nicholas A Pairolero, and Andrew A Toole (2022). "Identifying artificial intelligence (AI) invention: A novel AI patent dataset." *The Journal of Technology Transfer* 47.2, pp. 476–505.

Gifford, Daniel J and Robert T Kudrle (2011). "Antitrust approaches to dynamically competitive industries in the United States and the European Union." *Journal of Competition Law and Economics* 7.3, pp. 695–731.

Gilbert, Ben and David Rosenthal (2019). *Google Maps*. Acquired Podcast: Season 5, Episode 3. URL: https://www.acquired.fm/episodes/google-maps.

Google (2014). *Lenovo to Acquire Motorola Mobility*. URL: https://blog.google/inside-google/company-announcements/lenovo-to-acquire-motorola-mobility/.

Graham, Stuart JH and David C Mowery (2003). “Intellectual property protection in the US software industry.” *Patents in the Knowledge-based Economy* 219, p. 231.

Griliches, Zvi (1998). “Patent Statistics as Economic Indicators: a survey.” *R&D and productivity: the econometric evidence*, pp. 287–343.

Gugler, Klaus, Florian Szücs, and Ulrich Wohak (2023). *Start-up Acquisitions, Venture Capital and Innovation: A Comparative Study of Google, Apple, Facebook, Amazon and Microsoft*. Tech. rep. Tech. rep., Vienna University of Economics and Business, Department of Economics.

Hall, Bronwyn H, Adam Jaffe, and Manuel Trajtenberg (2005). “Market value and patent citations.” *RAND Journal of economics*, pp. 16–38.

Hovenkamp, Herbert (2008). “Schumpeterian Competition and Antitrust.” *U Iowa Legal Studies Research Paper* 08-43.

Jacobides, Michael G, Carmelo Cennamo, and Annabelle Gawer (2018). “Towards a Theory of Ecosystems.” *Strategic management journal* 39.8, pp. 2255–2276.

Jacobides, Michael G and Ioannis Lianos (2021). *Regulating Platforms and Ecosystems: An introduction*.

Kim, Bongsun et al. (2016). “The Impact of the Timing of Patents on Innovation Performance.” *Research Policy* 45.4, pp. 914–928.

Kopytoff, Verne (2014). *Motorola is a Gargantuan Mistake Only Google Could Afford to Make*. Time Magazine. URL: https://time.com/2650/motorola-is-a-gargantuan-mistake-only-google-could-afford-to-make/.

Krestel, Ralf et al. (2021). “A survey on deep learning for patent analysis.” *World Patent Information* 65, p. 102035.

Levin, Richard C et al. (1987). “Appropriating the Returns from Industrial Research and Development.” *Brookings papers on economic activity* 1987.3, pp. 783–831.

LexisNexis (Oct. 2023). *Universal Technology Taxonomy*. URL: https://www.lexisnexisip.com/wp-content/uploads/2023/10/Universal-technology-taxonomy-download.pdf.

Makinen, Marita, David Haber, and Anthony Raymundo (2012). “Acqui-Hires for Growth: Planning for Success.” *Venture Capital Review* 28. Accessed: 2025-02-25. URL: https://www.lowenstein.com/media/4480/ey-venture_capital_review.pdf.

McSweeny, Terrell and Brian O’Dea (Feb. 2018). “Data, Innovation, and Potential Competition in Digital Markets — Looking Beyond Short-Term Price Effects in Merger Analysis.” *Antitrust Chronicle 7, 11 (Competition Policy International)*.

Morton, Fiona M Scott and Carl Shapiro (2013). “Strategic Patent Acquisitions.” *Antitrust LJ* 79, p. 463.

New York Times (Sept. 2020). *This Deal Helped Turn Google Into an Ad Powerhouse. Is That a Problem?* URL: https://www.nytimes.com/2020/09/21/technology/google-doubleclick-antitrust-ads.html.

Pavitt, Keith (1982). “R&D, Patenting and Innovative Activities: A statistical exploration.” *Research Policy* 11.1, pp. 33–51.

Petit, Nicolas and David Teece (2020). “Taking Ecosystems Competition Seriously in the Digital Economy: A (Preliminary) Dynamic Competition/Capabilities Perspective.” *Capabilities Perspective (December 9, 2020)*.

Petit, Nicolas and David J Teece (2021). “Innovating Big Tech Firms and Competition Policy: Favoring dynamic over static competition.” *Industrial and Corporate Change* 30.5, pp. 1168–1198.

PhoneArena (2022). *Qualcomm patent puts wrench into Apple’s plan to build 5G modem*. URL: https://www.phonearena.com/news/qualcomm-patent-puts-wrench-into-apples-plan-to-build-5g-modem_id141078.

Posner, Richard (2006). “Do We Have Too Many Intellectual Property Rights?” *The Eighth Annual Honorable Helen Wilson Nies Memorial Lecture in Intellectual Property Law*.

Reuters (Oct. 2022). *Apple loses second bid to challenge Qualcomm patents at U.S. Supreme Court*. URL: https://www.reuters.com/technology/apple-loses-second-bid-challenge-qualcomm-patents-us-supreme-court-2022-10-03/.

Rinehart, William (2023). “Are ‘Killer Acquisitions’ by Tech Giants a Real Threat to Competition?” *The Center for Growth and Opportunity*.

Saltiel, Joseph (2019). *Five Years After Alice*. WIPO Magazine. URL: https://www.wipo.int/wipo_magazine/en/2019/04/article_0006.html.

Sharma, Ruchir (2023). *What’s Wrong with Tech Giants Riding the AI Wave*. URL: https://www.ft.com/content/31dc322a-b2a1-4ef3-9a9d-d7268d58fac5.

Sidak, J Gregory and David J Teece (2009). “Dynamic competition in antitrust law.” *Journal of Competition Law and Economics* 5.4, pp. 581–631.

Spiceworks (2022). *Qualcomm-Apple 5G Modem Deal*. URL: https://www.spiceworks.com/tech/innovation/news/qualcomm-apple-5g-modem-deal/.

Srebrovic, Rob and Jay Yonamine (2020). “Leveraging the BERT algorithm for Patents with TensorFlow and BigQuery.” *White paper*.

TechCrunch (Nov. 2011). *Amazon Snaps Up Yap and Its Voice Recognition Technology*. URL: https://techcrunch.com/2011/11/09/amazon-snaps-up-yap-and-its-voice-recognition-technology/.

— (Aug. 2013). *Facebook Acquires Mobile Technologies, Speech Recognition, and Jibbigo App Developer*. URL: https://techcrunch.com/2013/08/12/facebook-acquires-mobile-technologies-speech-recognition-and-jibbigo-app-developer/.

Teece, David J (1986). “Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy.” *Research policy* 15.6, pp. 285–305.

— (2007). “Explicating dynamic capabilities: the nature and microfoundations of (sustainable) enterprise performance.” *Strategic Management Journal* 28.13, pp. 1319–1350.

— (2010). “Technological innovation and the theory of the firm: the role of enterprise-level knowledge, complementarities, and (dynamic) capabilities.” *Handbook of the Economics of Innovation*. Vol. 1. Elsevier, pp. 679–730.

— (2012). “Next-Generation Competition: New concepts for understanding how innovation shapes competition and policy in the digital economy.” *JL Econ. & Pol’y* 9, p. 97.

Teece, David J, Gary Pisano, and Amy Shuen (1997). “Dynamic capabilities and strategic management.” *Strategic management journal* 18.7, pp. 509–533.

Toole, Andrew et al. (n.d.). “Inventing AI: Tracing the diffusion of artificial intelligence with US patents.” *US Patent and Trademark Office, Alexandria* 5 (), p. 2020.

Toole, Andrew A et al. (2019). “The Promise of Machine Learning for Patent Landscaping.” *Santa Clara High Tech. LJ* 36, p. 433.

Tunguz, Tom (2025). *Startup Acquihire Trends*. Accessed: 2025-02-25. URL: https://tomtunguz.com/startup-acquihire-trends/.

UK Competition and Markets Authority (2021). *Merger Assessment Guidelines (MAGs) for Public Consultation*. Accessed: 2025-01-02. URL: https://assets.publishing.service.gov.uk/media/61f952dd8fa8f5388690df76/MAGs_for_publication_2021_--_.pdf.

— (2022). *Final Report: Meta/Giphy Merger Investigation*. Accessed: 2025-01-02. URL: https://assets.publishing.service.gov.uk/media/635017428fa8f53463dcb9f2/Final_Report_Meta.GIPHY.pdf.

UK Competition Appeal Tribunal (2022). *Case No. 1429/4/12/21: Meta Platforms Inc vs. Competition and Markets Authority*. Accessed: 2025-01-02.

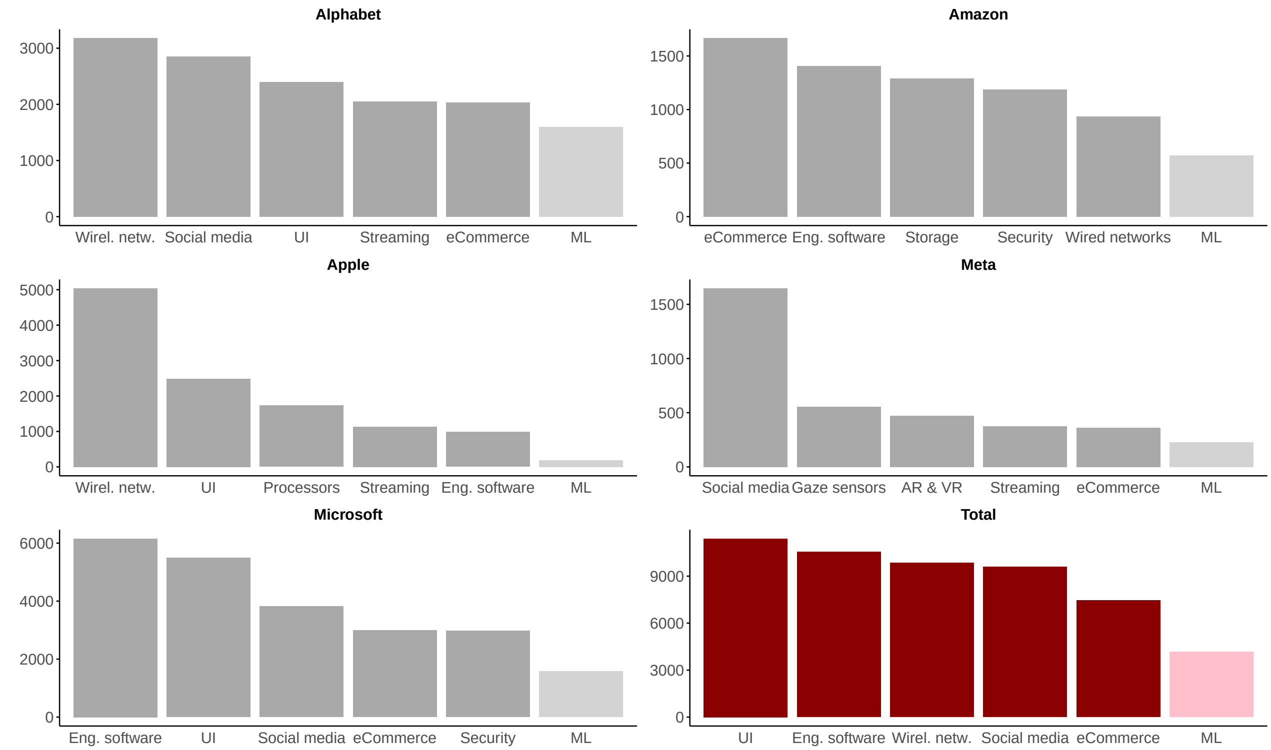
Vox (Feb. 2015). *Ten Years of Google Maps: From Slashdot to Ground Truth*. URL: https://www.vox.com/2015/2/8/11558788/ten-years-of-google-maps-from-slashdot-to-ground-truth.

Wikipedia (2023). *List of Mergers and Acquisitions by Apple*. Accessed: December 2023. URL: https://en.wikipedia.org/wiki/List_of_mergers_and_acquisitions_by_Apple.

# Appendix

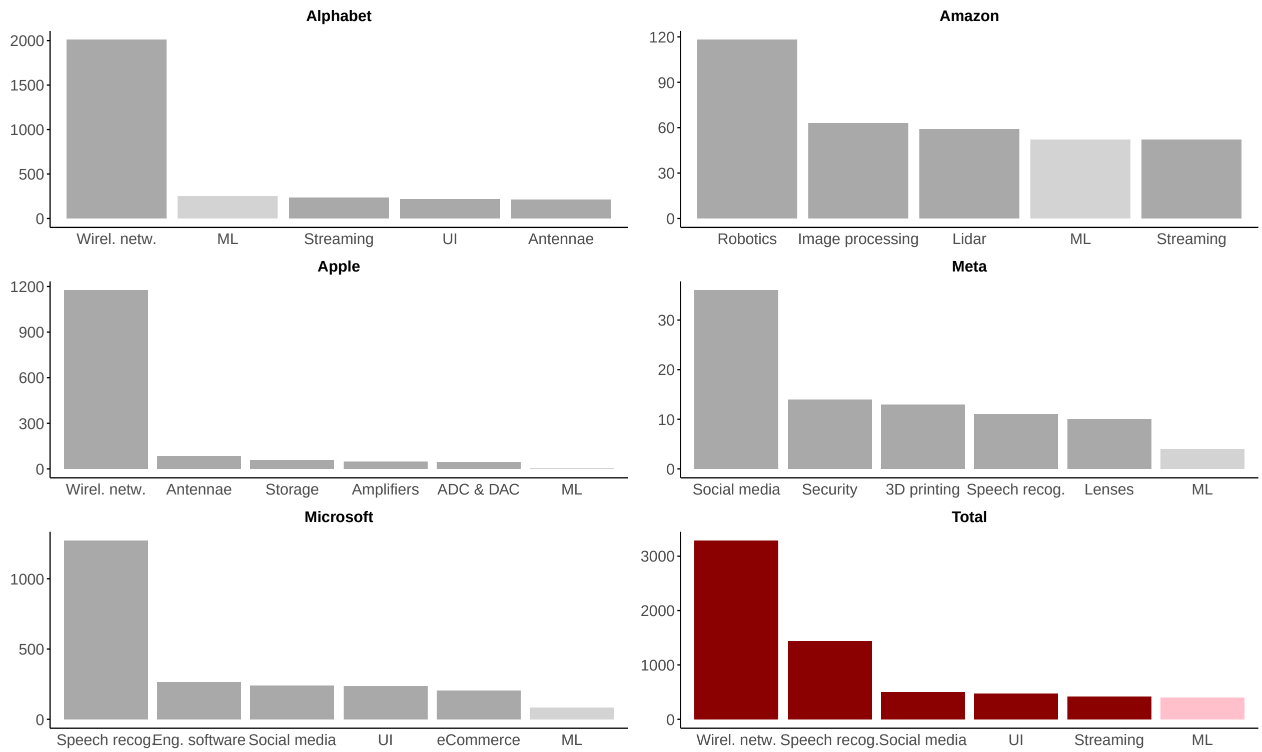
# A Further Descriptive Statistics

Figure 8. Big Tech's Top 5 Technology Sectors with Most Patents + Machine Learning



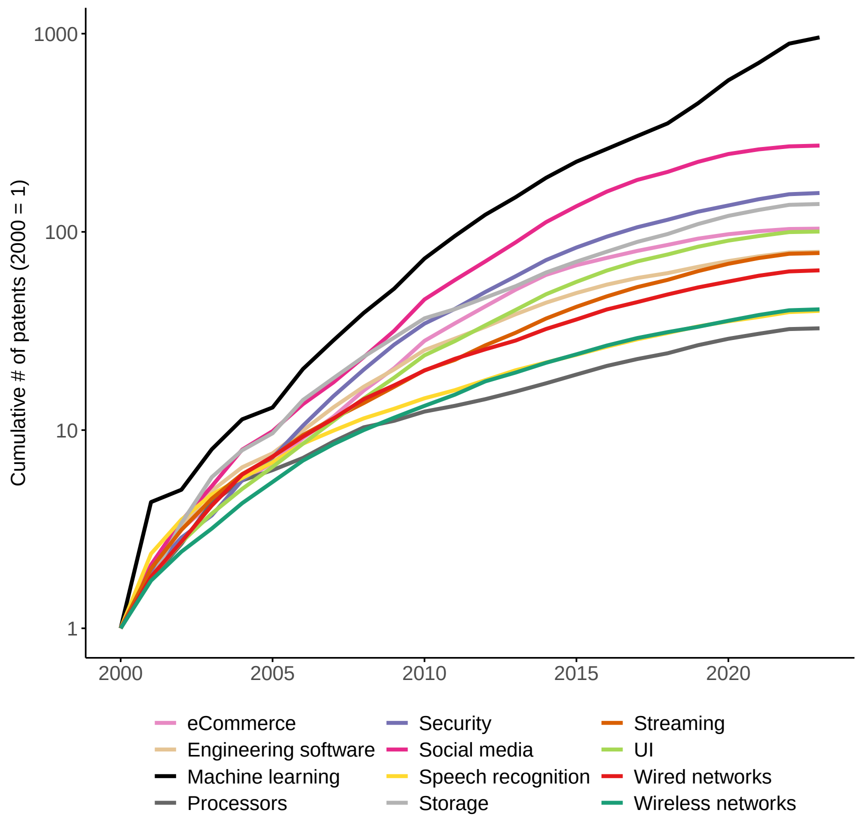

Note: *Top five technology sectors by total patents, along with the additional Machine Learning sector. Patent data is from Cipher/LexisNexis, including the UTT ML technology category. The M&A indicator is based on a matching process from our M&A database, which combines data from Refinitiv, Wikipedia, and web scraping.*

Figure 9. Big Tech's Top 5 M&A Technology Sectors with Most Patents + Machine Learning




Note: *Top five technology sectors by externally obtained patents through M&A, along with the additional Machine Learning sector. Patent data is from Cipher/LexisNexis, including the UTT ML technology category. The M&A indicator is based on a matching process from our M&A database, which combines data from Refinitiv, Wikipedia, and web scraping.*

Figure 10. Cumulative Big Tech Stock of Patents by Sector




Note: *Cumulative stock of patents of the Big 5 by sector, normalized with respect to 2000. The time span is from 2000 to 2022. Patent data is from Cipher/LexisNexis, including the UTT ML technology category.*

# B Cleaning Cipher Data

To construct our analytical dataset on a patent-family level, we utilize Cipher - a dataset from global patent databases which offers a wealth of information such as patent owners, assignees, technology areas, original assignees, and status. Recently bought by LexisNexis, the company manually creates the owner and parent categories. They also use ML to classify technology types. The Cipher dataset also includes expired patents. This appendix describes the steps we took to clean the data.

## B.1 Identifying Patent Originators

Our aim is to identify the patents that Big Tech owns as a result of acquisitions of other firms. However, we encounter three main issues in matching names in our acquisition-level dataset (Refinitiv) to Cipher. Firstly, our patent data does not clearly indicate if the patents that Big Tech owns come from another company. Secondly, it does not track assignee data, which means it may not recognize certain patents that Big Tech owns. Finally, the firm names are inconsistent between datasets, making it challenging to correctly identify each firm.

To overcome these obstacles, we merge the patent dataset and the acquisition dataset by first algorithmically creating a list of subsidiaries of Big Tech and the resulting patents assigned to them. However, we discovered three distinct scenarios that made the process more complex.

Firstly, many patents that are internally developed rely on externally acquired patents and human capital, which our study underestimates. To account for this, we ensure that any patents assigned to the acquired firm are properly matched with Big Tech, even if the assignee's name has not been formally changed. Secondly, we found that Big Tech owns patents that are not formally registered to them but have been acquired through M&A. After acquiring a company, Big Tech may retain the acquired company's patents as being registered under the acquired firm's name (assignee), with no formal change made to whom the owner of the patent is registered to. To address this scenario, we use our M&A list to match the acquired firm's patents with Big Tech, noting that the acquired firm and their patents are ultimately under the control of a Big Tech firm (even if not formally registered to them). Thirdly, we encountered scenarios where Big Tech owned patents that were registered under an internally developed subsidiary's name. A patent may be registered as belonging to a seemingly unrelated firm with no apparent links to Big Tech, but in fact, the seemingly unrelated firm was a subsidiary of Big Tech developed internally with no direct M&A links. To address this scenario, we applied our knowledge of Big Tech's corporate structure to discover the company was owned by Big Tech but not through acquisitions but instead through internal development.

We then thoroughly clean this list to address inconsistencies between the firms and other datasets we are using. For example, if a firm has "LLC" as part of its name in one dataset and not in another, we must ensure that the firms correctly match to avoid missing important data. Finally, we check this cleaned list against our M&A dataset from Refinitiv to trace the origins of the patents that Big Tech owns. This enables us to adequately demonstrate that all the companies which Big Tech has bought are actually listed as subsidiaries of Big Tech and which of the resulting patents are then assigned to Big Tech due to the acquisitions of those firms.

### B.2 Merging Cipher with Acquisition Data

One of the major challenges we faced in using Cipher was to standardize owner, assignee, organization and originator names to reconcile firm names for matching with other datasets, specifically M&A events. We did so by developing a multistep cleaning algorithm which cleaned and standardized all company names. The raw data had 146,664 patents and after processing, we were left with 127,299 patents in the dataset.

First, we filter the dataset to contain patent information from 1980 and onward. Then, we remove all instances of ".com" from the dataset. After, the misspellings of various firms and extra words are corrected (e.g., "amazong" needs to be "Amazon" and "Mela" needs to be "Meta"). This ensures that we can accurately identify and target the appropriate firms during the cleaning process. Then, we translate foreign firm names into English to facilitate accurate detection. The names are also then stripped of punctuation (commas and periods) and capitalization. We then remove words after the key Big Tech names (e.g., microsoft corporation changes to just microsoft, apple inc changes to just apple). This isolates the big tech company's stem name (the main body of the firm name) excluding any extraneous suffixes. A challenge that arose with this step is that in a few instances, some firms could be unintentionally removed from the data, so we have to manually fix them. For example, "metaswitch networks ltd" was left to be "meta", which is not an accurate representation of the original firm name so it must be reverted to "metaswitch networks ltd". There are other firms that need manual correcting as well. After this, we then find and take out extra words, legal entity endings, abbreviations, and redundant characters. After all these steps, there are still firms that may need specific alterations to ensure proper cleaning. Finally, we split firms with multiple owners or assignees into multiple columns, which are denoted using dashes, vertical lines, and brackets.

After all the cleaning, we then run the matching script and check the matching results manually to confirm accuracy.

## C Acquisition Data

We collect firm acquisition-level data from two sources. We first extract all announced and completed M&As (with complete information on acquirer and target firms) and announced and effective dates from Refinitiv which is provided by Refinitiv Desktop. This comprehensive dataset contains firms, their acquirers, ultimate parents, and deal status. However, upon further inspection, we found that this dataset was incomplete and missing many M&A firms, so we manually develop an additional dataset using the information found from Big Tech Mergers & Acquisitions Wikipedia pages. We manually compile the missing data from the Wikipedia pages into a new dataset with similar columns as Refinitiv. This dataset includes details such as the deal confirmation source, the acquirer ultimate parent, the deal status, and the ultimate parent. Furthermore, we believe some additional firms were not captured by Refinitiv and Wikipedia, so we manually search for them using various sources including news articles, company reports, and other publicly available data sources. We format the data in the same way as the Wikipedia data.

First, we clean the Refinitiv dataset. Since the same firm could appear in different databases under slightly different names, we create standardized and homogeneous names by removing extraneous words as well as stripping the names of capitalization. The process for cleaning this dataset is similar to that used for cleaning the patent-level data, but there are some key differences. With this dataset, we first add a few missing firms and then filter out deals made before 2000. Then, we simplify the names of all big tech firms (e.g., "Amazon.com Inc" becomes "Amazon", "Facebook Inc" becomes "Meta"). Next, joint buyouts where the ultimate acquirer is not listed as big tech is filtered out. We then remove instances of ".com" and split firms that have brackets and dashes into multiple columns. After, firms where the "Deal Type" is repurchased are extracted. Firms Expedia Inc, Giphy Inc, Delta Airlines Inc and WestJet Airlines Ltd are removed. Then, commas and periods are removed from the company names, and they are all changed to lowercase. We find and take out extra words, legal entity endings, abbreviations, and redundant characters. We also remove any special characters from the end of the firm names. Again, after all the cleaning, there are still numerous firms that may need specific alterations to ensure they are correctly retained and standardized. The raw data had 918 firms before processing and afterwards had 770 firms.

Next, we standardize and refine the data from Wikipedia. To create the dataset, we copied over the tables from the Wiki pages and format the information in a spreadsheet with similar formatting to the Refinitiv dataset. We removed redundant and unnecessary columns, such as “Country”, “Talent Acquired”, and “Related to”. Similar to how we process the Refinitiv data, we remove the deals made in 2023, extract instances of “.com”, and split firms with brackets and dashes. Then, we remove the punctuation, make everything lowercase and remove extraneous words. We also correct specific firms manually as they could be incorrectly manipulated in the cleaning process. Then we match these cleaned firms with the cleaned Refinitiv data and remove duplicates. Before merging the non-duplicate firms with the Refinitiv file to create a larger M&A file, we manually check the Deal Status and deal control variable to ensure that the deals are correct and actually occurred. Confirmation sources are also included in the data. The raw data had 865 firms before processing and afterwards had 210 firms.

Additionally, we cleaned other publicly available acquisition data in the same method as the Wikipedia dataset. This dataset was created by gathering information from reliable online resources that provided the firm and acquiror details, and then formatting the data in a very similar way to the Wikipedia data. We process this third dataset the exact way we did previously with Wikipedia. All deals before 2000 were filtered out, instances of “.com” were removed and firms were split from brackets and dashes. All capitalization, punctuation and unnecessary words were removed from the names. Manual corrections are made to specific firms to ensure they are not wrongly manipulated. These fully cleaned firms are then matched with the cleaned Refinitiv and Wikipedia data and any duplicates are removed. We then manually check the Deal Status and deal control variable to make sure these events are correct. The confirmation sources are recorded and included in the data. The non-duplicates are merged with the M&A file with the deals from the Refinitiv and Wikipedia datasets. The raw data had 457 firms before processing and afterwards had 21 firms that we added to the total M&A dataset.

After we add the missing data from the other two datasets into Refinitiv to create one complete M&A dataset, the dataset has a total of 995 acquired firms until 2022 (acquisition date). To the best of our knowledge, this combined process provides the most comprehensive database of acquisitions.

With the cleaned acquisition data compiled from both sources, we can accurately link acquisition events to their respective target firms and can begin the process of merging the patent and acquisition datasets. We combine our acquisition database with the patent data through a name-matching algorithm combined with manual checks. The merged patent and acquisition data show acquisition activities in our analytical dataset with 22.5 percent of acquisitions recorded in our patent database. This means that 225 M&A events from our acquisition dataset are also present in our patent dataset.

# D Merging Patent and Acquisition Data

In this section, we describe the process to merge patent and acquisition data with the Cipher patent database by matching company names with the owner, assignee, and original assignee names in the Cipher patent database. To minimize potential problems introduced by the minor discrepancies between different versions of the patent database and the M&A dataset, we run the cleaning algorithm to source the most standardized firm information. After this step, each company in the patent and acquisition database will have its original firm name and the target standardized name.

## D.1 Name Standardization and Cleaning

We begin by standardizing company names in our patent and acquisition database using our developed name standardization algorithm that is described in the appendices above. As some names are misspelled or include additional characters that prevent exact matching, this cleaning algorithm homogenizes these firm names and helps to isolate the company's stem name by removing redundant words, stripping punctuation, and making all into lowercase.

## D.2 The Matching Procedure

With these standardized and stem company names, we match the patent and M&A databases with the following procedure:

1. We first create a list of the columns we want to use to find matches in the Cipher patent dataset. As we are looking to match the M&A firms in the standardized target name ("target_no_brackets_undashed" column) with the potential patents they might hold, we must look for matches in Cipher's cleaned ownership, assignee, or original assignee columns. Specif-

ically, these columns are “owner1”, “owner2”, “owner3", “owner4”, “assignee1”, “assignee2”, “assignee3”, “assignee4”, “original_assignee1_no_bracket”, and “original_assignee2_bracket”.

2. Then, we initialize two data frames with empty columns to hold the matched data. For each row in Refinitiv, the algorithm searches each column in Cipher for a match.
   (a) If a match is found, it sets the value of the "match" column in the matched Refinitiv data frame to TRUE for that row in Refinitiv. It also sets the value of the "matches" column in the matched Cipher data frame to TRUE for the row(s) in Cipher that match the value in Refinitiv. Additionally, it populates the year of acquisition, the original target full name, the data source (Refinitiv, Wikipedia or other publicly available) and the cleaned target name we matched from in the matched Cipher data frame with the corresponding values from the M&A dataset.
   (b) If a match is not found, then the next row is searched.
3. After the exact matching process is complete, we calculate the percentage of matches found. We also manually check if the matches are correctly identified and if there are not false positives or false negatives.

Ex-post duplicate matches were removed to ensure that only the Big Tech acquirers patents was matched with the target firm, rather than other Big Tech companies who also might own patents from the target firm but which they did not acquire through M&A.