

“Should I Trust You?” Examining VR User’s Safety and Privacy Concerns

Shreyaan Nath, Texas A&M University

Motivation & Goals

Rapid development & acceptance of VR technology into people's lives since 2020

A near “monopolization” of the VR subspace by Meta with nearly 77% of the VR marketshare [1]

Objective:

Determine how users feel about the rise of VR developments and what are their perceptions about security and privacy.

Methodology

1. Identify All Subreddits & Keywords:

- r/virtualreality - privacy
- r/oculusquest - privacy
- r/privacy - VR

2. Implement a Reddit API into a Python Program for Data Extraction

3. Use Python’s NLTK Library and Sentence Transformer to Perform Sentiment Analysis & Identify Key Themes

4. Analyze All Collected Variables & Create Presentation



Findings

14% or 4.5k of Collected Comments Contained Opinions Related to VR Security & Privacy

Theme	# Comments	Percentage
Company Reputation	1579	20.57%
Future of VR Safety	1389	18.09%
Company Responsibility & Ethics	1269	16.53%
Report Issues	1084	14.12%
Legal & Compliance	969	12.62%
Privacy & Safety Advice	626	8.15%
Privacy Policy	180	2.34%
Surveillance	143	1.86%
Privacy Resignation	110	1.43%
Biometric Tracking	82	1.07%
Digital Safety	77	1.00%
Physical Tracking	57	0.74%
Profiling & Targeting	38	0.49%
Microphone & Audio	34	0.44%
Data Leakage	24	0.31%
Data Handling	16	0.21%

“It’s true that people overreact to things quite frequently these days and make things out to be worse than they are. However, there are a lot of things to be seriously concerned about regarding Facebook and how aggressive they are in their take over of the VR space.”

• A significant number of users have expressed concerns over Meta having a history of **privacy violations** and **prior data leaks**.

• However, many users also note that Meta’s influence will only continue to grow due to the **affordability** of its headsets in comparison to competitors.

“Privacy in VR is dead, Facebook owns the industry. Years of VR journalists and redditors and YouTubers insisting nothing was wrong brought us here. And will bring us right into the eye tracking and ads.”

Skills Learned

- Integrating a Reddit API into Python for Data Extraction.
- Integrating a sentiment analysis library (NLTK) into Python to categorize extracted data into significant themes relevant to VR.
- The differences between the NLTK Library & LLM API’s

Challenges

- Rate limits set in place by Reddit’s API (slowed data collection)
- Overcoming the biases of Python’s NLTK Library (overly positive assortment of comments)
- Extensive system resource requirements hindering the amount of data we could readily work upon.

Next Steps

- Expand sources to include more forums and subreddits (e.g., Meta community forums, VR-focused boards).
- Refine keyword and theme taxonomy. Use LLMs (e.g., ChatGPT) to validate and enrich labels.
- Cross-check user concerns against official privacy policies.

References

- [1]. Counterpoint Technology Market Research, <https://www.counterpointresearch.com/insight/global-xr-arvr-headsets-market-2024/>