Navigating the Kinematic Maze: Analyzing, Standardizing and Unifying XR Motion Datasets

Christian Rack[1]*, Vivek Nair[2], Lukas Schach[1], Felix Foschum[1], Marcel Roth[1], Marc Erich Latoschik[1]

[1]University of Würzburg
[2]UC Berkeley

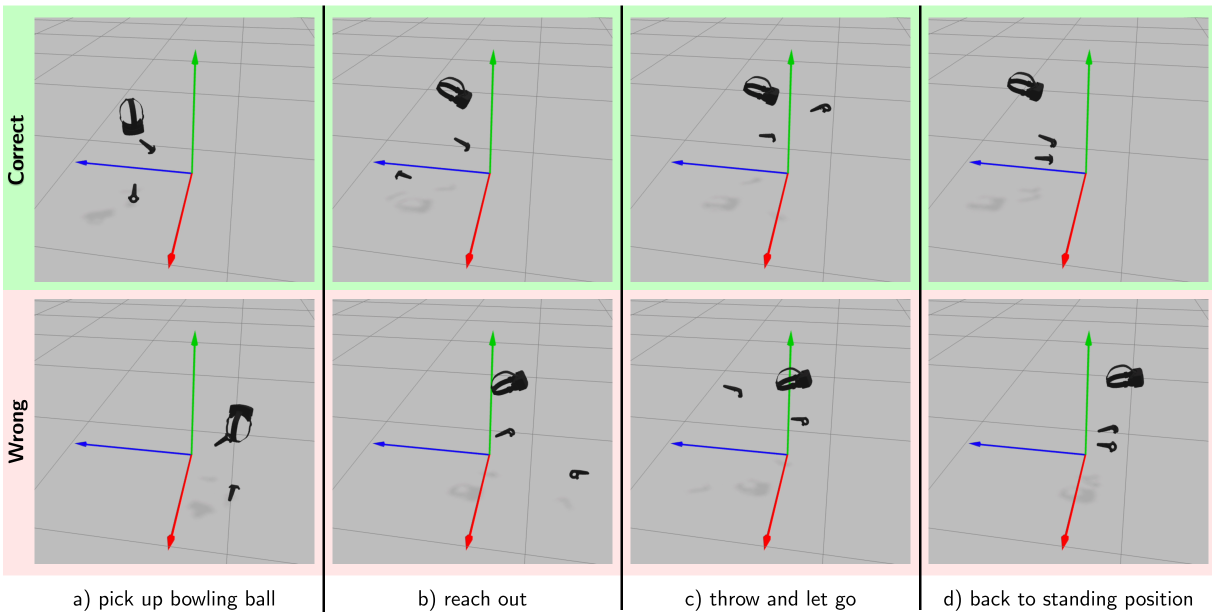

Figure 1: Each row shows identical frames from a VR bowling sequence taken from [14], reconstructed in our motion visualization tool. In the first row, the recording is correctly imported, showing accurate user movements. The second row, however, illustrates the impact of assuming an incorrect coordinate system (specifically, an inverted Z-axis): the user's positions are mirrored along the Z-axis and rotations are twisted.

Abstract

This paper addresses the critical importance of standards and documentation in kinematic research, particularly within Extended Reality (XR) environments. We focus on the pivotal role of motion data, emphasizing the challenges posed by the current lack of standardized practices in XR user motion datasets. Our work involves a detailed analysis of 8 existing datasets, identifying gaps in documentation and essential specifications such as coordinate systems, rotation representations, and units of measurement. We highlight how these gaps can lead to misinterpretations and irreproducible results. Based on our findings, we propose a set of guidelines and best practices for creating and documenting motion datasets, aiming to improve their quality, usability, and reproducibility. We also created a web-based tool for visual inspection of motion recordings, further aiding in dataset evaluation and standardization. Furthermore, we introduce the XR Motion Dataset Catalogue, a collection of the analyzed datasets in a unified and aligned format. This initiative significantly streamlines access for researchers, allowing them to download partial or entire datasets with a single line of code and without the need for additional alignment efforts. Our contributions enhance dataset integrity and reliability in kinematic research, paving the way for more consistent and scientifically robust studies in this evolving field.

Index Terms: Human-centered computing—Human computer interaction (HCI)—Kinematic Research—Datasets;

1 Introduction

The way we move tells a lot about us. Our motions alone can allow inferences about our identity [17,20,24] and personal attributes such as gender [13,22], age [12], or physique [21]. They can even aid in the detection of medical conditions like Parkinson's [26] disease or cybersickness [7–9]. Consequently, motion data has become a pivotal element in kinematic research, a field focused on the analysis and understanding of human motions.

With the advent of Virtual, Augmented, and Mixed Reality (VR, AR, MR, collectively referred to as Extended Reality or XR), the availability of motion data has significantly increased. These XR systems inherently capture user motions to facilitate immersive interactions, thus providing a rich source of data for kinematic research. This development has led to the creation of specialized XR user motion datasets, which serve as foundational resources for the field.

While initial studies in kinematic research often relied on private datasets, there is a growing trend towards open-access data, crucial for enabling the verification and replication of research findings. However, the mere availability of these datasets is not sufficient.

*e-mail: christian.rack@uni-wuerzburg.de

They must meet specific standards to be genuinely beneficial for kinematic research.

In contexts working with spatial data, explicit documentation of specifications like the used coordinate system, representation of rotations or units of measurement is essential. Otherwise, the correct interpretation of the dataset is largely left to chance. This is particularly significant for researchers using multiple datasets for their work, as they have to ensure a consistent data format —– without it, identical motions can yield different representations across datasets, leading to confusion or overfitting in analyses and machine learning applications. Surprisingly, existing motion datasets predominantly lack such essential documentation, significantly impeding their effective utilization in kinematic studies. This oversight is not just inconvenient, forcing researchers to deduce crucial specifications through trial and error, but also poses a real risk: incorrect assumptions about these specifications can lead to fundamental errors in analyses, leading to inflated claims and conclusions.

The impetus for this paper came from our own experience using multiple motion datasets in a recent research project. Gradually, we realized that our initial assumptions about their specifications were incorrect. This issue, which initially went unnoticed, could have led us to continue with flawed analyses. Recognizing the potential for similar challenges to affect other researchers, this paper seeks to bridge these gaps in practice. Our goal is to enhance the quality, reliability, and usability of future XR user motion datasets, making this a significant contribution to the field. Our objectives include:

1. Analyzing 8 existing XR motion datasets and revealing previously undisclosed details.

2. Introducing the XR Motion Dataset Catalogue[1]: this catalogue includes each of the analyzed datasets in an aligned and unified format for quick and easy access by researchers.

3. Offering guidelines and best practices for the creators of future datasets, as well as for authors and reviewers of future works.

4. Introducing a web-based Motion Visualization Tool to replay motion data recordings[2].

2 Analysis of Published XR Motion Dataset

The critical specifications we define in Section 4 are the result of our experiences and analysis of the existing landscape of XR motion datasets. This section describes our methodology for analyzing published datasets to uncover missing specifications. Here, our main objective was to faithfully reconstruct each motion sequence from any given dataset. We developed a web-based motion visualization tool for inspecting reconstructed motioned sequences, enabling us to validate our assumptions about missing critical dataset specifications visually.

2.1 Terminology and Characteristics of XR Motion Data

We use the following terminology in the context of this paper. A *recording* is a continuous sequence of data from one user. When there are multiple recordings for one user, they can be within the same or different recording *sessions*. Recordings can be of different *data types*: the primary data type we focus on in this paper is motion data, but often additional types get collected, like eye-tracking data, physiological data or application-specific data (e.g., game events). The duration of a recording can vary depending on the dataset, ranging from a few seconds to several hours. The structure of these recordings is tabular: every row represents an individual *frame*, and each column denotes a distinct *attribute* associated with that frame, for example spatial or rotational coordinates (X, Y, Z) or time information. Commonly, XR motion datasets provide coordinates for each *peripheral*, which usually describes the head mounted display (HMD) and left and right controllers in VR contexts, or the HMD and hands in AR contexts.

The datasets we have analyzed represent motion data in 3D Cartesian coordinate systems, using X, Y and Z axes. The choice of axis assignment varies and seems to align with the system used by the software where the tracking data was acquired. For instance, datasets from Unity applications typically map X to Right, Y to Up, and Z to Backward, while Steam OpenVR applications also map X to Right and Y to Up but differ by mapping Z to Forward. Additionally, there are different ways to represent rotations. The most common notations are Euler angles and quaternions. Euler angles consist of three values, and can be intrinsic or extrinsic, expressed in degrees or radians, and might follow different orders of axes (XYZ, YXZ, XZX, etc.), each combination resulting in different motion interpretations. Quaternions are composed of four values (X,Y, Z, and W) and offer a more consistent and less ambiguous method for rotation representation as there are no variants as with Euler angles. The timing of frames, often given in frames per second (fps), is another critical aspect, as it dictates the speed and fluidity of the reconstructed motion.

2.2 Motion Visualization Tool

Even though statistical analysis of motion data play an important role for identifying missing specifications, we found that visual replays of recordings are invaluable for confirming these specifications and revealing issues that statistical methods might miss. As part of our analysis, we developed a dedicated web-based motion visualization tool using HTML, JavaScript and the Three.js 3D library. This tool accepts motion recordings converted into a unified format provided by our conversion scripts (see Section 5), creating an interactive 3D scene that renders HMDs and hand controllers, thereby replicating user motions. This tool played a key role in our analyses, as misalignments in the recordings become immediately apparent: incorrect coordinate systems led to unnatural positions and animations; wrong units caused the peripherals to appear too near or too far; and inaccurate time coding notably affected replay speed. A live demonstration of this tool is available `https://cschell.github.io/kinematic-maze`. It showcases sample recordings from the datasets we examined and illustrates how different data misinterpretations can cause different types of distorted animations.

2.3 Analysis

In this study, we focused on collecting and analyzing datasets that provide motion data from XR users, specifically targeting datasets that contain typical tracking data from common VR and Augmented Reality AR setups. We consciously excluded datasets involving full-body motion capture, as these datasets possess unique characteristics requiring separate consideration. In addition to the eight datasets listed in Table 1, we attempted to acquire the dataset from Agac et al. [2]. Unfortunately, we were unable to obtain a functional copy from the authors within our publication timeframe. We also evaluated the AR dataset by Abdrabou et al. [1], but found it to be incomplete in its current form and were not able to produce reliable replays from it. We plan to include both the datasets from Agac et al. and Abdrabou et al. in our collection once we can access working versions.

In our exploration of the XR user motion datasets, we observed different approaches in dataset creation and documentation, each with its unique strengths and challenges. To collect critical specifications, we systematically downloaded each dataset and developed conversion scripts through an iterative process, constantly cross-verifying the outcomes using our visualization tool. Initial analysis

[1] `https://huggingface.co/datasets/cschell/xr-motion-dataset-catalogue`

[2] live demo: `https://cschell.github.io/kinematic-maze`

involved reviewing the corresponding paper and any available documentation for preliminary information on the specifications. Following this, we analyzed the data files and loaded them with differing specifications, until the animations in our motion visualization tool accurately represented the data. In this section, we provide a brief overview of each dataset, highlighting their distinct characteristics with respect to the discussed specifications. Table 1 summarizes these specifications, including those not previously reported.

The 'LiebersLabStudy21' [14] dataset focuses on VR users performing specific bowling and archery motions. Aimed at identifying VR users by their motion, it uses Unity's coordinate system (X: Right, Y: Up, Z: Backward). Each recording is stored in individual CSV files, with filenames indicating the corresponding body normalization. A Readme file accompanies the dataset, providing basic documentation of the CSV columns. The dataset employs Euler notation for rotations, which our analysis suggests is intrinsic, in degrees, and follows an XYZ order.

The 'LiebersHand22' dataset [15] includes interactions of 16 users with various interface elements (e.g., buttons and sliders), in AR and VR environments, again for motion-based user identification. Unfortunately, a comprehensive documentation is lacking. The data is stored in individual TSV files, incorporating both quaternion and Euler notation for rotations. This dataset also adheres to the Unity coordinate system.

'RMillerBall22' [18] captures the motions of 41 users performing ball throwing actions in VR, segregated by user, device, session, and peripheral. The CSV files combine multiple session recordings, demarcated by a special row, requiring some unique steps for the data import. The dataset adopts the Unity coordinate system and implies a fixed frame rate of 45 fps or 75 fps (depending on the device), though this is not explicitly detailed in the accompanying documentation.

The 'Who Is Alyx?' [24] dataset, collected from 71 users playing 'Half-Life: Alyx', provides a comprehensive documentation and user metadata. It employs the Steam OpenVR coordinate system (X: Right, Y: Up, Z: Forward) and provides individual files for each recording.

'BOXRR-23' [27], encompassing over 100,000 VR users playing Beat Saber and Tilt Brush, is notable for its scale and the use of the specialized binary format "XROR". This format not only includes motion data, but also metadata about user and the gaming session. A critical aspect of this dataset is the varying specifications: Beat Saber recordings uses meters and seconds, whereas Tilt Brush uses decimeters and milliseconds. It is crucial to recognize the distinct nature of recording files between these two applications. In Beat Saber, each file captures the entire sequence of a player interacting with a song, including comprehensive tracking of all peripherals. In contrast, Tilt Brush recordings consist of individual brush strokes made by the user, resulting in a series of short, isolated hand movements with significant time gaps between them. This unique structure of Tilt Brush data presents a different kind of motion analysis challenge, as it does not represent a continuous user activity. Importantly, these recordings only include motion data for one hand, omitting the other hand and the head-mounted display (HMD), which poses a unique challenge. Due to this limitation, it is difficult to verify with certainty whether the reconstructed recordings in our visualization tool are entirely accurate, so our conclusions about the coordinate system and units used in Tilt Brush have to be taken with a grain of salt.

The dataset 'LiebersBeatSaber23' [16] contains data from 15 users who played the game Beat Saber and is designed for user identification based on motion. The recordings of all users are provided within a single CSV file of nearly 2 GB in size. The dataset includes both, quaternion and Euler angle representations for rotations. There is some ambiguity regarding the time encoding, as it is unclear if the frame rate specification of 90 Hz is just an average value or constant, and there is no dedicated timestamp column.

'MooreCrossDomain23' [19] includes motion data from 45 VR users performing assembly tasks, and is also intended for user identification research. It provides a rich variety of rotation representations: quaternions, rotation matrices, and Euler angles, though without specifics on the Euler notation. The data repository is well structured with one CSV file for each recording, but lacks documentation.

'VR.net' [28] was designed for cybersickness research and provides 16 users playing various VR games. It uses the Steam OpenVR coordinate system but introduces a unique data encoding, with tracking data represented as transformation matrices. The absence of documentation necessitated a deeper exploration to understand and utilize the dataset effectively.

In summary, each of these datasets reflects the diversity and complexity inherent in XR user motion studies. Only few datasets provide exhaustively detailed documentation, and in many cases critical information is left to guesswork. This lack of detailed documentation might stem from a lack of awareness about the importance of such specifics or possibly from a primary focus on the additional accolade of publishing datasets rather than future work This emphasizes the necessity for standardized practices in dataset creation and documentation to facilitate ease of use, accurate interpretation, and robust research outcomes in kinematic studies.

3 Wrong Specifications Lead to Irreproducible Results

If researchers assume wrong dataset specifications, observed results will be based on distorted data and can lead to inflated estimates of analyses and model performances. These distortions became immediately apparent in our analyses using the motion player, but are easily overlooked by researchers who do not put effort into exhaustive visualizations and analyses.

Such errors can play out differently depending on whether researchers are dealing with a single dataset or attempting to compare multiple datasets. When working with a single dataset, certain issues may not seem as problematic, particularly if the dataset maintains internal consistency. For example, when training neural network models, the specific type of Euler angles used is trivial as long as the same notation is maintained across all input sequences. This consistency often circumvents the need for explicit specifications, explaining why many previous studies, which up until now typically rely on a single dataset, do not address these details. However, ambiguities become problematic when applying spatial data transformations, such as the data encodings proposed by Rack et al. [25], which are designed to remove unwanted noise from data to improve training of machine learning models. These methods require information about the dataset's specifications to correctly apply each spatial transformation. Consequently, inconsiderately adopting the wrong specifications will inevitably result in distorted data, leading to the opposite desired effect, even if only a single dataset is used.

The problems become more pronounced and complex when comparing across different datasets. If each dataset adopts a different structure, coordinate system, or unit of measurement, aligning and integrating these datasets becomes an indispensable task. For instance, in previous efforts we erroneously combined the 'Who is Alyx?' and 'BOXRR-23' datasets for user identification in machine learning models without appropriate conversion between both coordinate systems. This oversight led to an unexpected ease in user identification by the models. They seemed to distinguish users not just based on their motion profile, but also based on the coordinate system, effectively reducing the user pool for predictions during evaluations. Once we identified and rectified this error in our preprocessing, retraining the models on the correctly aligned data resulted in a significant drop in identification performance. This experience underscores the importance of dataset alignment to ensure the accuracy and validity of research findings.

Table 1: Overview over critical specifications for published XR user motion datasets; underlined items highlight previously undisclosed information; coordinate system notations: 'Unity' is 'X: Right, Y: Up, Z: Backward' and 'OpenVR' is 'X: Right, Y: Up, Z: Forward'; entries marked with '*' denote specifications about which we have some uncertainty.

Name	Coord. Sys.	Rot. Repr.	Units	Time	File Format
LiebersLabStudy21	Unity	Euler (deg;XYZ;extr..)	m	relative (ms)	CSV
LiebersHand22	Unity	Euler (deg;XYZ;extr.)/Quat.	m	relative (s and ms)	TSV
RMillerBall22	Unity	Quaternions	m	fixed 45 or 75 fps	Custom
Who Is Alyx?	OpenVR	Quaternions	cm	abs. (ISO 8601)	CSV
BOXRR-23 – Beat Saber	Unity	Quaternions	m	relative (s)	XROR
BOXRR-23 – Tilt Brush	Unity*	Quaternions	dm*	relative (ms)	XROR
LiebersBeatSaber23	Unity	Euler (deg;XYZ;extr.)/Quat.	m	90 Hz (presumably fixed)	CSV
MooreCrossDomain23	Unity	Euler (deg;XYZ;extr.)/Quat./6D	m	rel. (s)	CSV
VR.net	OpenVR	Transformation Matrix	m	absolute (unix)	CSV

In conclusion, the accuracy of dataset specifications is vital in ensuring the validity and reproducibility of research findings. Inadequate reporting of critical specifications by authors poses significant challenges for reviewers and other researchers in reproducing and verifying their methods, thereby impacting the reliability of the research. This situation underscores the crucial need for meticulous attention to dataset details, thorough documentation, and robust evaluation methods in kinematic research. Addressing these challenges is essential for researchers to maintain the integrity and reproducibility of their work, laying a solid foundation for producing reliable and impactful results in the field.

4 Critical Specifications for Motion Datasets

The following specifications are the results from our analyses and experiences we detailed in Section 2 and address datasets involving typical motion tracking data from XR setups. While these standards are broadly applicable and should hold true for other types of tracking data, such as image-based tracking [6] or acoustic navigation [4], they may necessitate additional considerations specific to each technology. These specifications aim to establish a foundational framework for XR motion datasets, with the understanding that the principles can be adapted and extended to accommodate a wider range of motion tracking methodologies.

[S1] Structure. The structure of a dataset, particularly how recordings are organized, significantly impacts data accessibility. A poorly structured dataset can lead to confusion about which files correspond to specific sessions or participants. For example, if a dataset combines multiple recordings into a single file without clear demarcation, it becomes challenging to isolate and analyze individual recordings. Conversely, if every motion sequence is saved as a separate file without a systematic naming convention or indexing, researchers might struggle to locate and aggregate relevant data for their studies. This issue is even more relevant in large datasets, where the sheer volume of recordings necessitates a well-defined organizational scheme to facilitate easy access and selection. Efficient data retrieval and analysis depend on a logical, well-documented structure that aligns with the research objectives.

[S2] File Format. The file format is vital in determining how recordings can be loaded and attributes correctly labeled. An unsuitable or poorly documented file format can lead to misunderstandings. This affects the integrity of the research, as conclusions drawn from improperly interpreted data are likely to be erroneous.

[S3] Coordinate System. Understanding the coordinate system used in a dataset is essential for accurately interpreting spatial data. If it is unclear how the X, Y, and Z coordinates correspond to axes like up, forward, and left/right, spatial relationships cannot be properly reconstructed. For example, the same motion will suddenly look very unrealistic if 'X' gets interpreted as 'up' instead of 'Y' – not only because positions are flipped, but also because rotations will be misinterpreted. The datasets analyzed for this work all use two similar coordinate systems, which only slightly differ: one is left-handed, so Z points 'right', the other right-handed, so Z points 'left'. While this difference might seem minor, as Figure 1 illustrates, an incorrect assumption about the coordinate system results in both mirrored positions and significantly distorted rotations of peripherals. It is important to note that while transforming positions between a left-handed and right-handed system is relatively straightforward (typically involving flipping the z-axis), the same does not apply to rotations. Rotations involve more complex transformations because they need to account for the changed orientation of the axes, which affects how objects orient themselves in space.

[S4] Units of Measurement. The units of measurement used in a dataset, whether meters, centimeters, custom units, etc., are fundamental for accurately assessing and comparing spatial data. For example, assuming centimeters instead of meters would lead to peripherals appearing a 100 times closer and motions 100 times slower.

[S5] Representation of Rotations. Misinterpreting rotations (e.g., Euler angles, quaternions, or transformation matrices) leads to incorrect reconstructions of motions. For instance, Euler notation seems straightforward at first glance, as it defines rotations around the X, Y, and Z axes. Yet, to apply it correctly, one needs to know whether the rotations are intrinsic (rotating about the axes of the moving coordinate system) or extrinsic (rotating about the axes of the fixed coordinate system), as well as the order of applying rotations along each axis. Like before, wrong assumptions regarding this are easy to miss, but will lead to incorrectly reconstructed motions.

[S6] Time Encoding. Accurate timing information is crucial for understanding the sequence of frames and duration of movements in motion data. This can be represented through timestamps or a fixed framerate. Without clear timing data, the dynamics of motion cannot be accurately analyzed. Assuming the wrong timing of frames will effectively lead to reconstructed motions to be too fast or too slow.

5 The XR Motion Dataset Catalogue

In the current landscape, accessing the individual datasets poses a significant logistical challenge. Even if researchers have the discussed specifications at hand, they have still have to navigate through a cumbersome process of retrieving each dataset from the respective websites, understand file structure and organization, and laboriously import and convert the files into a usable format for their research. This not only demands considerable time and effort but also introduces a considerable risk of errors and inconsistencies. To streamline this process and enhance the accessibility and usability of motion datasets, we introduce the XR Motion Dataset Catalogue. This catalogue represents a concerted effort to collect all the analyzed datasets and convert them into a unified format. We host the final collection of aligned datasets on the Hugging Face platform to streamline access and usability. Hugging Face is a widely recognized platform known for hosting machine learning models and providing a robust

Table 2: Specifications we use for the XR Motion Dataset Catalogue.

Coordinate System	X: Right, Y: Up, Z: Forward (OpenVR)
Units	Centimeters
Rotation Rep.	Quaternions
Time Encoding	relative (milliseconds)
File Format	CSV

infrastructure for sharing datasets. This platform is particularly beneficial for researchers as it offers easy and flexible access to datasets. One significant advantage is the ability to selectively download specific parts of a dataset rather than the entire collection, which can be highly efficient for targeted research needs. As a result, it is possible to download complete datasets, or just selected recordings, with just one line of code.

The foundation of this catalogue is a collection of conversion scripts, which we also publish on GitHub[3]. For each dataset, we have implemented a conversion script that defines the correct mapping of attribute names (i.e., column names) for data import, and performs all necessary actions to align the datasets in terms of their specifications, as detailed in Table 2. The choice of these parameters for alignment was influenced primarily by the default settings of Three.js, the 3D engine that powers our visualization tool. This repository is a valuable starting point for any researcher planning to work with either of these datasets, as it demonstrates how to quickly import and convert each dataset. By providing these ready-to-use scripts that align with the standards proposed in our guidelines, we substantially reduce the initial workload —– and room for error —– involved in dataset preparation.

Altogether, the XR Motion Dataset Catalogue simplifies the process of acquiring and working with these datasets. Researchers can now effortlessly retrieve any dataset from the catalogue in a standardized format, reducing the complexity and error potential inherent in the previous methods of dataset acquisition and preparation. The XR Motion Dataset Catalogue marks a substantial step forward in making XR motion data more accessible, consistent, and user-friendly for the research community.

6 Guidelines for Improving Motion Dataset Practices

At first glance, it might look surprising why creators of these motion datasets did not resort to established standards. After all, other domains such as computer graphics, game development, and digital animation, already provide standards like HAnim, COLLADA, and FBX for exchanging motion data. These standards accommodate complex requirements, encompassing motion as well as additional elements like textures, lighting, and 3D modeling details. However, in kinematic research, these comprehensive standards are often ignored in favor of custom solutions. The primary reason for this is that the intricacy of these data standards, while advantageous for visually intensive domains, adds unnecessary complexity and overhead. Kinematic research primarily concentrates on analyzing motion data in its purest form, without the auxiliary features these standards offer.

As a result, kinematic research tends to favor simpler, more accessible, and space-efficient data exchange formats. Formats like CSV or basic binary encodings are preferred because they simplify the exporting and importing processes and eliminate the need for additional dependencies. This ensures a focused examination of motion data, enhancing ease of handling, analysis, and replication in kinematic studies. Nevertheless, as the field transitions from the initial use of private datasets to the current trend of larger public datasets, the absence of standardized practices and conventions in this domain emerges as a significant challenge.

Against this backdrop, we propose a list of guidelines creators and users of future motion datasets for kinematic research should follow to make their datasets accessible. These guidelines are the result of our analyses described in Section 2 and best practices we have established over time within our team for creating, utilizing, and evaluating XR user motion datasets. Addressing dataset creators, as well as authors and reviewers, the guidelines aim to foster transparency, consistency, and accessibility, which is essential for the integrity and advancement of research in this field. By adhering to these practices, researchers can significantly improve the quality of their datasets, ensure the reproducibility of results, and facilitate more effective and accurate analyses in kinematic studies.

6.1 Guidelines For Creators of Motion Datasets

As the field of kinematic research advances, the importance of standardized practices in dataset creation and documentation cannot be overstated. To facilitate this, we propose the following guidelines and best practices for future datasets in XR user motion studies.

[GC1] Use accessible standards and report critical requirements. First and foremost, it is imperative that future datasets comprehensively report all relevant specifications discussed in Section 4. For file formats, we recommend adopting common formats like CSV, or binary equivalents such as HDF5 [11] or Parquet [3] for tabular data, or formats like JSON, YAML, or BSON (binary variant of JSON). Custom formats, like the XROR format used by the BOXRR-23 dataset, can also be a sensible solution for datasets with very specific characteristics and unique requirements. Regardless of the format chosen, the dataset should be accompanied with a thorough documentation of each data attribute and how it has been labeled. We advocate using quaternions to represent rotations, as Euler angles require additional specifications and are easily misinterpreted. Providing a timestamp column with the passed time since the start of the recording in milliseconds offers a clear way to specify the timing of each frame. It is advisable to omit redundant data such as velocities and accelerations from the dataset unless they are crucial for specific applications. Including such derived data can lead to inconsistencies, as they can be calculated from the primary data (like position) with different methods, potentially introducing discrepancies. Furthermore, the exclusion of these derivatives simplifies the dataset, enhancing its accessibility and reducing potential confusion for users. Clear documentation of all of these aspects is crucial for accurate data interpretation and replication of research.

[GC2] Account for sensible file structure. Motion recordings should be methodically arranged in an accessible and transparent file structure, ideally dedicating a single file to each recording and data type. It is common for XR datasets to include additional time series data such as events from the application or other data types like eye-tracking. Since these data types often operate at different frequencies and may not synchronize perfectly with the motion data, they should be stored in separate files, each with its own timestamp column. This organization ensures that the timestamps for each data type are accurately maintained and align with the motion data, provided these timestamps originate from the same or synchronized clocks. While there are data formats that allowed storing everything within one file, such as XROR, BSON, HDF5, etc., we argue that this separation keeps the organization simple and allows researchers to work with less complex formats and selectively load only the needed data type, enhancing the dataset's usability and flexibility. Combined with a clear naming scheme, this should make it straightforward to select individual recordings without having to inconveniently extract them from the rest of the dataset. Not only does this save time, resources and frustration, it also makes the routines for importing recordings less susceptible to bugs.

[GC3] Allow easy and permanent dataset access. Other re-

[3] https://github.com/cschell/xr-motion-dataset-conversion-scripts

searchers should be able to easily and permanently access and download published datasets. Several of the analyzed datasets require to download the full data, which is especially cumbersome if the total size is prohibitively large. Often, only certain data types or recordings of a few users are needed, so dataset creators should look for ways to not only provide bulk downloads but also partial downloads. Tools like Git paired with Git LFS or DVC allow straightforward ways to easily manage even large datasets, offering a convenient alternative to single zip archives. For hosting, there are options like Hugging Face, GitHub, Kaggle, Zenodo, etc. that not only provide free hosting, but often tooling for up- and downloading datasets, which improves accessibility for both, creators and users of datasets.

[GC4] Provide demo code. Additionally, datasets should include example scripts that demonstrate how to properly load the data and correctly identify each attribute. These scripts not only serve as a practical tool for other researchers but also act as a form of documentation, offering insights into the intended use and interpretation of the data.

[GC5] Offer visualizations. An ideal enhancement for future datasets is the provision of options for data visualization. This significantly eases the process of understanding and analyzing motion data, making datasets more accessible and user-friendly. To aid in this, we publish the code for our motion visualization tool and provide instructions for how to set it up.

[GC6] Add contextual information. Beyond the discussed fundamental specifications, there is additional contextual information that can greatly benefit researchers. For example, providing background information about the used data source, such as whether it was collected from Unity, Steam OpenVR, or other platforms, can give important context information about the dataset's characteristics. Moreover, awareness of application-specific traits is crucial. For instance, scenarios where users are teleported within a scene can result in abrupt and seemingly inexplicable 'jumps' in the data. Similarly, it should be clarified if data for certain peripherals, like hand tracking, are available only under specific conditions (like when hands are visible to the camera). This helps to distinguish between intentional data absences and potential errors. Additionally, understanding whether users might place their controllers or HMDs down during a session, or whether they are seated or standing, can offer valuable insights into the dataset's dynamics. These nuances, although seemingly minor, can have substantial implications for the accuracy and reliability of research outcomes, underscoring the importance of comprehensive dataset documentation. Creators of datasets should also be aware that their datasets might be used for different purposes, so they should not just focus on the specific requirements of their own research.

[GC7] Disclose recording methods. We strongly recommend that dataset creators disclose and discuss the software used to generate their datasets. This transparency not only allows for the reproduction, verification, and extension of research but also fosters an environment of open collaboration and innovation in the field. By sharing the tools and methods used for data collection, researchers can contribute to a more robust and dynamic understanding of kinematic data in XR environments.

[GC8] Make information easily accessible. In line with general best practices for dataset creation, it is highly advisable to utilize established dataset labeling frameworks, such as Dataset Nutrition Labels [5], Data Cards [23], or Datasheets for Datasets [10]. These frameworks provide structured and standardized ways to present critical information about datasets, promoting transparency and ease of use. By incorporating these labeling frameworks, dataset creators can ensure that users are well-informed about the nature and characteristics of the data. These frameworks can easily be augmented with the aforementioned specific information relevant to XR user motion studies. This practice not only allows researchers to quickly understand the datasets without spending time and resources for tedious and error prone analyses, but also fosters a culture of clarity and accountability in data sharing within the kinematic research community. Implementing such comprehensive labeling approaches will significantly contribute to the rigor and reproducibility of future research in this field.

[GC9] Consider ethical implications. Collecting and sharing XR motion datasets research entails significant ethical considerations, particularly regarding participant privacy. As research has shown, motion data can inadvertently reveal personal information [17, 20, 21], making it crucial to implement protective measures. Even if users are fine with being openly recognized within the dataset collection study, they should be aware that they could be re-identified in different scenarios where they want to stay anonymous, just based on their motion data. Hence, researchers must account for informed consent, pseudonymize data as soon as possible, and comply with relevant data protection laws to safeguard participant privacy. Additionally, Data Use Agreements (DUAs) can regulate access, outlining specific conditions for data use and ensuring ethical handling.

6.2 Guidelines For Authors and Reviewers

Not only creators of motion datasets should follow best practices, but also other researchers working with these datasets. The following guidelines are intended as best practices for researchers working on publications that use motion datasets, and as checklist for reviewers evaluating submissions in this field.

[GAR1] Review and Exploration. For researchers engaging with XR user motion studies, a thorough understanding and exploration of key dataset specifications are imperative. These specifications, as outlined in previous sections, are crucial for accurate data interpretation and experimental reproducibility. If a dataset's documentation lacks these details, researchers should make efforts to acquire this information directly from the dataset creators or conduct their own analysis to determine these specifics. Furthermore, any such efforts and findings must be transparently disclosed in their publications.

[GAR2] Conversion. Researchers must document any conversions applied to the motion data. Detailed documentation of these conversions is essential for clarity and integrity of the research. It ensures that each step of data handling is accurately conveyed and that the data is not inadvertently distorted during the process. This transparency is critical for other researchers who may wish to replicate or build upon the work.

[GAR3] Alignment. In studies utilizing multiple datasets, it is essential to disclose the key differences between them and the measures taken to align each dataset. This includes aligning coordinate systems, normalizing units of measurement, standardizing rotation representations and frame timing. Researchers must clearly outline how they harmonized disparate datasets to ensure consistent and accurate analysis.

[GAR4] Critical Analysis of Results. Authors must critically analyze and potentially discuss how the specifications of the datasets could have influenced their results. This examination should consider whether the results might be skewed — either overly optimistic, such as when machine learning models overfit to dataset-specific signals, or overly pessimistic, due to erroneous preprocessing or misinterpretation of the data. Such a critical evaluation helps in contextualizing the findings and provides a more nuanced understanding of the study's implications and limitations.

[GAR5] Publication of Codebase. Publishing the codebase is a fundamental requirement. This includes code for data import, alignment, preprocessing, and analysis. Making the codebase publicly available allows for independent verification of the methodology and ensures that data handling has been executed correctly. It fosters transparency, reproducibility, and collaborative advancement in the field. Without access to code, it is impossible for reviewers and other researchers to validate, replicate, or extend the findings, thereby impeding scientific progress in kinematic research.

7 Conclusion

In this paper, we presented an in-depth analysis of existing XR motion datasets, revealing critical gaps in documentation and standardization practices.

The guidelines we proposed are intended to serve as a valuable resource for future creators of datasets, as well as authors and reviewers of publications that utilize such datasets. They offer a framework for concise documentation, rigorous discussion, and insightful feedback, ultimately enhancing the quality and integrity of research in this domain. Our aim is to foster a culture of transparency, consistency, and ethical responsibility in the collection, sharing, and utilization of motion data.

By introducing the XR Motion Dataset Catalogue and providing comprehensive guidelines for dataset creation and usage, we aim to address these gaps. Our work underscores the importance of meticulous documentation, standardized data formatting, and ethical considerations in the creation and sharing of motion datasets. We hope that our efforts lay the foundation for more accessible, reliable, and improved motion datasets in kinematic research.

Moving forward, we plan to continuously update the XR Motion Dataset Catalogue with emerging datasets. We invite researchers in the field to collaborate and contribute to this initiative, thereby enriching the catalogue and aiding the progression of kinematic research. Together, we hope to create a more standardized, accessible, and practical landscape for motion datasets in kinematic research and related fields.

Acknowledgments

The authors wish to thank Dawn Song, James F. O'Brien, Louis Rosenberg, and Bjoern Hartmann. This work has in parts been funded by the Minderoo Foundation, Meta Reality Labs, the National Science Foundation, the National Physical Science Consortium, the Fannie and John Hertz Foundation, and the Berkeley Center for Responsible, Decentralized Intelligence.

References

[1] Y. Abdrabou, L. Mecke, R. Rivu, S. Prange, Q. D. Nguyen, V. Voigt, F. Alt, and K. Pfeuffer. How Unique do we Move? Understanding the Human Body and Context Factors for User Identification. In *Mensch Und Computer 2023*, pp. 127–137. ACM, Rapperswil Switzerland, Sept. 2023. doi: 10.1145/3603555.3603574

[2] S. Agac and O. D. Incel. User Authentication and Identification on Smart Glasses with Motion Sensors. *SN Computer Science*, 4(6):761, Sept. 2023. doi: 10.1007/s42979-023-02202-4

[3] Apache Foundation. Apache Parquet.

[4] C. Chen, U. Jain, C. Schissler, S. V. A. Gari, Z. Al-Halah, V. K. Ithapu, P. Robinson, and K. Grauman. SoundSpaces: Audio-visual navigation in 3D environments. In A. Vedaldi, H. Bischof, T. Brox, and J.-M. Frahm, eds., *Computer Vision – ECCV 2020*, pp. 17–36. Springer International Publishing, Cham, 2020.

[5] K. S. Chmielinski, S. Newman, M. Taylor, J. Joseph, K. Thomas, J. Yurkofsky, and Y. C. Qiu. The Dataset Nutrition Label (2nd Gen): Leveraging Context to Mitigate Harms in Artificial Intelligence. In *Workshop on Dataset Curation and Security*, 2022.

[6] A. I. Dell, J. A. Bender, K. Branson, I. D. Couzin, G. G. de Polavieja, L. P. Noldus, A. Pérez-Escudero, P. Perona, A. D. Straw, M. Wikelski, and U. Brose. Automated image-based tracking and its application in ecology. *Trends in Ecology & Evolution*, 29(7):417–428, 2014. doi: 10 .1016/j.tree.2014.05.004

[7] M. S. Dennison, M. D'Zmura, A. V. Harrison, M. Lee, and A. J. Raglin. Improving motion sickness severity classification through multi-modal data fusion. In T. Pham, ed., *Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications*, p. 27. SPIE, Baltimore, United States, May 2019. doi: 10.1117/12.2519085

[8] M. S. Dennison, A. Z. Wisti, and M. D'Zmura. Use of physiological signals to predict cybersickness. *Displays*, 44:42–52, Sept. 2016. doi: 10.1016/j.displa.2016.07.002

[9] A. Garcia-Agundez, C. Reuter, H. Becker, R. Konrad, P. Caserman, A. Miede, and S. Göbel. Development of a Classifier to Determine Factors Causing Cybersickness in Virtual Reality Environments. *Games for Health Journal*, 8(6):439–444, Dec. 2019. doi: 10.1089/g4h.2019.0045

[10] T. Gebru, J. Morgenstern, B. Vecchione, J. Wortman Vaughan, H. Wallach, H. Daumé III, and K. Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, Dec. 2021.

[11] HDF5 Group. HDF5 Library and Format.

[12] E. Jain, L. Anthony, A. Aloba, A. Castonguay, I. Cuba, A. Shaw, and J. Woodward. Is the Motion of a Child Perceivably Different from the Motion of an Adult? *ACM Transactions on Applied Perception*, 13(4):1–17, July 2016. doi: 10.1145/2947616

[13] L. T. Kozlowski and J. E. Cutting. Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6):575–580, Nov. 1977. doi: 10.3758/BF03198740

[14] J. Liebers, M. Abdelaziz, and L. Mecke. Understanding user identification in virtual reality through behavioral biometrics and the effect of body normalization. In *Conference on Human Factors in Computing Systems - Proceedings*, pp. 1–11. Association for Computing Machinery, New York, NY, USA, May 2021. doi: 10.1145/3411764.3445528

[15] J. Liebers, S. Brockel, U. Gruenefeld, and S. Schneegass. Identifying Users by Their Hand Tracking Data in Augmented and Virtual Reality. *International Journal of Human–Computer Interaction*, pp. 1–16, Oct. 2022. doi: 10.1080/10447318.2022.2120845

[16] J. Liebers, C. Burschik, U. Gruenefeld, and S. Schneegass. Exploring the stability of behavioral biometrics in virtual reality in a remote field study: Towards implicit and continuous user identification through body movements: Towards implicit and continuous user identification through body movements. In *29th ACM Symposium on Virtual Reality Software and Technology*, VRST 2023. Association for Computing Machinery, New York, NY, USA, 2023. doi: 10.1145/3611659.3615696

[17] M. R. Miller, F. Herrera, H. Jun, J. A. Landay, and J. N. Bailenson. Personal identifiability of user tracking data during observation of 360-degree VR video. *Scientific Reports*, 10(1):1–10, 2020. doi: 10.1038/s41598-020-74486-y

[18] R. Miller, N. K. Banerjee, and S. Banerjee. Combining Real-World Constraints on User Behavior with Deep Neural Networks for Virtual Reality (VR) Biometrics. In *Proceedings - 2022 IEEE Conference on Virtual Reality and 3D User Interfaces, VR 2022*, pp. 409–418. Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/VR51125.2022.00060

[19] A. G. Moore, T. D. Do, N. Ruozzi, and R. P. McMahan. Identifying virtual reality users across domain-specific tasks: A systematic investigation of tracked features for assembly. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 396–404, 2023. doi: 10.1109/ISMAR59233.2023.00054

[20] V. Nair, W. Guo, J. Mattern, R. Wang, J. F. O'Brien, L. Rosenberg, and D. Song. Unique Identification of 50,000+ Virtual Reality Users from Head & Hand Motion Data. Feb. 2023.

[21] V. Nair, C. Rack, W. Guo, R. Wang, S. Li, B. Huang, A. Cull, J. F. O'Brien, M. Latoschik, L. Rosenberg, and D. Song. Inferring Private Personal Attributes of Virtual Reality Users from Head and Hand Motion Data, June 2023.

[22] F. E. Pollick, J. W. Kay, K. Heim, and R. Stringer. Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance*, 31(6):1247–1265, Dec. 2005. doi: 10.1037/0096-1523.31.6.1247

[23] M. Pushkarna, A. Zaldivar, and O. Kjartansson. Data Cards: Purposeful and Transparent Dataset Documentation for Responsible AI. In *2022 ACM Conference on Fairness, Accountability, and Transparency*, pp. 1776–1826. ACM, Seoul Republic of Korea, June 2022. doi: 10.1145/3531146.3533231

[24] C. Rack, T. Fernando, M. Yalcin, A. Hotho, and M. E. Latoschik. Who is Alyx? A new behavioral biometric dataset for user identification

in XR. *Frontiers in Virtual Reality*, 4, 2023. doi: 10.3389/frvir.2023.1272234

[25] C. Rack, A. Hotho, and M. E. Latoschik. Comparison of Data Encodings and Machine Learning Architectures for User Identification on Arbitrary Motion Sequences. In *2022 IEEE International Conference on Artificial Intelligence and Virtual Reality, AIVR 2022*. IEEE, 2022.

[26] T. Steinmetzer, I. Bönninger, M. Reckhardt, F. Reinhardt, D. Erk, and C. M. Travieso. Comparison of algorithms and classifiers for stride detection using wearables. *Neural Computing and Applications*, 6, 2019. doi: 10.1007/s00521-019-04384-6

[27] Vivek Nair, Wenbo Guo, Rui Wang, James F. O'Brien, Louis Rosenberg, and Dawn Song. Berkeley Open Extended Reality Recording Dataset 2023 (BOXRR-23), 2023. doi: 10.25350/B5NP4V

[28] E. Wen, C. Gupta, P. Sasikumar, M. Billinghurst, J. Wilmott, E. Skow, A. Dey, and S. Nanayakkara. VR.net: A Real-world Dataset for Virtual Reality Motion Sickness Research, June 2023.