Berkeley Open Extended Reality Recordings 2023 (BOXRR-23): 4.7 Million Motion Capture Recordings from 105,852 Extended Reality Device Users

Vivek Nair
UC Berkeley
Berkeley, CA 94720
vcn@berkeley.edu

Wenbo Guo
UC Berkeley
Berkeley, CA 94720
henrygwb@berkeley.edu

Rui Wang
UC Berkeley
Berkeley, CA 94720
ruiwang813@berkeley.edu

James F. O'Brien
UC Berkeley
Berkeley, CA 94720
job@berkeley.edu

Louis Rosenberg
Unanimous AI
Pismo Beach, CA 93448
louis@unanimous.ai

Dawn Song
UC Berkeley
Berkeley, CA 94720
dawnsong@berkeley.edu

Abstract

Low-cost extended reality (XR) devices such as the Meta Quest 2 have seen a recent surge in popularity, with motion tracking "telemetry" data lying at the core of nearly all XR and metaverse experiences. Researchers are just beginning to understand the implications of this data for security, privacy, usability, and more, but currently lack large-scale human motion datasets to study. The BOXRR-23 dataset contains 4,717,215 motion capture recordings, voluntarily submitted by 105,852 XR device users from over 50 countries. BOXRR-23 is over 200 times larger than the largest existing motion capture research dataset and uses a new, highly-efficient and purpose-built XR Open Recording (XROR) file format.

1 Introduction

For decades, human motion capture (MoCap) recordings have been an important resource in a variety of fields, ranging from animation and computer-generated imagery (CGI) to authentication and human-computer interaction (HCI). Recently, the proliferation of extended reality (XR) devices has created a prominent new application for this data, with motion data being central to almost all XR and "metaverse" experiences. Since 2002, at least 25 motion capture datasets have been created based on laboratory studies of up to a few hundred users to facilitate research in this important domain.

An emerging area of interest for security and privacy researchers is the passive identification and authentication of XR users based on their movement patterns. Until recently, XR identification and authentication studies have been limited to a few hundred users due to the lack of large-scale human motion datasets. By contrast, studies involving traditional biometrics, such as fingerprints or facial recognition, typically use datasets involving 100,000 or more subjects [3].

In this paper, we introduce the BOXRR-23 dataset, which contains 4,717,215 motion capture recordings uploaded by 105,852 XR device users from over 50 countries. In addition to being more diverse and ecologically valid than laboratory studies, BOXRR-23 is over 200 times larger than the largest known public motion capture dataset. We recently used this dataset, for the first time, to

Submitted to the 37th Conference on Neural Information Processing Systems (NeurIPS 2023) Track on Datasets and Benchmarks. Do not distribute.

demonstrate that XR motion data provides a biometric signal on par with fingerprints [32]. However, we envision the potential uses of this data may go far beyond security and privacy to include areas such as motion synthesis, human-computer interaction, and theoretical machine learning research.

In addition to assembling this dataset from three public sources and enriching it with additional metadata, we developed a new "Extended Reality Open Recording" (XROR) file format due to the lack of an existing standard format suitable for this use case. The XROR format is about 30% more space efficient than the original file formats, without loss of precision.

To help interested researchers evaluate this dataset, we provide documentation pursuant to a number of open standards, including Datasheets for Datasets [17] and Dataset Nutrition Labels [19]. Furthermore, we conducted a large-scale survey ($N = 1,006$) of the users contained in this dataset to better understand their demographics, the results of which are summarized herein.

2 Background

Since the 1990s, computerized motion tracking systems have been used for animation and CGI in a large number of popular movies, television series, and video games. A typical commercial motion capture solution uses optical tracking or inertial measurement units (IMUs) to measure the location of various body parts, with prices ranging from $10,000 to over $250,000 for a full-body tracking system. Conventional motion capture datasets have involved expensive laboratory studies with up to 300 subjects paid to perform a variety of tasks while wearing a professional motion capture setup.

Motion capture data is also central to the operation of extended reality (XR) systems, which include devices supporting augmented reality (AR), virtual reality (VR), and mixed reality (MR) technologies. XR has experienced a recent surge in popularity with the release of affordable self-contained VR devices like the Meta Quest 2. Most consumer-oriented virtual reality systems include a head-mounted display (HMD) and two hand-held controllers. The system uses either external or onboard sensors to measure the position and orientation of these devices in 3D space, providing six degrees of freedom (6DoF), captured at a rate of between 60 and 144 times per second. In essence, XR devices have recently become an affordable and widely-adopted form of motion tracking system.

The motion data generated by an XR device is used by a client-side application, such as "Beat Saber" or "Tilt Brush," to render auditory, visual, and haptic stimuli, creating an immersive 3D experience. In some cases, users capture and share recordings of the motion data generated during an XR usage session to allow other users to "replay" the same virtual experience.

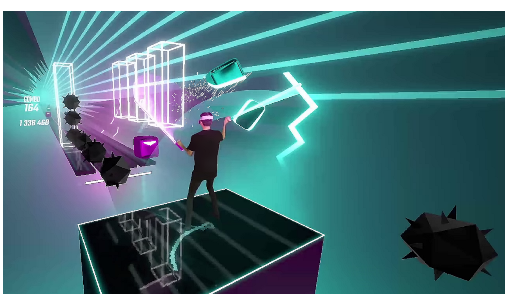

Figure 1: "Beat Saber" – VR rhythm game.

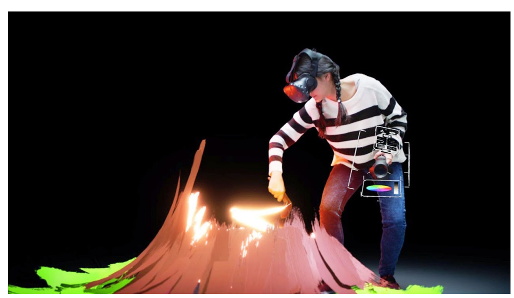

Figure 2: "Tilt Brush" – VR painting app.

2.1 Beat Saber

"Beat Saber" [16], shown in Figure 1, is a VR rhythm game where players slice blocks representing musical beats with a pair of sabers they hold in each hand. With over 6 million copies sold, Beat Saber is the most popular VR application of all time [43]. The game contains a number of "maps," which consist of an audio track and a series of objects presented to the user in time with the audio. These objects include "blocks," which the player must hit at the correct angle with the correct saber, "bombs," which the player must avoid hitting with their sabers, and "walls," which the player must avoid with their head. The player is given a score based on their accuracy in completing these tasks. Reacting to these events typically requires users to deploy fast ballistic movements [42, 14].

While hundreds of maps are included in the base game, over 100,000 user-created maps can be played by installing open-source game modifications. Beat Saber enthusiasts may choose to install open-source leaderboard extensions in order to compete with other players to achieve a higher "rank" on the leaderboards for popular maps. Two of the most popular Beat Saber leaderboard services are "BeatLeader" [36] and "ScoreSaber" [7], with a combined 4 million scores being submitted to the platforms to date. When submitting a score to either of these services, users attach a motion capture recording of them playing the corresponding Beat Saber map, which is then made publicly available on the BeatLeader or ScoreSaber website to allow others to audit the legitimacy of the claimed score.

2.2 Tilt Brush

"Tilt Brush" [10], shown in Figure 2, is a VR painting game created by Google that allows users to create 3D virtual objects using a variety of brushes and tools. Users can then export their drawings in various file formats, along with a motion capture recording of them creating the object, allowing other users to re-watch the original painting process. From 2017 to 2021, Google hosted "Google Poly," a free service for sharing virtual creations (and accompanying motion capture recordings) from Tilt Brush. After the shutdown of Google Poly in 2021, the "PolyGone" project [6] was created to host a free archive of over 50,000 user-submitted creations from Google Poly under a CC-BY license. Contrary to Beat Saber, Tilt Brush motion consists primarily of precise fine motor movements.

3 Data Collection

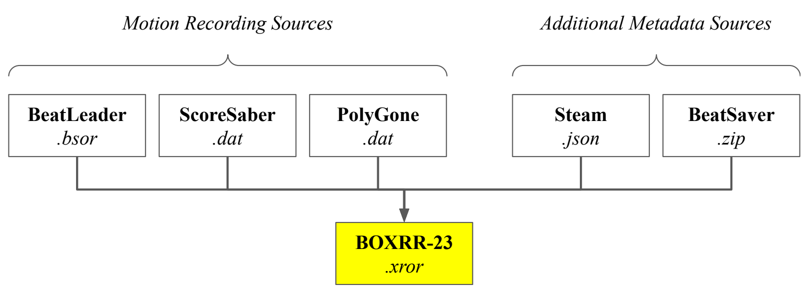

Figure 3: Data collection and processing pipeline for BOXRR-23 dataset.

Figure 3 shows the data collection process used to produce the BOXRR-23 dataset. We downloaded over 4.7 million publicly-available motion capture recordings stored on the BeatLeader, ScoreSaber, and PolyGone websites, and obtained additional metadata information, such as player experience levels and in-game events, from the public web APIs of Steam [9] and BeatSaver [2]. We then removed identifiable details like player IDs and pseudonyms to protect the identity of each user. Finally, we converted all recordings from their original formats into our purpose-built XROR format, described in §5. The sizes of each of the sources, and of the dataset, are summarized in Table 1. We performed this data collection process in April 2023 and have included all valid, non-corrupt recordings submitted to all three platforms between November 1st, 2017 and April 15th, 2023.

Table 1(A): Sources for data in BOXRR-23 dataset.

Source	**Application**	**Users**	**Recordings**	**Format**	**Size**
BeatLeader	Beat Saber	95,192	3,525,456	.bsor	6.25 TB
ScoreSaber	Beat Saber	55,331	1,136,581	.dat	1.44 TB
PolyGone	Tilt Brush	27,693	55,178	.tilt	1.87 TB

Table 1(B): Output characteristics of BOXRR-23 dataset.

Dataset	**Users**	**Recordings**	**Format**	**Size**
BOXRR-23 Dataset	105,852	4,717,215	.xror	4.71 TB

4 Related Work

We searched for existing datasets relating to “motion capture,” “telemetry,” “VR motion,” “XR motion,” etc., on dataset hosting platforms like Kaggle, Zenodo, and Dryad, as well as for academic papers relating to motion capture data and experiments. We found over 25 existing datasets containing human motion recordings. The majority of these datasets come from conventional non-XR motion tracking systems, as listed in Table 2(A), while several originate from XR-based laboratory studies, listed in Table 2(B). The largest existing study contained 511 subjects [27], with a single session captured from each subject. By contrast, our dataset, summarized in Table 2(C), contains over 105,000 subjects and 4.7 million recordings from the three sources described in §3.

In addition to being over 200 times larger than the largest existing dataset, we found that all of the existing datasets come from a laboratory study in which participants used a small number of homogeneous devices and were generally physically present in a narrow geographical area. Thus, the BOXRR-23 dataset is more useful for obtaining a representative sample of XR users, as it originates from real XR users using their own devices in their own homes. As a result, it contains diverse data from over 40 types of XR devices, and includes users from over 50 countries around the world.

Table 2(A): Current motion capture datasets outside XR.

Dataset	Organization	Year	Subjects	Recordings
BMLrub [40]	Ruhr University Bochum	2002	111	3,061
HDM05 [30]	Max Planck Society	2007	4	215
CMU-MMAC [22]	Carnegie Mellon University	2008	5	5
EYES Japan [5]	EYES Japan	2009	12	750
HumanEva [38]	University of Toronto	2010	3	28
SFU MoCap [8]	Simon Fraser University	2012	7	44
ACCAD [1]	Ohio State University	2012	20	252
Sleight of Hand [20]	Trinity College Dublin	2012	1	62
Human3.6m [21]	Romanian Academy	2013	11	44
MoSh [24]	Max Planck Society	2014	19	77
MPI Limits [12]	Max Planck Society	2015	3	35
KIT MoCap [26]	Karlsruhe Institute of Technology	2016	232	2,925
Total Capture [41]	University of Surrey	2017	5	37
AMASS [25]	Max Planck Society	2019	300	11,000
CMU MoCap [4]	Carnegie Mellon University	2019	144	2,605
MoVi [18]	Queen’s University	2021	90	1,890

Table 2(B): Current motion capture datasets inside XR.

Dataset	Organization	Year	Subjects	Recordings
Behavioural Biometrics [34]	Bundeswehr University Munich	2019	22	88
TTI [27]	Stanford University	2020	511	511
Body Normalization [23]	University of Duisburg-Essen	2021	16	48
Obfuscation [29]	University of Central Florida	2021	60	120
Body Sway [13]	Purdue University	2021	28	336
You Can’t Hide [39]	University of Padova	2022	35	69
Motion Matching [35]	Technical University of Catalonia	2022	1	12
Personal Identifiability [28]	Stanford University	2023	232	1856
Who is Alyx [37]	University of Würzburg	2023	71	142

Table 2(C): Our new XR motion capture dataset.

Dataset	Organization	Year	Subjects	Recordings
BOXRR-23	University of California, Berkeley	2023	105,852	4,717,215

5 XROR Format

As detailed in §3, the data included in the BOXRR-23 dataset was scraped from three separate sources (BeatLeader, ScoreSaber, and PolyGone), each using three separate custom file formats designed specifically for those platforms (.BSOR, .DAT, and .TILT, respectively, summarized in Table 3(A)). We felt that the experience of future consumers of this dataset would be improved if the recordings were all converted to a single file format that could be analyzed and ingested via a unified pipeline.

We began by evaluating open-source motion capture file formats such as .BVA, .BVH, and .MVNX. Unfortunately, we found that the existing formats were unsuitable for this database for a variety of reasons. Some formats, such as .BVA and .BVH, only have support for motion data, and did not allow us to embed the rich metadata and event data streams we wished to include in the dataset. Others, like .MVNX, did support the inclusion of arbitrary metadata and event data streams, but used an inefficient underlying text-based file format (.XML) that would have caused the dataset to balloon to over 300 TB in size. Finally, some proprietary formats did contain all of the necessary features in an efficient binary format, but were not open-source and required paid tools or licenses to utilize them. Overall, we found that none of the existing open-source file formats were unsuitable for this dataset.

A formal specification of the XROR format, using the BSON version of the JSON Schema notation, is provided here: `https://rdi.berkeley.edu/metaverse/boxrr-23/dict.json`.

To address the issues with existing open-source file formats, we introduce the new "Extened Reality Open Recording (XROR)" file format. XROR files contain metadata as well as rich event and motion data streams, and are based internally on BSON (Binary JSON), a flexible, widely-supported format with libraries in dozens of languages. Metadata is stored as JSON key-value pairs, while event data and motion data streams are converted to 2D floating-point arrays and compressed using fpzip, a lossless compressor of multidimensional floating-point arrays designed by Lawrence Livermore National Laboratory specifically for the efficient storage and transmission of scientific datasets.

To evaluate the relative efficiency of our new format, we converted a portion of our dataset into a variety of existing open formats, summarized in Table 3(B), as well as our proposed XROR format, as shown in Table 3(C). Even compared to the original, purpose-built formats shown in Table 3(A), XROR achieves space savings of at least 30% with no loss in precision.

Table 3(A): Source file formats for motion data.

Format	Metadata	Motion Data	Event Data	Compression	Avg. Size
.tilt	✓	✓	✓		33.89 MB
.bsor	✓	✓	✓		1.77 MB
.dat	✓	✓	✓		1.27 MB

Table 3(B): Existing general file formats for motion data.

Format	Metadata	Motion Data	Event Data	Compression	Avg. Size
.mvnx	✓	✓	✓		61.90 MB
.bvh		✓			25.79 MB
.bva		✓			13.98 MB

Table 3(B): Proposed new open file format for motion data.

Format	Metadata	Motion Data	Event Data	Compression	Avg. Size
.xror	✓	✓	✓	✓	0.99 MB

Due to the advantages of our new XROR format over the existing alternatives, the entire BOXRR-23 dataset is offered exclusively as XROR files. To help researchers process this format, we have provided open-source tools to parse XROR files, and convert them to and from a variety of formats, including .TILT, .BSOR, .DAT, and .JSON: `https://github.com/metaguard/xror`.

6 Access Instructions

Researchers interested in using the BOXRR-23 dataset are invited to visit `https://rdi.berkeley.edu/metaverse/boxrr-23/`. The permanent DOI is `https://doi.org/10.25350/B5NP4V`. For ease of access, the dataset has been split into 106 .zip files, each containing up to 1,000 users. Each user is represented by a folder, containing one or more recordings from that user in .xror format.

We developed the licensing terms for this dataset in conjunction with the Committee for Protection of Human Subjects (CPHS) and Intellectual Property & Industry Research Alliances (IPIRA) groups at UC Berkeley, with the chief goal of protecting the human subjects contained in this dataset. The dataset is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license, and is additionally subject to an ethical data use agreement (DUA) that prohibits unethical uses of the data, such as attempts to deanonymize the subjects. Access to the dataset is automatically granted upon agreeing to the CC BY-NC-SA 4.0 license and DUA.

7 Intended Use Cases

As detailed in §1, we originally produced this dataset for use in a VR authentication study, which required a large number of users for comparison with traditional biometrics. However, there are a number of interesting uses for this dataset beyond security and privacy research.

7.1 Notable Known Uses

Until recently, this dataset has only been available for internal use at UC Berkeley. Thus far, we have published three papers using this dataset in the XR security and privacy domain:

- We conducted a study that uniquely identified over 55,000 VR users based on their head and hand motion [32]. By using the BOXRR-23 dataset, this study was over 200 times larger than the next largest VR identification study, and the first to demonstrate parity with biometrics like fingerprints.
  - Result: After training a classification model on 5 minutes of data per person, a user can be uniquely identified amongst the entire pool with 94.33% accuracy from 100 seconds of motion.
  - Availability: The source code and documentation required to replicate this result using the BOXRR-23 dataset can be found at `https://github.com/metaguard/identification`.
- In another study, we combined the BOXRR-23 dataset with a survey to demonstrate that a large number of sensitive data attributes can be inferred from VR users based on motion alone [33].
  - Result: Using simple machine learning models, over 35 private data attributes could accurately and consistently be inferred from VR users using head and hand motion data alone.
- In a third paper, we presented "MetaGuard," [31] a differential privacy-based tool for protecting user data privacy in the metaverse, which we evaluated using the BOXRR-23 dataset.
  - Result: We show a significant degradation of attacker capabilities when using MetaGuard.
  - Availability: The source code and documentation required to replicate this result using the BOXRR-23 dataset can be found at `https://github.com/metaguard/metaguard`.

7.2 Future Directions

While the dataset was originally intended for use in the security and privacy domain, and has thus far only been used in this field, we can envision a number of additional interesting applications for this data. Historically, motion capture data has primarily been used for computer graphics, animation, and CGI, and our data could also be used in this domain. For example, it could be used to train large-scale generative machine learning models for natural human motion synthesis tasks. It may also be of interest to researchers studying human-computer interaction in XR. For example, researchers could use the data to investigate interaction patterns likely to cause discomfort or injury.

Finally, this dataset presents a challenging and unique opportunity for theoretical machine learning research, because it consists of long, sequential data, with sequence lengths often in excess of 100,000. Most existing deep learning algorithms are not well equipped to handle sequential data of this size. Currently, our dataset is a rare instance of a task in which classical ML algorithms seem to out-perform deep learning methods [32]. Developing models that can accurately and efficiently ingest the data contained in this dataset may require theoretical advances in machine learning techniques.

7.3 Prohibited Uses

In consultation with CPHS, the Institutional Review Board (IRB) for UC Berkeley, the ethical data use agreement accompanying this dataset prohibits the following unethical uses:

- Researchers should not use the information in the dataset, alone or in combination with any other data, to deanonymize or contact the individuals who are data subjects.
- Researchers should not use the information in the dataset, alone or in combination with any other data, to infer sensitive information about subjects, including protected health information.

8 Population Survey

To shed additional light on the demographics of the users within our dataset, we conducted a large-scale online survey of VR users. The survey contained about 50 questions and received 1,006 responses, of which 830 users were present in the BOXRR-23 dataset. It was conducted in coordination with BeatLeader and other Beat Saber organizations, and thus did not reach the 1% of BOXRR-23 users from Tilt Brush. The full results of this survey are available at `https://arxiv.org/abs/2305.14320`, and are summarized in Figure 4 below.

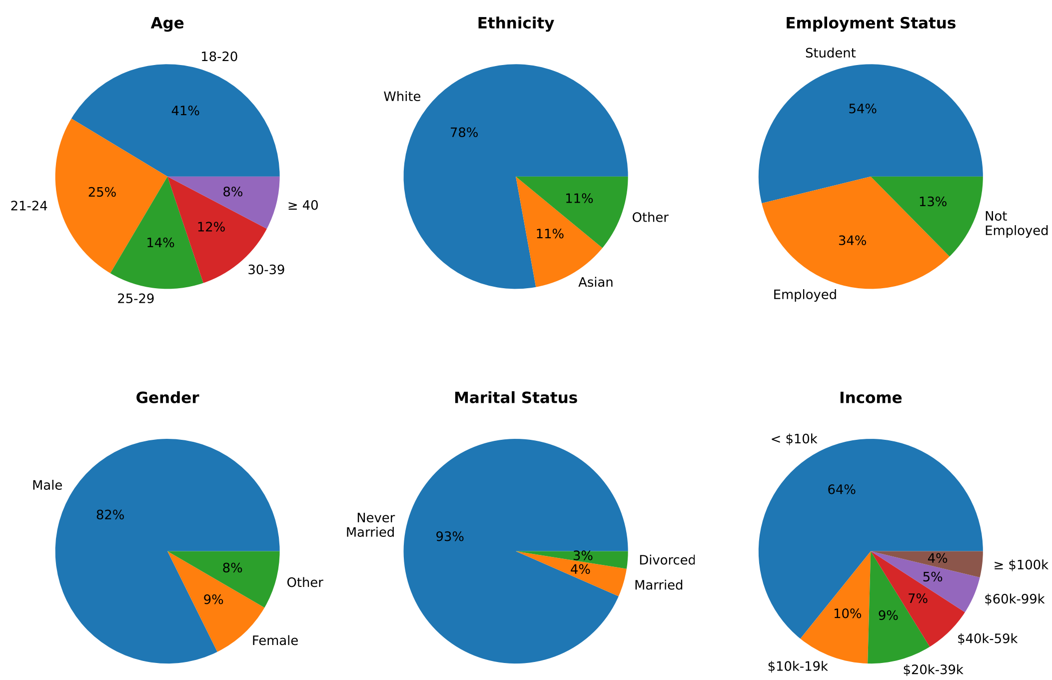

Figure 4: Survey results from 830 users present in the BOXRR-23 dataset.

9 Limitations

As may be evident by the survey results provided in §8, the users included in our dataset are not necessarily representative of a general population. While the subjects are demographically similar to the overall population of VR device users [11], they consist entirely of users who chose to upload a BeatSaber performance or TiltBrush drawing to a public platform. As such, we believe enthusiast or expert-level users to likely be overrepresented in the dataset. Furthermore, the data is derived from just two VR applications, Beat Saber and Tilt Brush, with almost 75% of the users and 99% of the recordings being from Beat Saber alone. Overall, researchers should be cautious when attempting to use this dataset to draw conclusions about larger populations than the ones directly included.

Additionally, there are some risks associated with the dataset being derived from ordinary XR users. Some metadata values, such as Beat Saber song titles or Tilt Brush drawing descriptions, may contain objectionable content due to their user-submitted nature. Metadata constituting user-configured settings like height and handedness should be considered self-reported, and are subject to the typical response biases associated with self-reported values. Finally, because the data is from "the wild" rather than a laboratory study, it originates from a wide variety of heterogeneous XR devices and physical environments, and may include more noise and tracking errors than a lab-created dataset.

10 Ethical Considerations

Because our dataset consists entirely of motion capture recordings from human subjects, significant attention was given to ethics throughout the process of designing and collecting the dataset. Our collection of this dataset was overseen by the UC Berkeley Office for Protection of Human Subjects (OPHS), an OHRP-certified Institutional Review Board (IRB), approved as protocol #2023-03-16120.

We note that in producing this dataset, the authors had no direct contact with human subjects. Instead, our data is derived from three public sources. All data utilized in this study was already broadly, publicly available, to any person in the world with an internet connection, without the need for permissions, credentials, authentication, or any special tools or applications, via the websites of ScoreSaber, BeatLeader, and PolyGone. No new data is being made accessible to the public in the publication of this dataset; our contribution is in finding, scraping, aggregating, reprocessing, enriching, and distributing this existing data, and in surveying the underlying population.

Despite the public nature of the data and the IRB approval, we chose to obtain written permission from ScoreSaber, BeatLeader, and PolyGone before proceeding out of an abundance of caution and respect for the communities from which this data originates. We did not begin collecting data until authorized to do so by these communities, and sought their input throughout the collection process.

Users of the ScoreSaber, BeatLeader, and PolyGone platforms must voluntary install custom software to share their motion recording data with these platforms. They are fully aware of the nature of the data being shared, as uploading and publicly sharing XR data is the explicit purpose of these platforms. They also consent to their recordings being made publicly available in the privacy policies of these platforms. For example, the BeatLeader Privacy Policy, which can be found at `https://www.beatleader.xyz/privacy`, states that "Replays may contain personally identifiable information... Your data, including associated personally identifiable information, will be broadly publicly available to anyone with an internet connection via the BeatLeader website." Users of Google Poly (and PolyGone) consent to making their data publicly available under a CC-BY license.

Beyond consenting to the publication of their data in privacy policies and license agreements, we made further attempts to notify users of their involvement in academic research. For example, we worked in collaboration with the BeatLeader team to inform users of their inclusion in academic research via the official social media channels of the platform, and to develop an opt-out mechanism.

Although users knowingly consented to the public availability of their motion data, we took two additional steps to protect the privacy of data subjects. First, all known explicit identifiers, such as usernames and user IDs, have been removed from the dataset. No potentially sensitive information, such as protected health information, is included in the data or metadata. Second, the dataset is offered under a data use agreement (DUA) that prohibits researchers from attempting to deanonymize or contact the users, or to infer private attributes of the users that may be deemed sensitive. We voluntarily followed the strictest PII data handling standards and guidelines offered by our institution throughout the dataset collection process to preclude the accidental release of non-anonymized data.

Participants originally submitted their motion data to the ScoreSaber, BeatLeader, and PolyGone platforms for purposes other than academic research. Namely, they chose to make their data freely publicly available for reasons such as competitive e-sports or collaborative artwork; as such, users were not compensated for their original submissions, nor for their inclusion in the dataset. Moreover, any participant risks associated with the use of an extended reality device would have been realized by the users regardless of the later inclusion of the resultant motion recordings in this dataset.

While it is impossible to entirely eliminate the risks associated with a new dataset, we believe the additional risk posed by our dataset is minimal in light of the fact that all of the included data was already public. On the other hand, the data has the potential to facilitate significant advances in fields like graphics, HCI, XR, AI/ML, and computer security and privacy. We have taken significant steps to mitigate the potential harms of this dataset while maximizing its utility for beneficial research. Overall, we believe this research constitutes a net benefit to the subjects whose data was included by shedding light on the implications of the motion capture data which they have already, independently chosen to publish. For instance, security and privacy research using this dataset benefits society by highlighting the magnitude of the VR privacy threat and motivating future work on countermeasures.

11 Conclusion

We have presented the BOXRR-23 dataset, a 4.7 TB dataset of extended reality motion capture recordings from users around the world. Unlike existing motion capture datasets, BOXRR-23 is derived from recordings submitted by participants using their own XR devices, rather than a laboratory setup. As a result, it contains over 200 times more users, and over 400 times more recordings, than all known comparable datasets, while simultaneously being more diverse and ecologically valid.

The two XR applications included in BOXRR-23, Beat Saber and Tilt Brush, provide highly complementary motion data. Beat Saber consists almost entirely of fast ballistic movements while Tilt Brush consists almost entirely of fine motor movements, each controlled by a separate part of the brain [15]. By combining these sources, BOXRR-23 provides researchers a diverse collection of motion patterns.

For the first time, BOXRR-23 allows the identifiability of human motion data to be directly compared with biometrics like fingerprints and facial recognition, which have long enjoyed large public datasets. As such, we hope to see new advances in passive authentication mechanisms and privacy-preserving systems for XR, in addition to potential deployments in fields ranging from graphics and animation to usability and human-computer interaction.

In addition to identifying three new sources of motion data not previously widely known to academic researchers, we contributed a new XROR format to enable the efficient storage and transmission of this data. Our XROR format is approximately 30% more efficient than the three original data formats, without any loss in precision, while also being more versatile than most existing open-source formats. Documentation for our dataset is offered according to widely-recognized open standards, including Datasheets for Datasets [17] and Dataset Nutrition Labels [19]. We also conducted a large survey of over 800 users present in the dataset to help researchers understand its demographic constituency.

As advances in extended reality allow this technology to reach increasingly large audiences, human motion data will remain vital to the operation XR and “metaverse” systems for the foreseeable future. In particular, augmented reality (AR) technology promises to be the next major medium of human-computer interactions, potentially even replacing the use of mobile devices such as smartphones. If this reality comes to pass, it is vital that we improve our understanding of the uses and implications of the motion data that these devices are designed to generate. We look forward to seeing future work that deploys our dataset to advance public knowledge in a variety of important fields, and to drive improvements to XR and metaverse experiences that benefit the field of extended reality as a whole.

Acknowledgments and Disclosure of Funding

This work was supported in part by the National Science Foundation (NSF), the National Physical Science Consortium (NPSC), the Fannie and John Hertz Foundation, and the Berkeley Center for Responsible, Decentralized Intelligence (RDI). To the best of our knowledge, none of the authors or funding organizations have competing interests to declare.

References

[1] Accad mocap system and data. URL: `https://accad.osu.edu/research/motion-lab/mocap-system-and-data`.

[2] BeatSaver. URL: `https://beatsaver.com/`.

[3] Biometric accuracy standards. URL: `https://csrc.nist.gov/CSRC/media/Events/ISPAB-MARCH-2003-MEETING/documents/March2003-Biometric-Accuracy-Standards.pdf`.

[4] Cmu graphics lab motion capture database. URL: `http://mocap.cs.cmu.edu/`.

[5] mocapdata.com. URL: `http://mocapdata.com/`.

[6] Polygone Art. URL: `https://polygone.art/`.

[7] ScoreSaber. URL: `https://www.scoresaber.com/`.

[8] Sfu motion capture database. URL: `https://mocap.cs.sfu.ca/`.

[9] Steam. URL: `https://store.steampowered.com/`.

[10] Tilt Brush by Google. URL: `https://www.tiltbrush.com/`.

[11] Report: Vive Users Are 95 Percent Male And Spend 5 Hours Per Week in VR, February 2017. URL: `https://www.uploadvr.com/vive-users-94-9-percent-male-spend-5-hours-week-vr-average/`.

[12] Ijaz Akhter and Michael J. Black. Pose-conditioned joint angle limits for 3d human pose reconstruction. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1446–1455, 2015. `doi:10.1109/CVPR.2015.7298751`.

[13] Shaquitta Dent, Kelley Burger, Skyler Stevens, Benjamin Smith, and Jefferson Streepey. The effect of music on body sway when standing in a moving virtual environment. *PLOS ONE*, 16:e0258000, 09 2021. `doi:10.1371/journal.pone.0258000`.

[14] S.A. Douglas and A.K. Mithal. *The Ergonomics of Computer Pointing Devices*. Applied Computing. Springer London, 2012. URL: `https://books.google.com/books?id=clLlBwAAQBAJ`.

[15] Christoph Fromm and Edward V Evarts. Relation of motor cortex neurons to precisely controlled and ballistic movements. *Neuroscience letters*, 5(5):259–265, 1977.

[16] Beat Games. Beat Saber. `https://beatsaber.com/`. URL: `https://beatsaber.com/`.

[17] Timnit Gebru, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumeé III, and Kate Crawford. Datasheets for Datasets. *arXiv:1803.09010 [cs]*, January 2020. `arXiv:1803.09010`.

[18] Saeed Ghorbani, Kimia Mahdaviani, Anne Thaler, Konrad Kording, Douglas James Cook, Gunnar Blohm, and Nikolaus F. Troje. MoVi: A large multi-purpose human motion and video dataset. *PLOS ONE*, 16(6):e0253157, jun 2021. URL: `https://doi.org/10.1371%2Fjournal.pone.0253157`, `doi:10.1371/journal.pone.0253157`.

[19] Sarah Holland, Ahmed Hosny, Sarah Newman, Joshua Joseph, and Kasia Chmielinski. The dataset nutrition label: A framework to drive higher data quality standards, 2018. `arXiv:1805.03677`.

[20] Ludovic Hoyet, Kenneth Ryall, Rachel McDonnell, and Carol O'Sullivan. Sleight of hand: Perception of finger motion from reduced marker sets. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, I3D '12, page 79–86, New York, NY, USA, 2012. Association for Computing Machinery. `doi:10.1145/2159616.2159630`.

[21] Catalin Ionescu, Dragos Papava, Vlad Olaru, and Cristian Sminchisescu. Human3.6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(7):1325–1339, 2014. `doi:10.1109/TPAMI.2013.248`.

[22] Fernando De la Torre, Jessica K. Hodgins, Adam W. Bargteil, Xavier Martin, J. Robert Macey, Alex Tusell Collado, and Pep Beltran. Guide to the carnegie mellon university multimodal activity (cmu-mmac) database. 2008.

[23] Jonathan Liebers, Lukas Mecke, Alia Saad, Jonas Auda, Uwe Gruenefeld, Florian Alt, Stefan Schneegaß, and Mark Abdelaziz. Understanding user identification in virtual reality through behavioral biometrics and the effect of body normalization. 04 2021. `doi:10.1145/3411764.3445528`.

[24] Matthew Loper, Naureen Mahmood, and Michael J. Black. Mosh: Motion and shape capture from sparse markers. *ACM Trans. Graph.*, 33(6), nov 2014. `doi:10.1145/2661229.2661273`.

[25] Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. AMASS: Archive of motion capture as surface shapes. In *International Conference on Computer Vision*, pages 5442–5451, October 2019.

[26] Christian Mandery, Ömer Terlemez, Martin Do, Nikolaus Vahrenkamp, and Tamim Asfour. The kit whole-body human motion database. In *2015 International Conference on Advanced Robotics (ICAR)*, pages 329–336, 2015. `doi:10.1109/ICAR.2015.7251476`.

[27] Mark Miller, Fernanda Herrera, Hanseul Jun, James Landay, and Jeremy Bailenson. Personal identifiability of user tracking data during observation of 360-degree vr video. *Scientific Reports*, 10, 10 2020. `doi:10.1038/s41598-020-74486-y`.

[28] Mark Roman Miller, Eugy Han, Cyan DeVeaux, Eliot Jones, Ryan Chen, and Jeremy N. Bailenson. A large-scale study of personal identifiability of virtual reality motion over time, 2023. `arXiv:2303.01430`.

[29] Alec G. Moore, Ryan P. McMahan, Hailiang Dong, and Nicholas Ruozzi. Personal identifiability and obfuscation of user tracking data from vr training sessions. In *2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 221–228, 2021. `doi:10.1109/ISMAR52148.2021.00037`.

[30] Meinard Müller, Tido Röder, Michael Clausen, Bernhard Eberhardt, Björn Krüger, and Andreas Weber. Documentation mocap database hdm05. 06 2007.

[31] Vivek Nair, Gonzalo Munilla Garrido, and Dawn Song. Going incognito in the metaverse, 2023. `arXiv:2208.05604`.

[32] Vivek Nair, Wenbo Guo, Justus Mattern, Rui Wang, James F. O'Brien, Louis Rosenberg, and Dawn Song. Unique identification of 50,000+ virtual reality users from head & hand motion data, 2023. `arXiv:2302.08927`.

[33] Vivek Nair, Christian Rack, Wenbo Guo, Rui Wang, Shuixian Li, Brandon Huang, Atticus Cull, James F. O'Brien, Louis Rosenberg, and Dawn Song. Inferring private personal attributes of virtual reality users from head and hand motion data, 2023. `arXiv:2305.19198`.

[34] Ken Pfeuffer, Matthias Geiger, Sarah Prange, Lukas Mecke, Daniel Buschek, and Florian Alt. Behavioural biometrics in vr: Identifying people from body motion and relations in virtual reality. pages 1–12, 04 2019. doi:10.1145/3290605.3300340.

[35] Jose Luis Ponton, Haoran Yun, Carlos Andujar, and Nuria Pelechano. Combining Motion Matching and Orientation Prediction to Animate Avatars for Consumer-Grade VR Devices. *Computer Graphics Forum*, 41(8):107–118, 2022. doi:10.1111/cgf.14628.

[36] Viktor Radulov. BeatLeader. URL: https://www.beatleader.xyz/.

[37] Christian Schell, Fabian Sieper, Lukas Schach, and Marc E. Latoschik. cschell/who-is-alyx: v2.0, February 2023. doi:10.5281/zenodo.7663984.

[38] Leonid Sigal, Alexandru Balan, and Michael Black. Humaneva: Synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion. *International Journal of Computer Vision*, 87:4–27, 03 2010. doi:10.1007/s11263-009-0273-6.

[39] Pier Paolo Tricomi, Federica Nenna, Luca Pajola, Mauro Conti, and Luciano Gamberini. You can't hide behind your headset: User profiling in augmented and virtual reality, 2022. arXiv:2209.10849.

[40] Nikolaus F. Troje. Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *Journal of vision*, 2 5:371–87, 2002.

[41] Matt Trumble, Andrew Gilbert, Charles Malleson, Adrian Hilton, and John Collomosse. Total capture: 3d human pose estimation fusing video and inertial sensors. In *2017 British Machine Vision Conference (BMVC)*, 2017.

[42] Shiv Naga Prasad Vitaladevuni. Human Movement Analysis: Ballistic Dynamics, and Edge Continuity for Pose Estimation. October 2007. URL: http://hdl.handle.net/1903/7610.

[43] Jan Wöbbeking. Beat Saber generated more revenue in 2021 than the next five biggest apps combined, August 2022. URL: https://mixed-news.com/en/beat-saber-generated-more-revenue-in-2021-than-the-next-five-biggest-apps-combined/.

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See Section 9.
   (c) Did you discuss any potential negative societal impacts of your work? [Yes] See Section 10.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes] See Section 10.
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A]
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments (e.g. for benchmarks)...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] See Section 7.1.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [N/A]
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [N/A]

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [N/A]

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] See Section 3.
   (b) Did you mention the license of the assets? [Yes] See Section 2.
   (c) Did you include any new assets either in the supplemental material or as a URL? [Yes] See Section 6.
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [Yes] See Section 10.
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [Yes] See Sections 9 and 10.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [Yes] See Section 10.
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [Yes] See Section 10.
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [Yes] See Section 10.

A Appendix

1. **Required supplementary materials:**
   (a) **Dataset documentation and intended uses.** We have included supplemental materials following the "datasheets for datasets" and "dataset nutrition" labels frameworks.
   (b) **URL to website/platform where the dataset/benchmark can be viewed and downloaded by the reviewers.** `https://rdi.berkeley.edu/metaverse/boxrr-23/`
   (c) **Author statement that they bear all responsibility in case of violation of rights, etc., and confirmation of the data license.** We bear all responsibility in case of any violation of rights during the collection of the data or other work, and will take appropriate action when needed, e.g. to remove data with such issues. The dataset is offered under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.
   (d) **Hosting, licensing, and maintenance plan.** Our current cloud data storage provider is Wasabi. Long-term hosting and maintenance of the dataset will be facilitated and funded by the Berkeley Center for Responsible, Decentralized Intelligence (RDI).
2. **Supplementary materials for accessibility:**
   (a) **Links to access the dataset and its metadata.** The dataset can be downloaded at `https://rdi.berkeley.edu/metaverse/boxrr-23/`. The metadata description is available at `https://github.com/metaguard/xror`.
   (b) **Provide a detailed explanation on how the dataset can be read.** We use the new XROR format described in §5 of this paper. The tools for parsing the XROR files are provided at `https://github.com/metaguard/xror`.
   (c) **Long-term preservation.** We are committed to the long-term availability of the dataset, which will be maintained and funded by the Berkeley Center for Responsible, Decentralized Intelligence (RDI) for the foreseeable future.
   (d) **Explicit license.** The dataset is offered under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.
   (e) **Structured metadata.** The dataset landing page contains structured metadata. The dataset schema is available at `https://rdi.berkeley.edu/metaverse/boxrr-23/dict.json`.
   (f) **Persistent dereferenceable identifier.** DOI: `https://doi.org/10.25350/B5NP4V`.
3. **Supplementary materials for reproducibility:** We do not include any original machine learning experiments in this paper. Instead, we reference three prior results using this dataset, the source code and documentation for which are described in §7.1.