In-Depth Analysis of Windows 11 Telemetry

Yazılım Mühendisliği Anabilim Dalı

Dönem Projesi

Cengizhan Öner

Proje Danışmanı: Prof. Dr. Femin Yalçın Küçükbayrak

Ağustos 2024

In-Depth Analysis of Windows 11 Telemetry

Abstract

Telemetry is the process of collecting and sending information electronically from objects that are far away. In technology, it often means gathering data from users, sometimes without their direct consent. This research looks into how Windows 11, Microsoft's latest operating system, collects data from users. It focuses on the types of data collected, where this data goes, how Microsoft uses it, and the related legal and ethical issues. Windows 11 collects two main types of data, required diagnostic data and optional diagnostic data. Required data, collected no matter what, helps Microsoft monitor system performance and stability. Optional data, collected by default, includes detailed error and crash reports to improve user experience. This study aims to explain the impact of these practices on user privacy and data security and to explore the legal regulations around data collection, highlighting the balance between technological progress and ethical considerations.

Keywords: Telemetry, privacy, operating system, diagnostic data, bloatware

Windows 11 Telemetrisinin Detaylı Analizi

Öz

Telemetri bilginin uzaktaki nesnelerden elektronik olarak toplanma ve gönderilme sürecidir. Teknoloji bağlamında ise telemetri, açık rızaları olmaması rağmen kullanıcıdan toplanan veri anlamına gelmektedir. Bu araştırmada Microsoft'un son çıkan işletim sistemi olan Windows 11'in kullanıcılardan veri toplarken kullandıkları yöntemler incelenecektir. Toplanan verilerin türü, verilerin nereye iletildiği, Microsoft'un bu verileri nasıl kullandığı ve konuyla ilgili yasal ve etik hususlar üzerinde yoğunlaşılacaktır. Windows 11'in topladığı veri türü iki ana başlığa ayrılır, gerekli tanılama verileri ve isteğe bağlı tanılama verileri. Zorunlu olarak toplanan gerekli tanılama verileri, Microsoft'un sistem performansı ve kararlılığını gözlemlemesine yardımcı olur. İsteğe bağlı tanılama verileri varsayılan olarak toplanır ve kullanıcı deneyimini geliştirmek adına detaylı hata ve çökme raporları içerir. Bu çalışma, söz konusu bu yöntemlerin kullanıcıların gizliliği ve veri güvenliği üzerindeki etkisini açıklamak, veri toplama ile ilgili yasal mevzuatı incelemek ve teknolojik gelişmeler ile etik anlayış arasındaki dengeyi vurgulamak için yapılmıştır.

Anahtar Sözcükler: Telemetri, gizlilik, işletim sistemi, tanılama verileri, gereksiz programlar

Table of Contents

List of Figures

List of Tables

List of Abbreviations

CPU	Central Processing Unit
CSV	Comma Separated Values
GB	Gigabyte
JSON	JavaScript Object Notation
RAM	Random Access Memory
SHA	Secure Hash Algorithm

Introduction

1.1 What Is Telemetry?

Telemetry, derived from the Greek words *"tele,"* meaning remote and *"metron,"* meaning measure, is a technological process that involves the automated collection, transmission, and analysis of data from remote or inaccessible points to a centralized system. Cambridge Dictionary defines the word *"telemetry"* as:

> *"The science or process of collecting information about objects that are far away and sending the information somewhere electronically."* (2024)

Originally used in industries such as aerospace and meteorology for monitoring and transmitting data from distant locations, telemetry has evolved significantly with the advent of digital technology. Today, it plays a crucial role in various sectors, including healthcare, automotive and most prominently, information technology. In the context of information technology, it means the information collected from user(s) with or without their consent by the developer individual(s) or company (Tan et al., 2021). This data can range from basic diagnostic information about system performance to detailed records of user behavior and interactions with the software (Wickramasinghe, 2023). This information is collected for various reasons, such as to enable developers and companies to monitor system health, improve functionality, optimize user experiences, but with the most frequent reason being to tailor advertising efforts more effectively (Wickramasinghe, 2023). During this process, the data is sold to third party companies looking for potential customers for the product(s) on sale, in full or in part, depending on the requirements of the buyer company. However, the pervasive use of telemetry has raised significant concerns about user privacy, data security, and the ethical implications of such practices. As software developers, including major corporations like Microsoft, increasingly rely on telemetry to gather insights from users, the boundaries between necessary data collection

for system improvement and intrusive data gathering for commercial purposes have become a focal point of debate. Understanding the role of telemetry, particularly within the context of widely used operating systems like Windows 11, is essential for assessing its impact on user privacy and the broader ethical considerations it invokes.

1.2 Types of Telemetry in Windows 11

In the Privacy Statement presented during the installation process of Windows 11, Microsoft's telemetry data collection process is outlined. Microsoft categorizes the telemetry collected from the users under two titles: required diagnostic data and optional diagnostic data and they are explained as such:

- *Required diagnostic data* is used to monitor the overall performance and stability of the OS itself and the programs installed and if they are up-to-date (Watts, 2024). It also helps Microsoft understand how the update process goes. This type of data will be sent to Microsoft whether optional diagnostic data option is toggled or not. Required diagnostic data is the foundational telemetry that Windows 11 collects to ensure the basic functionality, security, and performance of the operating system. This data is collected continuously and automatically, regardless of the user's preferences or settings (Watts, 2024). Microsoft emphasizes that this type of data is essential for maintaining the core operations of the system. It includes information such as device configuration, specifications, system stability and whether the operating system and programs are up-to-date. Additionally it tracks basic system activities, like the success or failure of updates and hardware capabilities, which helps Microsoft identify and address compatibility issues. By monitoring these parameters, Microsoft can promptly detect and resolve issues that may affect the user experience, ensuring that the operating system remains stable and secure. Importantly, even if users opt out of optional diagnostic data, the collection of required diagnostic data cannot be disabled, as it is deemed critical for the ongoing maintenance and improvement of Windows 11.

- *Optional diagnostic data* provides a more detailed and nuanced view of how Windows 11 and the programs installed on it are being used. Optional diagnostic data is used to determine what user and/or process behavior caused a program to crash or give error as it includes more information in the error reports and crash dumps. By default, the optional diagnostic data toggle is on unlike required diagnostic data, users have the option to toggle this data collection on or off, though it is enabled by default. This type of telemetry is more comprehensive and includes information related to app usage, feature engagement and the specific conditions that lead to system errors or crashes (Watts, 2024). For instance, when a program crashes or encounters an error, optional diagnostic data captures detailed error reports, crash dumps and logs that help Microsoft identify the root causes of these issues. This data is invaluable for developers aiming to improve the software's stability and performance, as it offers insights into user behavior and environmental factors that contribute to problems. Additionally, optional diagnostic data can include user interactions with the features of operating system, helping Microsoft to understand how users engage with their products and what areas might require further refinement. However, the comprehensive nature of this data collection raises concerns about user privacy, as it can include sensitive information related to the habits of user and interactions with their device.

These categories help Microsoft maintain and enhance the operating system, while also informing their broader product development strategies. Understanding these distinctions is crucial for evaluating the privacy implications and potential risks associated with telemetry. Beyond these two categories, Microsoft also gathers data related to system usage and feedback, particularly through user-initiated diagnostic reports and feedback tools within Windows 11. While these are not classified under the standard telemetry categories, they further contribute to the company's understanding of user experiences and system performance. Users can manually submit feedback or reports when they encounter issues, which may include screenshots or detailed logs that provide context for the problem (Watts, 2024).

Despite the clear utility of telemetry data in improving software quality and user experience, the breadth and depth of the information collected, especially under optional diagnostic data, have led to ongoing discussions about the balance between innovation and privacy. Users and privacy advocates express concerns that the default settings, which favor extensive data collection, may inadvertently expose personal information or be used for purposes beyond technical improvement, such as targeted advertising or other commercial activities. As such, understanding the specifics of what data is collected, how it is used and the options available to users is critical for informed decision-making and protecting individual privacy in an increasingly data-driven world.

Literature Review

The theoretical foundation for this analysis on Windows 11 telemetry is built upon a comprehensive understanding of telemetry data's in digital forensics and organizational data-driven practices, as explored in two key studies: "Forensic Analysis of the Windows Telemetry for Diagnostics" by Barik et al., 2016 and "The Bones of the System: A Case Study of Logging and Telemetry at Microsoft." by Jauhyeok et al., 2020. These studies provide crucial insights into how telemetry data is collected, processed and utilized both within forensic investigations and organizational contexts, offering a valuable framework to examine the telemetry practices of Windows 11.

2.1 Telemetry in Digital Forensics

The study "Forensic Analysis of the Windows Telemetry for Diagnostics" by Barik et al., (2016) delves into the role of telemetry data within the realm of digital forensics. Telemetry, as defined in this context, is the automated collection and transmission of data from a remote device, which in the case of Windows systems, includes vital information stored in RBS files. The research emphasizes the importance of these files in forensic investigations, highlighting that they store data beyond what is traditionally captured by standard Windows artifacts like the Registry, Prefetch and Event Logs. Unlike these conventional artifacts, which record data only during specific user actions or system events, RBS files contain periodic and continuous data, offering a more comprehensive view of system activities (Barik et al., 2016).

This continuous data collection provides forensic investigators with a more reliable and holistic understanding of a target system's activities. The capacity of RBS files to store information such as hardware serial numbers, external storage connection records and traces of executed processes makes them invaluable for verifying and complementing evidence obtained from other digital artifacts. This study underscores the potential of

Windows telemetry data to enhance forensic analysis, particularly in scenarios where traditional artifacts may not provide sufficient evidence or when certain traces need corroboration.

2.2 Telemetry in Organizational Data-Driven Cultures

The second study, "The Bones of the System: A Case Study of Logging and Telemetry at Microsoft," explores telemetry within the broader context of organizational data practices at Microsoft. This research examines how large organizations leverage event data platforms to support data-driven decision-making across various job roles. Through qualitative interviews and quantitative surveys, the study uncovers the complexities and challenges associated with the widespread use of telemetry and event data within an organization.

One of the key findings of this study is the tension that arises from the interaction of event data between different activities within the organization. Privacy, security, and compliance concerns often clash with the need for data accessibility and usability across different teams and roles (Jauhyeok et al., 2020). The study also highlights the challenges of balancing data retention with the need for proactive system monitoring, as well as the difficulties in identifying the producers and consumers of event data within a large, decentralized organization. These findings illustrate the intricate dynamics at play when telemetry data is integrated into organizational workflows, where the same data may serve multiple, sometimes conflicting, purposes.

2.3 Linking to the Current Study

These two studies provide a critical lens through which to analyze the telemetry practices of Windows 11. The forensic perspective from the first study highlights the depth and potential utility of telemetry data stored in Windows systems, reinforcing the importance of understanding what data is collected and how it can be accessed and analyzed. In contrast, the organizational perspective from the second study underscores the challenges of managing and utilizing telemetry data within large systems, where issues of privacy, data retention and inter-departmental communication become increasingly complex.

In the context of this research on Windows 11, these studies inform our understanding of how telemetry data is collected and used, both from a user privacy standpoint and within broader organizational practices. The forensic implications emphasize the need for transparency and user control over data collection, while the organizational challenges highlight the potential trade-offs between system functionality, data security and user privacy. By situating this study within the theoretical frameworks provided by these two key pieces of research, we can better understand the impact of Windows 11 telemetry on users and organizations and propose strategies for balancing the benefits of telemetry with the ethical considerations of data privacy.

Methodology

To conduct a thorough analysis of the telemetry data collection practices in Windows 11, a structured and systematic approach was employed, focusing on replicating a typical user experience. This methodology ensures that the findings are representative of the data that an average user would generate during normal use of the operating system.

1. Acquisition and Verification of Windows 11 ISO File

The first step involved obtaining the latest version of Windows 11 (version 23H2, English [United States]) from the official Microsoft website. This ensures that the study is based on the most current and widely-used version of the operating system. To verify the integrity and authenticity of the downloaded ISO file, the SHA256 hash of the file was compared against the hash provided by Microsoft on their website. This step is crucial to confirm that the file has not been tampered with and is the exact version released by Microsoft, thereby ensuring the reliability of the subsequent analysis.

2. Selection and Setup of Virtual Environment

For this study, the Windows 11 “Home” edition was selected and installed on a virtual machine (VM) created using Oracle VM VirtualBox. The decision to use the “Home” edition was based on its prevalence among average users, as opposed to the “Pro” and “Education” editions, which are more specialized for professional and educational environments, respectively. The virtual machine was configured with 8 Central Processing Unit (CPU) cores, 8 Gigabyte (GB) of Random Access Memory (RAM) and 128 GB of disk space, specifications that are representative of modern consumer

hardware. This setup ensures that the findings are applicable to a wide range of typical Windows 11 installations.

3. Configuration During Installation

During the installation process, certain steps were carefully managed to mimic the setup of a typical user while maintaining control over variables that could affect the telemetry data collected. A dedicated email address was created specifically for this project, as Windows 11 requires users to sign in with a Microsoft account and there is no officially supported method to bypass this step. This approach ensures consistency and allows for controlled data collection throughout the study. The “Let’s customize your experience” page was skipped using the built-in option, as it primarily serves advertising purposes and is not relevant to the telemetry analysis. Similarly, the “Use your phone from your PC” and “Keep your phone’s photos safe with OneDrive” options were skipped, as these features fall outside the scope of this research. However, the “Always have access to your recent browsing data” option was accepted to explore its impact on telemetry, while offers for Microsoft 365 were declined to avoid introducing variables unrelated to the operating system's core functions.

4. Privacy Settings and Data Collection Tools

To ensure that the telemetry data collected reflects a typical user experience, all privacy settings were left at their default configurations. Altering these settings could potentially influence the volume and type of data collected, thereby skewing the results. The "Diagnostic Data Viewer" option, located under the “Diagnostics & feedback” section, was enabled, as it is an essential tool for monitoring and analyzing the telemetry data sent from the device. This tool allows for real-time tracking of the diagnostic data being collected and provides transparency into what data is being sent to Microsoft.

5. System Updates and Preparation

Once the installation was complete, all pending operating system updates were downloaded and installed via the Windows Update feature and all program updates available through the Microsoft Store were applied. The system was then rebooted to ensure that the analysis was conducted on the most up-to-date version of Windows 11, as updates often include changes to telemetry settings and data collection practices. This step is vital for ensuring that the findings of study are current and applicable to users who keep their systems updated.

6. Execution of Pre-installed Software

A specialized PowerShell script (which can also be found on https://github.com/onerc/scripts/tree/main/w11) was used to execute all pre-installed software listed in the Start menu. Each application was run for 15 seconds before being closed, simulating a basic interaction with these programs. This process was included because pre-installed software is an integral part of the operating system and interactions with these programs may generate additional telemetry data. The use of a standardized script ensures that each application is treated uniformly, allowing for consistent and replicable results.

7. Continuous Internet Connection and Account Log-in

Throughout the duration of the project, the internet connection was kept online and the Microsoft account used during the installation remained logged in. This decision was made to simulate a continuous and uninterrupted user experience, as maintaining an active internet connection is crucial for the real-time transmission of telemetry data to Microsoft's servers. Additionally, keeping the account logged in ensures that all telemetry processes, which may require a persistent connection to the user's account, are fully

operational. This approach allows for a comprehensive analysis of the telemetry data generated under typical usage conditions, reflecting what an average user would experience.

Live captions	Microsoft Store
Magnifier	Microsoft To Do
Narrator	Movies & TV
On-Screen Keyboard	News
Voice access	Notepad
Windows Speech Recognition	OneDrive
Calculator	Outlook (new)
Calendar	Paint
Camera	Phone Link
Clock	Photos
Copilot	Quick Assist
Cortana	Settings
Family	Snipping Tool
Feedback Hub	Solitaire & Casual Games
File Explorer	Sound Recorder
Game Bar	Sticky Notes
Get Help	Terminal
Get Started	Tips
Mail	Weather
Maps	Windows Backup
Media Player	Windows Security
Microsoft 365 (Office)	Windows Tools
Microsoft Clipchamp	Xbox
Microsoft Edge	

Table 3.1: List of pre-installed software

This methodology, encompassing the careful setup of a virtual machine environment, the controlled configuration of system settings and the systematic execution of software, provides a comprehensive framework for analyzing Windows 11's telemetry practices. By mimicking a typical user experience while maintaining rigorous control over variables, this study aims to produce findings that are both accurate and broadly applicable to Windows 11 users.

Findings

Upon booting the Windows 11 operating system, it was observed that several pre-installed programs and processes automatically start running in the background. This automatic startup behavior is a default feature of Windows 11, designed to ensure that the system is ready to provide a seamless user experience right from the moment the operating system is launched. As they are closed source/proprietary software, researching and reverse engineering process is harder and more time consuming, if not impossible. On the contrary, some of these programs are named self-explanatorily and individuals interested in this area can decode these names and take educated guesses about the purpose(s) of these programs. This lack of transparency can be attributed to the proprietary nature of the software, which prevents users from easily understanding or modifying the processes involved.

Microsoft includes various background processes in Windows 11 to manage essential system functions, such as security updates, system health monitoring, and the facilitation of cloud services. For instance, processes related to Windows Defender and Windows Update are critical for maintaining system security and performance. However, other processes may be less apparent in their purpose, leaving users to speculate on their functions based on process names alone. Some processes, like "SearchIndexer.exe," suggest their purpose through their names, indicating that they handle tasks such as indexing files for faster search results.

However, many of these processes are not as transparent. Because these programs are proprietary, the community of users, researchers, and developers face significant challenges in understanding the full extent of what these processes do. Unlike open-source software, where code is publicly available for inspection, modification and improvement, proprietary software hides its internal workings. This makes it difficult, if not impossible,

to fully reverse-engineer or analyze these processes without violating legal restrictions or requiring substantial time and resources.

Moreover, the automatic startup of these processes raises questions about resource utilization and potential privacy concerns. Users might be unaware that certain programs are running in the background, potentially consuming system resources or transmitting telemetry data to Microsoft without explicit consent. For example, processes related to telemetry may collect diagnostic data or user activity information to be sent back to Microsoft for analysis. While some of this data collection is necessary for maintaining system stability and security, the opacity of these processes means users have limited control or understanding of what data is being collected and for what purposes.

AggregatorHost	RuntimeBroker
audiodg	SearchHost
backgroundTaskHost	SearchIndexer
conhost	SecurityHealthService
csrss	SecurityHealthSystray
ctfmon	services
dllhost	ShellExperienceHost
dwm	sihost
explorer	smartscreen
fontdrvhost	smss
gamingservices	spoolsv
gamingservicesnet	SppExtComObj
identity_helper	sppsvc
Idle	StartMenuExperienceHost
LocationNotificationWindows	svchost
lsass	System
Memory Compression	taskhostw
MicrosoftEdgeUpdate	upfc
MpDefenderCoreService	userinit
msedge	Widgets
MsMpEng	WidgetService
NisSrv	WindowsTerminal
OneDrive	wininit
PhoneExperienceHost	winlogon
Registry	WmiPrvSE

Table 4.1: List of processes running in the background

4.1 Background Processes Table

To provide a clearer picture of the processes observed during startup, a detailed table has been compiled listing the names of these processes, their associated services and any identifiable functions. This table serves as a foundation for understanding the complex web of activities that occur behind the scenes in Windows 11. While some processes may have self-explanatory names that give hints about their roles (e.g., "ShellExperienceHost.exe" likely relates to the user interface and shell experience), others may require further investigation to determine their exact purpose.

The second finding is on telemetry data collection and export. During interactions with programs on Windows 11, it was observed that the operating system generates a Comma-Separated Values (CSV) file containing event data, including timestamps, event titles and associated telemetry information formatted in JavaScript Object Notation (JSON). This data is collected to monitor system performance, track user interactions and improve the overall functionality of the operating system. The generated CSV files, which can be accessed using the "Diagnostic Data Viewer" application available in the Windows Store, provide a straightforward way to view and export this telemetry data.

The process of creating and storing telemetry data in CSV format allows for a detailed and organized record of events and associated data. Each CSV file entry includes critical information such as the event timestamp, a descriptive event title, and the telemetry data in JSON format. This JSON data contains structured information related to the specific event, including details about the system state and user interactions at the time of the event. The use of CSV for exporting this data means that it is stored in a plain-text format, making it easily accessible and readable. Users can open these files with any text editor to review the raw telemetry data.

One significant aspect of this finding is that the telemetry data stored in these CSV files is not obfuscated or encrypted. This lack of encryption means that once the data is exported, it is readily available in an unaltered form, which could pose potential privacy concerns. The open nature of the data allows anyone with access to the exported files to view or analyze the information contained within them without requiring specialized tools or knowledge. This transparency facilitates user access to their own data but also raises questions about the security of sensitive information, as unencrypted data can be more vulnerable to unauthorized access or misuse.

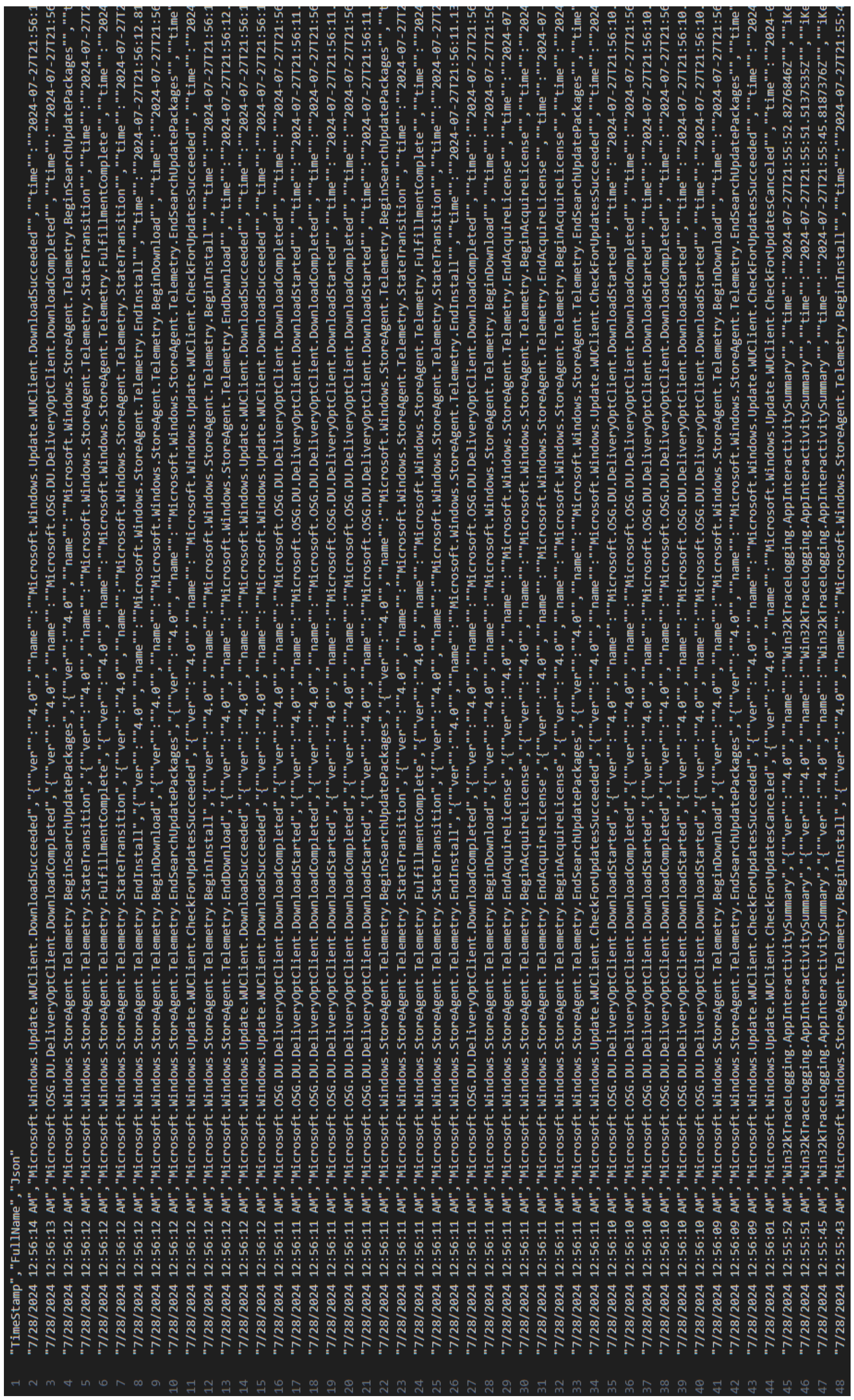

Figure 4.1 A section of CSV file

The aforementioned JSON data includes but not limited to:

Data	Explanation of Microsoft
"MonitorWidth": 1024	Number of horizontal pixels in the application host monitor resolution
"MonitorHeight": 768	Number of vertical pixels in the application host monitor resolution
"WindowWidth": 768	Number of horizontal pixels in the application window
"WindowHeight": 519	Number of vertical pixels in the application window
"bootId": 8	Total boot count since the operating system was installed
"KeyboardInputSec": 0	Total number of seconds during which there was keyboard input
"MouseInputSec": 1	Total number of seconds during which there was mouse input
"TouchInputSec": 0	Total number of seconds during which there was touch input
"InFocusDurationMS": 9735	Total time (in milliseconds) the application had focus
"AppVersion": 11.2405.13.0_x64_!2024/06/07: 17: 18: 34!199187!notepad.exe	Version of the application that produced this event

Table 4.2: A section of JSON file along with their corresponding explanations

Discussion

As it is clearly against their business model and as they have implicitly stated in their privacy statement, Microsoft does not allow their users to completely disable data collection.

Microsoft's data collection practices, as detailed in their privacy statements and implemented in Windows 11, reflect a business model that prioritizes continuous data collection as a fundamental aspect of their service offering. Despite user concerns about privacy, Microsoft does not provide an option to completely disable telemetry, aligning with their business objectives of improving system performance and user experience through data-driven insights. This approach, while beneficial for system optimization and support, raises significant privacy concerns, particularly for users who are sensitive to the extent of data collected.

Previous personal attempts of uninstalling these programs software containing telemetry and bloatware ended up with an OS that has multiple visually and functionally critical issues rendering it unusable. Removing or disabling these components often leads to visual and functional issues within the operating system, rendering it unstable or unusable. This unintended consequence underscores the tight integration of telemetry and system components within Windows 11, which complicates efforts to manage or limit data collection without adversely affecting system functionality.

Furthermore, even for users under the age of 13, once their parents or guardian gives consent, their account is treated like any other account (Microsoft Privacy Statement, 2024), meaning they will be subjected to the same telemetry policies someone at the age of consent may consider privacy-violating.

> *"Once parental consent or authorization is granted, the child's account is treated much like any other account.* (Microsoft Privacy Statement, 2024)*"*

Once consent is granted, child accounts are subject to the same telemetry practices as adult accounts, potentially exposing younger users to privacy concerns that might be considered inappropriate by their guardians. This broad application of telemetry policies reflects Microsoft's commitment to a standardized data collection approach but also highlights potential ethical dilemmas regarding the extent of data collected from minors.

Even programs designed to assist users with disabilities, such as "Narrator," are subject to these telemetry policies, emphasizing that Microsoft's data collection practices are pervasive across all functionalities and user groups. This uniform application of telemetry raises questions about the balance between providing useful features and respecting user privacy.


```
""app"": {
   ""id"": ""W:
0000f519feec486de87ed73cb92d3cac802400000000!0000227ff82d17f77267
9a9048e08a60f732f0a01dd6!narrator.exe"",
   ""ver"": ""2053/12/13: 22: 13: 35!9EDFA!narrator.exe"",
   ""is1P"": 5,
   ""asId"": 261
}
```


Table 5.1: A section of JSON file containing data about "Narrator" program

In response to these privacy concerns, some users seek alternatives to mitigate or bypass Microsoft's data collection practices. These methods include:

Third-Party Software: Users may turn to third-party programs that claim to disable or reduce telemetry. However, the effectiveness and security of these tools can vary, and relying on them may introduce additional risks or complexities.

Modified ISO Files: Some individuals seek out modified versions of Windows 11 ISO files that purportedly omit telemetry and bloatware. While these modified versions may offer a reduction in data collection, they also carry significant risks, including potential security vulnerabilities and lack of official support.

Free and Open-Source Operating Systems: Alternatives such as GNU/Linux or BSD distributions offer the advantage of open-source code, which allows users to audit and verify the privacy and security of their operating systems. These free software options can provide greater transparency and control over data collection, though they may require a learning curve and adaptation from users accustomed to proprietary systems.

Overall, the persistence of telemetry in Windows 11 and its broad application across different user accounts and functionalities reflect Microsoft's stance on data collection as a core component of their operating system. This approach, while beneficial for system improvement and support, continues to provoke debate over privacy and user control. As users seek ways to manage their data privacy, the trade-offs between system functionality, security, and personal privacy remain central to the ongoing discussion about modern operating systems and their data practices.

Limitations

This analysis of Windows 11's telemetry practices presents several limitations that may affect the comprehensiveness and applicability of the findings:

The study focused primarily on observable telemetry practices and background processes. It did not encompass all potential data collection mechanisms or proprietary technologies employed by Microsoft. As a result, there may be additional forms of data collection or data types that were not identified or analyzed in this study. The proprietary nature of many of the background processes and telemetry components limited the ability to fully investigate their functionality. Without access to the source code or detailed technical documentation, understanding the complete range of data collected and its processing methods was challenging. This limitation constrained the depth of analysis and the ability to offer a comprehensive evaluation of all telemetry practices.

The use of a virtual machine for testing Windows 11, while practical for controlled experimentation, may not perfectly replicate the behavior of the operating system on physical hardware. Differences in system performance, resource management, and hardware interactions could potentially influence the results, making them less representative of a typical user experience on physical machines. The analysis of third-party tools and modified ISO files for bypassing or disabling telemetry was limited to available options and their claims. The effectiveness and security of these tools were not thoroughly tested or verified within the scope of this study, which could affect the reliability of recommendations regarding alternative solutions. The study assumed a standard user interaction model and did not account for variations in individual user behaviors or configurations. Differences in usage patterns, system settings, and application interactions could influence telemetry data collection and system performance, potentially affecting the generalizability of the findings.

The analysis was constrained by legal and ethical considerations regarding reverse engineering and data access. These constraints limited the ability to conduct a more in-depth examination of the telemetry components and their interactions with the operating system. As Windows 11 and its telemetry practices are subject to updates and changes over time, the findings of this study may become outdated. Future updates to the operating system could alter telemetry practices, background processes, or data handling methods, which may not be reflected in the current analysis.

These limitations suggest that while the study provides valuable insights into Windows 11's telemetry practices, it is important to consider these factors when interpreting the results and to recognize the potential need for further research to address these constraints.

Conclusion

This analysis of Windows 11's telemetry practices has provided a detailed examination of the data collection mechanisms of the operating system and their implications for user privacy and system performance. The findings reveal that Microsoft employs a comprehensive approach to telemetry, with automatic background processes and detailed data collection mechanisms designed to enhance system functionality and user experience. However, these practices also raise significant concerns regarding transparency, privacy and control.

The analysis highlighted two primary aspects: the automatic initiation of background processes and the generation of telemetry data files. The background processes, while essential for system stability and performance, are proprietary and lack detailed documentation, limiting user understanding and control. The telemetry data, collected and exported in accessible CSV and JSON formats, provides transparency but also poses privacy risks due to the lack of encryption.

The study also discussed the broader implications of Microsoft's telemetry policies, including their uniform application across all user accounts, including those of minors, and the impact on users seeking to bypass or mitigate data collection through third-party tools or alternative operating systems. These observations underscore the tension between maintaining a high-performing system and addressing user concerns about data privacy and security.

While the analysis offers valuable insights into the nature and scope of Windows 11's telemetry practices, it is subject to limitations such as the proprietary nature of software components, the use of virtual machines and the evolving nature of the operating system. Future research and continued scrutiny are necessary to address these limitations and to provide a more comprehensive understanding of the implications of telemetry practices.

In conclusion, the findings reflect the complexity of balancing system functionality with privacy considerations in modern operating systems. As Windows 11 and its telemetry practices continue to evolve, ongoing dialogue and research will be essential to ensure that user privacy is adequately protected while maintaining system performance and usability.

This analysis is crucial because it sheds light on the intricate balance between system functionality and user privacy in Windows 11. By providing a detailed examination of the operating system's telemetry practices, this study offers valuable insights into how user data is collected, processed and stored. Understanding these practices enables users to make informed decisions about their privacy and take appropriate measures to manage their data. For researchers, this analysis contributes to the broader field of data privacy and system security, highlighting the need for greater transparency and user control. It also provides a foundation for further investigation into the evolving practices of data collection and the development of solutions to enhance user privacy in modern operating systems. Overall, this analysis serves as a critical resource for both users and researchers seeking to navigate and address the complexities of telemetry and data privacy in contemporary technology.

References

(2023, February 17). Diagnostic Data Viewer Overview. Microsoft. https://learn.microsoft.com/en-us/windows/privacy/diagnostic-data-viewer-overview

(2023, October 05). Telemetry 101: An Introduction to Telemetry | Splunk. (n.d.). Splunk. https://www.splunk.com/en_us/blog/learn/what-is-telemetry.html

(2024, February 29). Required diagnostic events and fields for Windows 11, versions 23H2 and 22H2. Microsoft. https://learn.microsoft.com/en-us/windows/privacy/required-diagnostic-events-fields-windows-11-22h2

(2024, June). Microsoft Privacy Statement. Microsoft. https://privacy.microsoft.com/en-US/privacystatement

Barik, T., DeLine, R., Drucker, S., & Fisher, D. (2016). The bones of the system: A case study of logging and telemetry at Microsoft. 2016 IEEE/ACM 38th International Conference on Software Engineering Companion (ICSE-C), 92-101.

Han, J., Park, J., Chung, H., & Lee, S. (2020). Forensic analysis of the Windows telemetry for diagnostics. Digital Investigation, 32, 200900.

Tan, L., Su, W., Zhang, W., Lv, J., Zhang, Z., Miao, J., … Li, N. (2021). In-band Network Telemetry: A Survey. Computer Networks, 186, 107763.doi:10.1016/j.comnet.2020.107763

Watts, L., & Watts, L. (2024, July 2). Telemetry Data: Examples & Types of data Collected. Teramind Blog | Content for Business. https://www.teramind.co/blog/telemetry-data-examples-types/

Appendix

"7/28/2024 12:53:06 AM",

"Win32kTraceLogging.AppInteractivitySummary","

{

""ver"": ""4.0"",

""name"": ""Win32kTraceLogging.AppInteractivitySummary"",

""time"": ""2024-07-27T21: 53: 06.7304062Z"",

""iKey"": ""o: 0a89d516ae714e01ae89c96d185e9ae3"",

""ext"": {

""utc"": {

""shellId"": 281572472535515136,

""eventFlags"": 258,

""pgName"": ""WIN"",

""stId"": ""82522D31-6258-4372-8C33-557A613B606A"",

""flags"": 469762661,

""epoch"": ""805140"",

""seq"": 539

},

""privacy"": {

""dataType"": 33554432,

""isRequired"": false,

""dataCategory"": 1,

""product"": 1

},

```
 ""metadata"": {
        ""f"": {
          ""PartATransform_AppSessionGuidToUserSid"": 8,
          ""AppSessionId"": 8,
          ""AggregationStartTime"": 9,
          ""ViewFlags"": 5,
          ""EventSequence"": 5
        },
        ""privTags"": 33554432,
        ""policies"": 0
      },
      ""app"": {
        ""id"": ""U:Microsoft.Paint_11.2404.1020.0_x64__8wekyb3d8bbwe!App"",
        ""ver"": ""11.2404.1020.0_x64_!2024/06/21: 13: 44: 37!0!mspaint.exe"",
        ""is1P"": 5,
        ""asId"": 497
      },
      ""os"": {
        ""bootId"": 8,
        ""name"": ""Windows"",
        ""ver"": ""10.0.22631.3958.amd64fre.ni_release.220506-1250"",
              ""expId"": ""MD: 283BAEF,ME: 2E3922F,FX: 13023E93,ME: 2E30985,FX:
13127526,ME: 2E30A2A,FX: 1316017C,MD: 3036FD6""
      },
      ""device"": {
        ""localId"": ""s: 511919E3-DD97-42E5-9103-A9E32CCEA92C"",
        ""deviceClass"": ""Windows.Desktop""
      },
```

```
""protocol"": {
        ""devMake"": ""innotek GmbH"",
        ""devModel"": ""VirtualBox"",
        ""ticketKeys"": [
          ""2262550""
        ]
      },
      ""user"": {
        ""localId"": ""m:cf8ae4213a2e6dbd""
      },
      ""loc"": {
        ""tz"": ""+03: 00""
      }
    },
    ""data"": {
      ""AppId"": ""U:Microsoft.Paint_11.2404.1020.0_x64__8wekyb3d8bbwe!App"",
      ""AppVersion"": ""11.2404.1020.0_x64_!2024/06/21: 13: 44: 37!0!mspaint.exe"",
      ""CommandLineHash"": 2463376551,
      ""AppSessionId"": ""00002974-0001-0008-3DC3-10576FE0DA01"",
      ""AggregationStartTime"": ""2024-07-27T21: 46: 16.4398625Z"",
      ""AggregationDurationMS"": 410282,
      ""InFocusDurationMS"": 9625,
      ""FocusLostCount"": 1,
      ""NewProcessCount"": 1,
      ""UserActiveDurationMS"": 9625,
      ""UserOrDisplayActiveDurationMS"": 9625,
      ""UserActiveTransitionCount"": 0,
      ""InFocusBitmap"": ""0x0000100000000000"",
      ""InputSec"": 7,
      ""KeyboardInputSec"": 0,
      ""SipKeyboardInputSec"": 0,
      ""InjectedKeyboardInputSec"": 0,
      ""MouseInputSec"": 7,
      ""InjectedMouseInputSec"": 0,
```

```
      ""TouchInputSec"": 0,
      ""InjectedTouchInputSec"": 0,
      ""PenInputSec"": 0,
      ""InjectedPenInputSec"": 0,
      ""ExtPenInputSec"": 0,
      ""PrecisionTouchpadInputSec"": 0,
      ""InjectedPrecisionTouchpadInputSec"": 0,
      ""HidInputSec"": 0,
      ""WindowWidth"": 819,
      ""WindowHeight"": 614,
      ""MonitorWidth"": 1024,
      ""MonitorHeight"": 768,
      ""MonitorFlags"": 0,
      ""WindowFlags"": 16,
      ""InteractiveTimeoutPeriodMS"": 60000,
      ""AggregationPeriodMS"": 1200000,
      ""BitPeriodMS"": 20000,
      ""AggregationFlags"": 49,
      ""SummaryRound"": 0,
      ""SpeechRecognitionSec"": 0,
      ""GameInputSec"": 0,
      ""TargetAsId"": 497,
      ""CompositionRenderedSec"": 9,
      ""CompositionDirtyGeneratedSec"": 8,
      ""CompositionDirtyPropagatedSec"": 8,
      ""BackgroundMouseSec"": 0,
      ""AudioInMS"": 0,
      ""AudioOutMS"": 0,
      ""ViewFlags"": 0,
      ""SinceFirstInteractivityMS"": 9641,
      ""EventSequence"": 46
    }
}
```