m-body.ai

Open Source Data And Tools For Generative Animation

GENERATOR at Sheridan
Screen Industries Research & Training Centre

CRIN
Centre de développement et de recherche en intelligence numérique

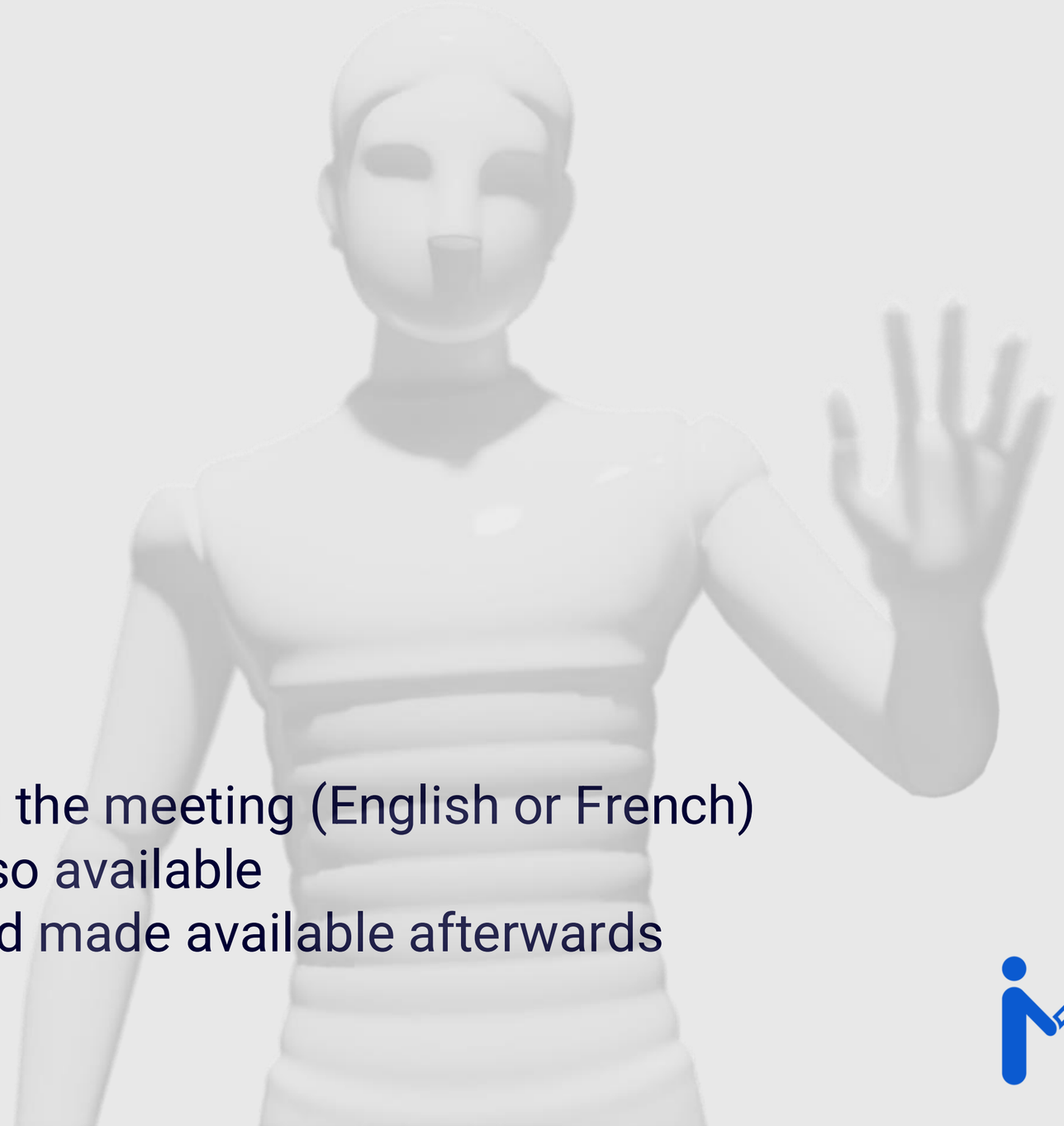
DURHAM COLLEGE
SUCCESS MATTERS

mrcstudio
MIXED-REALITY CAPTURE STUDIO

llio

AGENDA

1. Project and Team Introduction
2. Dataset Capture
3. Dataset Processing
4. Generative Animation Overview
5. Open-source Software Tools
6. Collaboration
7. Work conducted already
8. Q & A



- Post questions in the chat during the meeting (English or French)
- French version of the slides is also available
- Presentation will be recorded, and made available afterwards

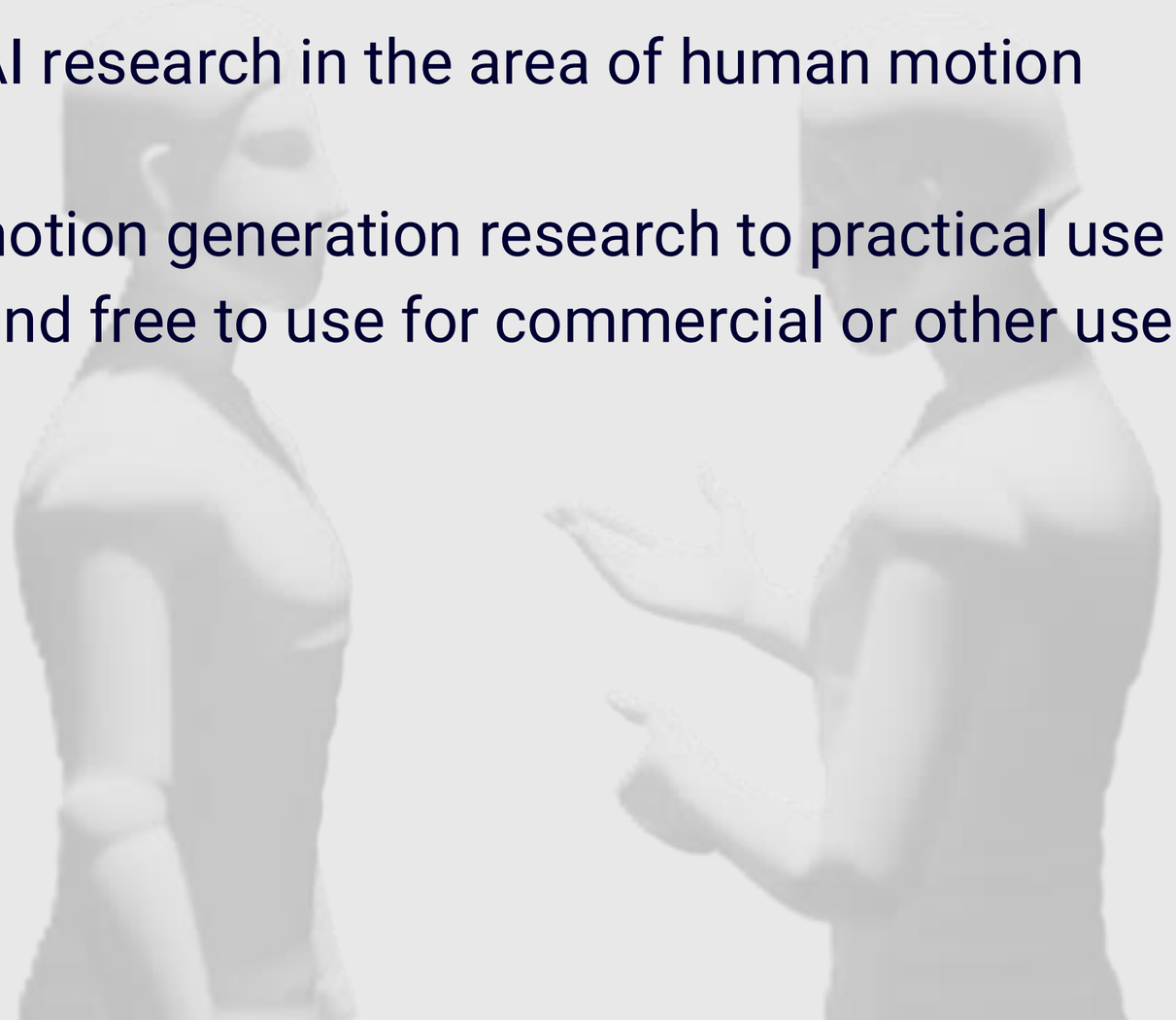
M-body.ai is a bilingual applied research project aiming to accelerate the advancement and adoption of generative human performance technologies.

M-body.ai is federally-funded by the National Sciences and Engineering Research Council of Canada (NSERC).

M-BODY.AI GOALS

What are the main goals of the M-body.ai project?

- Support the advancement of AI research in the area of human motion generation
- Make it easier to put human motion generation research to practical use
- All outputs publicly available and free to use for commercial or other uses



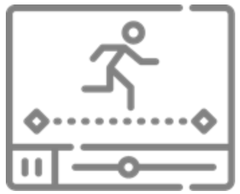
M-BODY.AI DELIVERABLES

For Academic and Applied Researchers

A free, novel **dataset of multi-modal, multi-agent interactions** to support the training of various ML models

For Animation Tool Developers

Open-source **software systems** to simplify the integration of generative character performance models into industry-standard content creation tools



For Animators

Open-source, reference **generative character animation tools** enabling animators to improve the efficiency and quality of their daily work

RESEARCH TEAM

GENERATOR at Sheridan
Screen Industries Research & Training Centre

Sheridan College - SIRT
Screen Industries Research and Training Centre

CDRIN
Centre de développement et de recherche en intelligence numérique

Cégep de Matane - CDRIN
Le Centre de Développement et de Recherche en Intelligence Numérique

DURHAM COLLEGE SUCCESS MATTERS | mrcstudio MIXED-REALITY CAPTURE STUDIO

Durham College - MRC, AI Hub
Mixed Reality Capture Studio

llio

Cégep de Rivière-du-Loup - LLio
Le Laboratoire en Innovation Ouverte

Sheridan College's Screen Industries Research and Training Centre (SIRT)

Empowering Ontario's screen industries with cutting-edge research and technology as a federally funded Technology Access Centre.

- Pinewood Toronto Studios in Toronto, Ontario
- Targeting the screen industries, focusing on the film, television, and interactive media sectors
- Supported by a dedicated team of 16 experts from various domains

S GENERATOR at Sheridan

Screen Industries Research & Training Centre

Sheridan College

SIRT – Screen Industries Research and Training Centre

Virtual Production

- Integrating cutting-edge motion capture systems and real-time technology to create real-time visuals on set

Animation and Virtual Humans

- Developing digital human pipelines and workflows to enhance character interaction and engagement

Software Development

- Integrating and developing software solutions for media and entertainment (interactive, Mixed Reality, and film)

In the M-body project:

- Multi-modal **motion capture** pipeline using cutting-edge **IR**-based technology, **electromagnetic** field, and **RGBD** motion capture systems.
- **Character architecture** and **standards** for **production-quality animation**, designed for machine learning **datasets**
- **Systems architecture** and implementation of **software tools** and plugins

MRC - Mixed-Reality Capture Studio

AI Hub - Applied Research in Artificial Intelligence for Business Solutions

Business insights while providing intelligent and autonomous solutions

- Oshawa, ON
- Targeting Immersive media, Indie Game Studios, Training & Simulation & Data rich sectors (Health, Finance, marketing, Supply Chain).
- MRC - 2 PIs, 3 research associates / AI Hub - 4 PIs, 6 research associates

DURHAM COLLEGE SUCCESS MATTERS | mrcstudio MIXED-REALITY CAPTURE STUDIO | AI/HUB

Durham College

MRC - Mixed-Reality Capture Studio

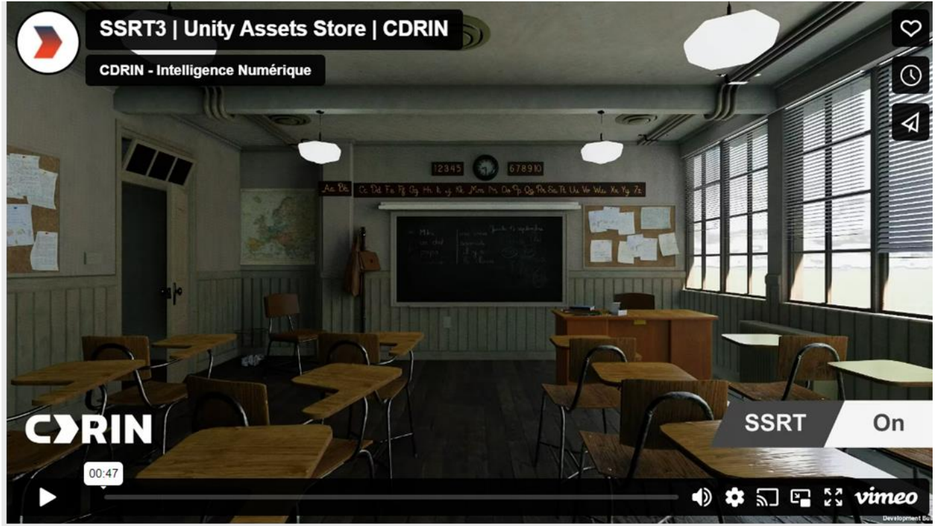
AI Hub - Applied Research in Artificial Intelligence for Business Solutions

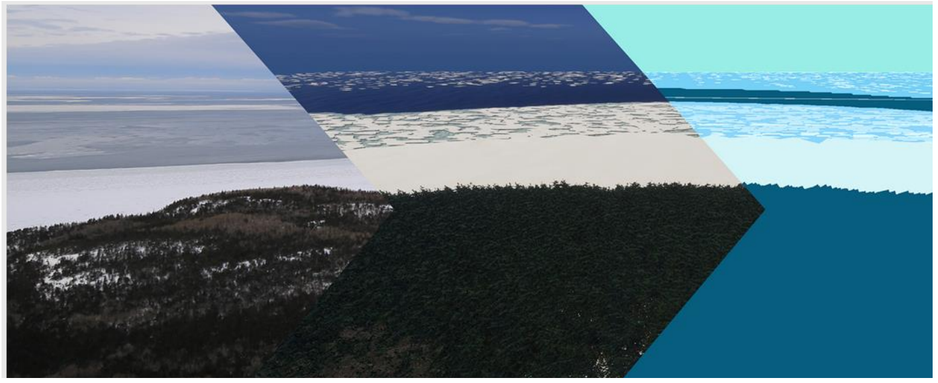
Training the next generation of talent with our paid work-integrated learning opportunities

- offers industry partners access to technical expertise
- state-of-the-art facilities and platforms
- student talent

In the M-body project:

- **Vision based Motion Capture** with rgb and ai system
- **Machine Learning** and Patterning
- Model testing and implementation with **artificial Intelligence**
- **student** integration and management





CDRIN

Promoting digital intelligence for creative, cultural, education, and health sectors, driving Quebec's economic growth through R&D and collaboration.

- Matane, Montreal, Quebec City
- Targeting the entertainment sector including video games and animation. (Other areas: environment, health and education)
- We are a dedicated team of 20 people (AI researchers, R&D programmers, 3D technical artist, and more)

CDRIN
Centre de développement
et de recherche en
intelligence numérique

Cégep de Matane

CDRIN- Le Centre de Développement et de Recherche en Intelligence Numérique

Offering:

- Technology Monitoring & Foresight
- Proof of Concept, Prototyping
- Implementation
- Competence Transfer & Specialized Training

In the M-body project

- **ML/AI**
  - Technology monitoring
  - AI model selection
  - Data processing
  - Model evaluation
  - Development

LLio - Le Laboratoire en Innovation Ouvert

Foster the adoption of open innovation practices involving the user

- Affiliated to Cégep de Rivière-du-Loup
- College transfer center in innovative social practice
- 18 people, a diverse team of researchers, facilitators and designers

Cégep de Rivière-du-Loup

LLio - Le Laboratoire en innovation ouverte

- Enhance the capacity to innovate
- Bring together several stakeholders around a common challenge
- Design thinking, design fiction and systems thinking
- Community governance and community of practice

In the M-body project:

- Design center around the **users**
- Consult the **community** and implement tools to **mobilize** it around the project
- **Licences**
- A **student research** and **creation** project: IdeAs

M-body.ai
Dataset Capture

DATASET CHALLENGES

Why is additional human conversational gesture data required?

- **Low quality** and **low diversity** of existing animation datasets
- Lack of **commercially** usable animation datasets
- **Inconsistent** body geometry (proportion vs normalization)
- **Inaccurate** hand and finger capture
- Limited channels of data

Human Motion Dataset

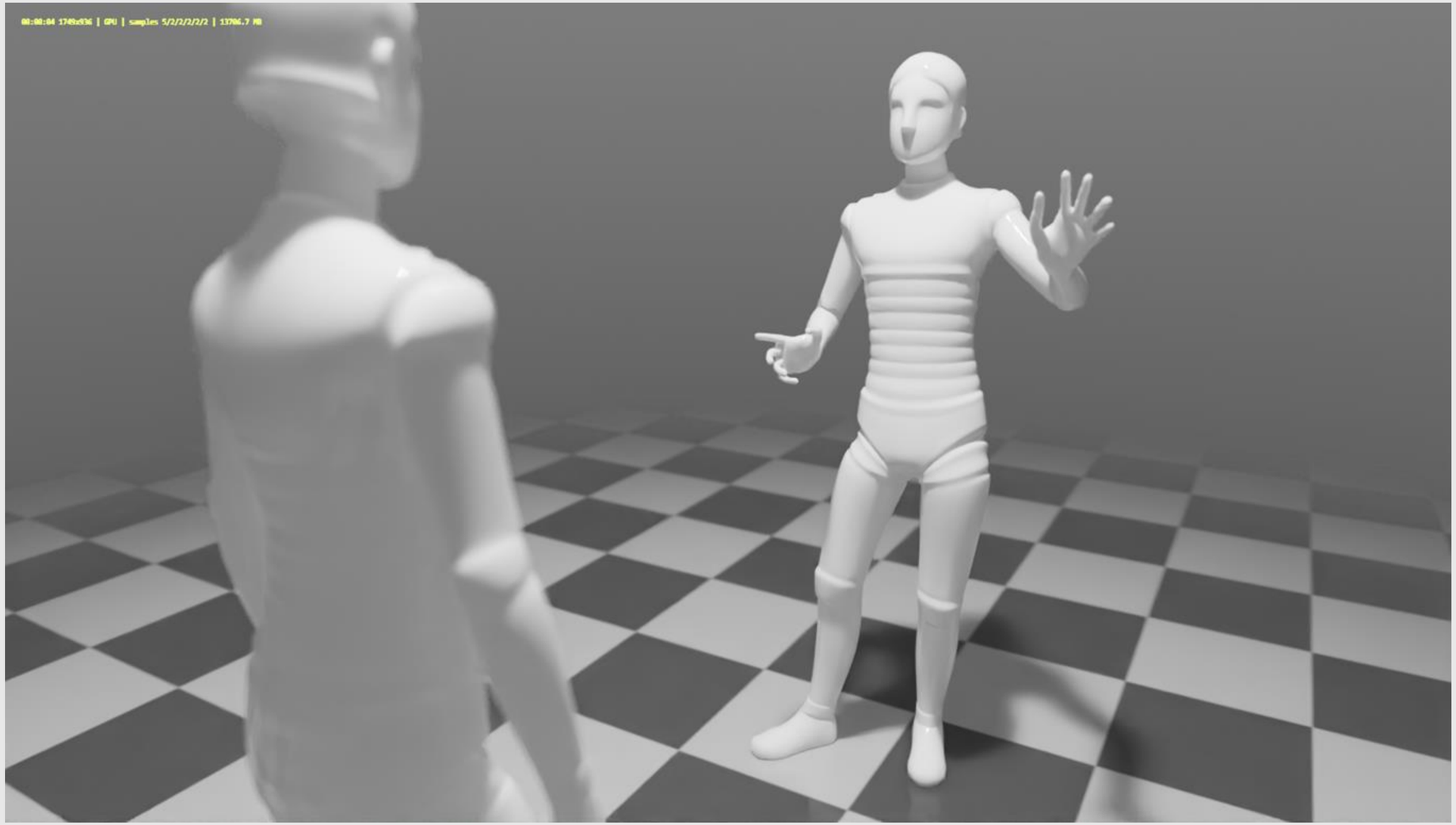
Planned Dataset Specifications

- At least 20 unique performers - 10 male, 10 female, diverse ethnicity
- 30 hours of 2-agent conversational animation (60 hours single-agent)
- Scenario: two agents standing and having general conversations

Datastreams included

- **Skeletal** body animation, including **hands**/**fingers**/**head**
- **Spoken audio** for each **separate** agent
- **Timed transcript** of spoken audio
- **Raw facial** performance capture (video and audio)
- **Reference** video
- **Topologically consistent** body geometry

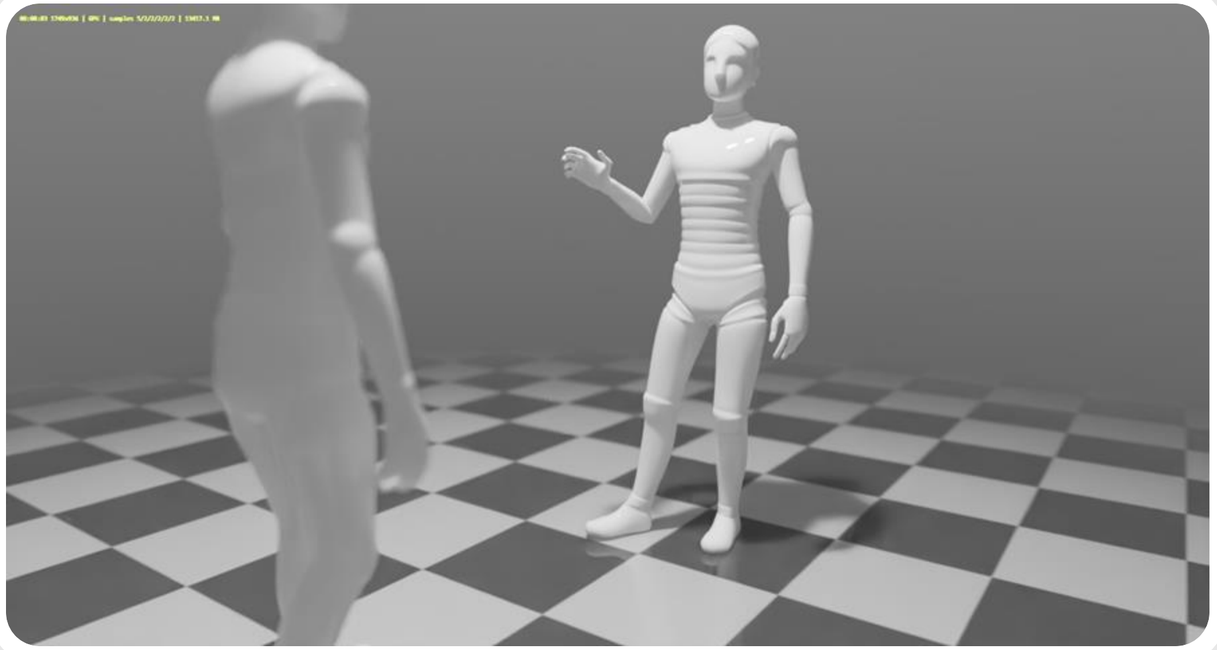
HUMAN MOTION DATASET - SAMPLE



HUMAN MOTION DATASET - SAMPLE CAPTURE

Raw Footage

3D Digital Double



HUMAN MOTION DATA - CAPTURE METHOD

Motion Capture (Full body including hand tracking)

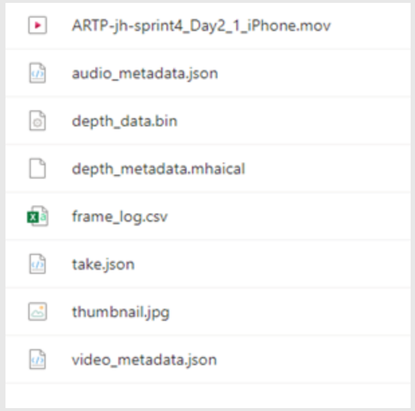
- Optitrack Motive (44 camera passive system), Manus Gloves (Hand Tracking)
- Captury (16 camera markerless mocap system)

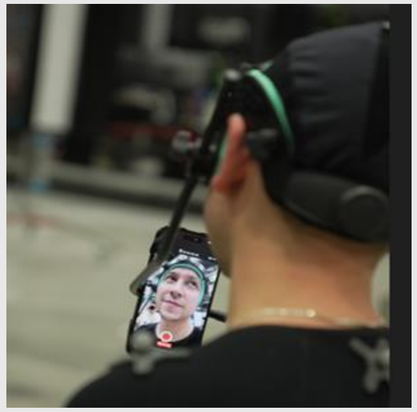
Facial Capture

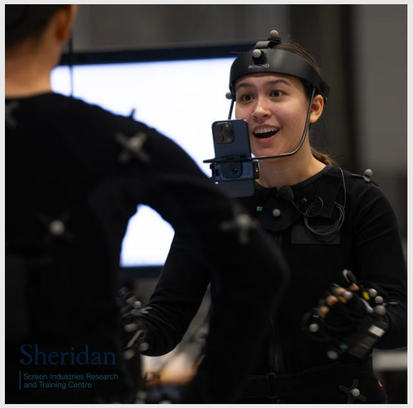
- iPhone (RGBD)

Audio + Video

- Reference Video, Audio captured with shotgun & lavalier mic (discrete channels)







M-body.ai
Dataset Processing

DATASET PROCESSING CHALLENGES

Consistent Animation Problems being solved through our processing

Posture correction

- usually from retargeting to skeletons with different proportions
- incorrect solves from motion capture
- incorrect marker matching, sync issues, Euler angle solve

Foot sliding

- micro movements correction
- geometry intersection with ground, improper foot roll

Root shifting (inertia systems)

- vertical positioning is off multi level, or just not calibrated correctly)

Noise

- hidden markers, joint axis flipping

Geometry Intersection

- accurate body proportions to account for accurate positions

CHARACTER ARCHITECTURE

Improve data storage in a skeleton hierarchy to boost machine learning analysis of natural motion in conversations.

Production quality System

- procedural character pipeline utilizing **standardized Topology**, **Rig** and **Controls**

Accurate proportions

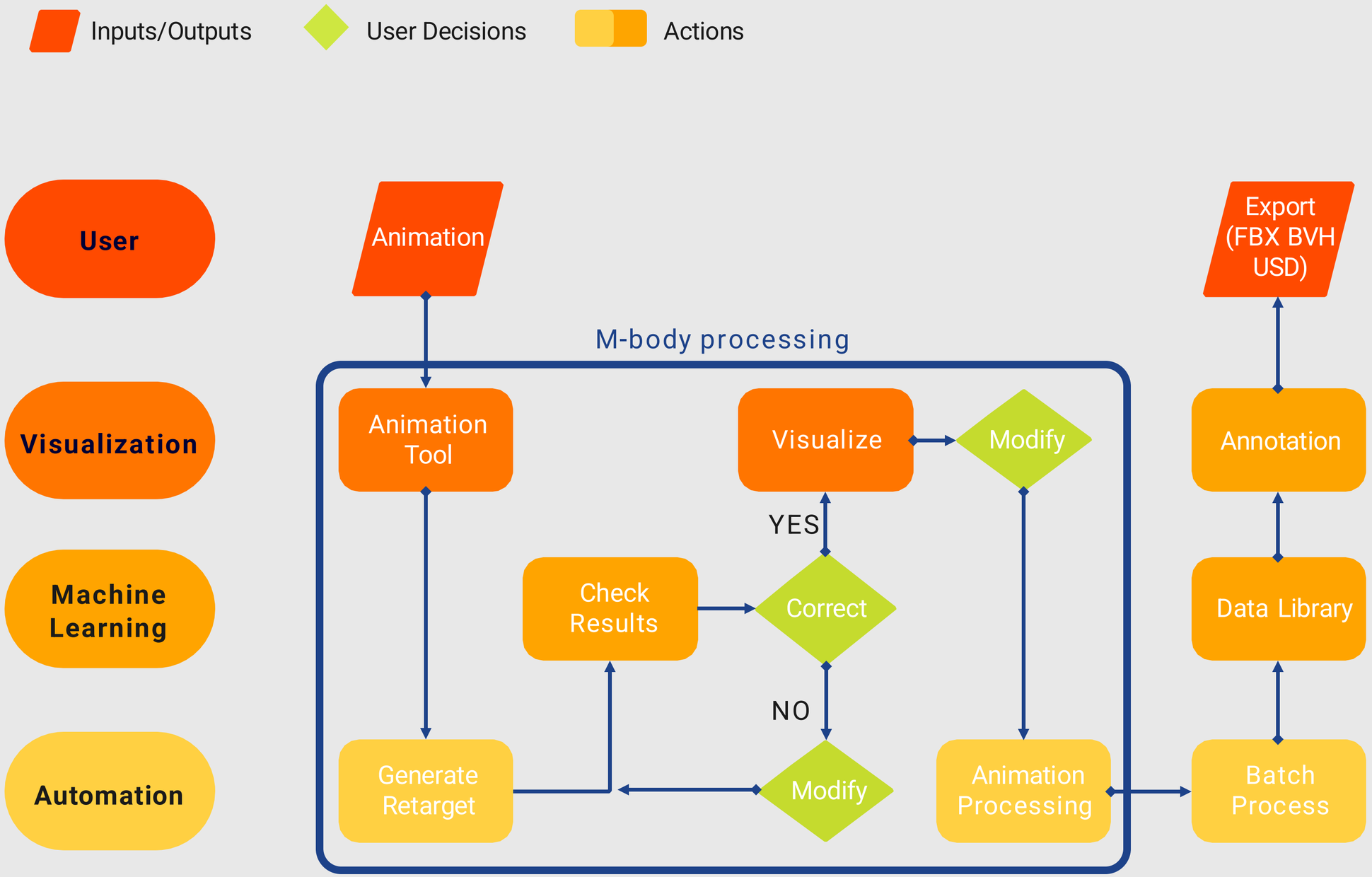
- captures **precise human movements** and variations in conversational actions between individuals

Error Mitigation

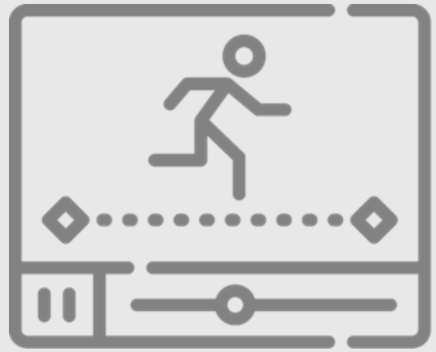
- automated **batch processing** to minimize human error

Retargeting System

- retargeting to characters with different skeletal proportions and structures
- Ability to **modify** and **control** characters animation **post generation**


Inputs/Outputs
User Decisions
Actions
User
Visualization
Machine Learning
Automation
Animation
M-body processing
Animation Tool
Visualize
Modify
YES
Check Results
Correct
NO
Generate Retarget
Modify
Animation Processing
Export (FBX BVH USD)
Annotation
Data Library
Batch Process

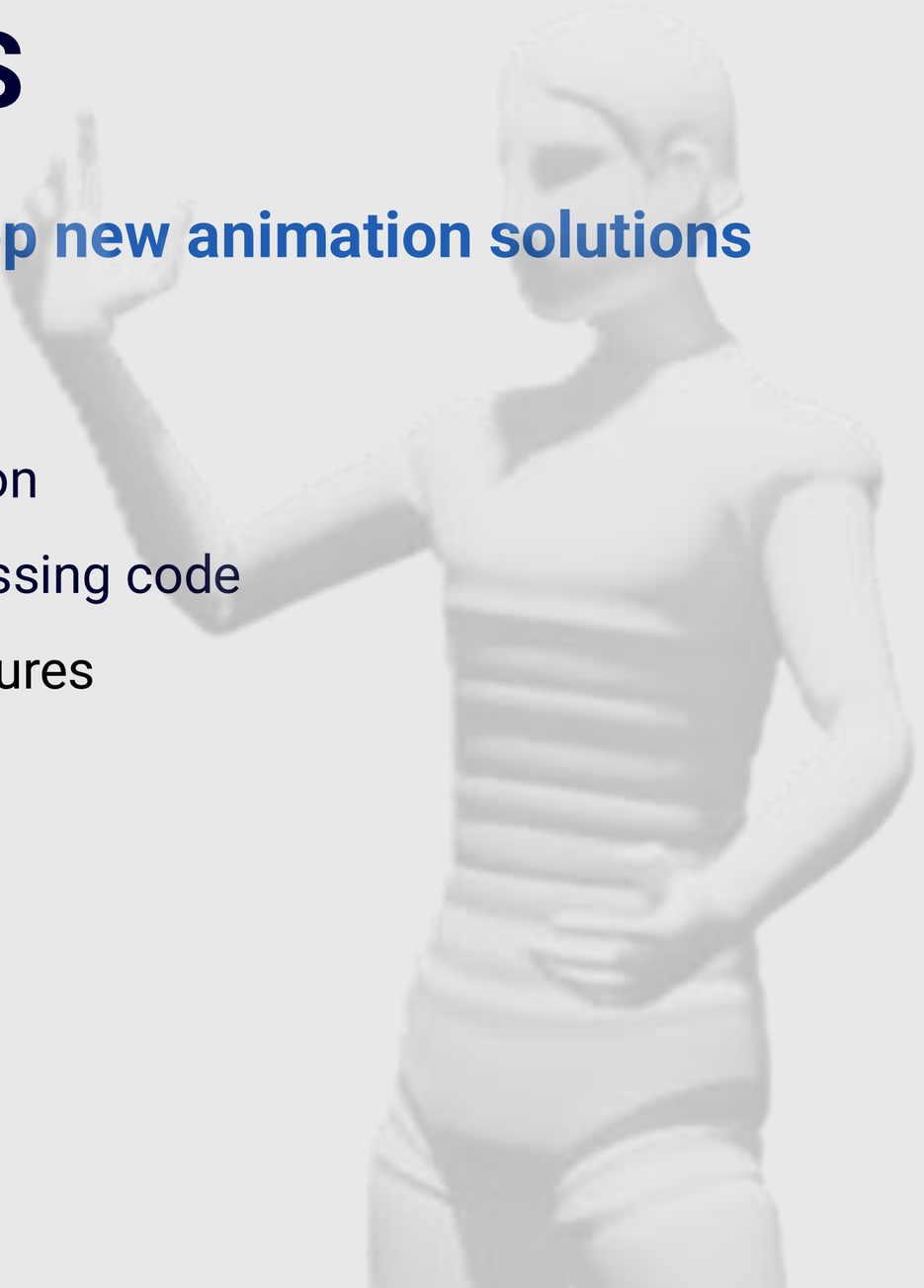
M-body.ai
Generative Animation



RESEARCH CHALLENGES

Why is it difficult for researchers to develop new animation solutions utilizing ML?

- Difficulty **validating** solutions for use in production
- Requires development of **custom** data pre-processing code
- **Idiosyncratic** and complex nature of human gestures
- Small and non-diverse **datasets**
- **Contextual** integration
- Challenges in **evaluation**



ML – GENERATIVE ANIMATION

Use cases

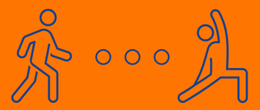
Motion synthesis

Inputs:

- Text
- Audio (speech, music)
- Character Identity or Style
- Motion Category
- Initial pose
- Scene, Objects

Motion Completion /In-betweening

Generating intermediate poses between the start pose and the end pose.



Style Transfer

Use a style motion learned by the model, and a content motion unseen by the model.

Crowd Animation

Generating animation for a large group of agents.

Rigging for Articulated Characters

Given a 3D model, predict a skeleton that matches the animator expectations in joint placement and topology

ML – SAMPLE MODELS

A large-scale evaluation of gesture generation models in monadic and dyadic settings.

GENEA Challenge 2023

DiffuseStyleGesture+

- One of the top-tier models in the 2023 challenge.
- MIT-License
- obtained the reproducibility award by GENEA committee

ZeroEGGs

- One of the top-tier models in the 2022 challenge.
- A variational framework for speech-driven gesture generation with zero-shot style control by example.

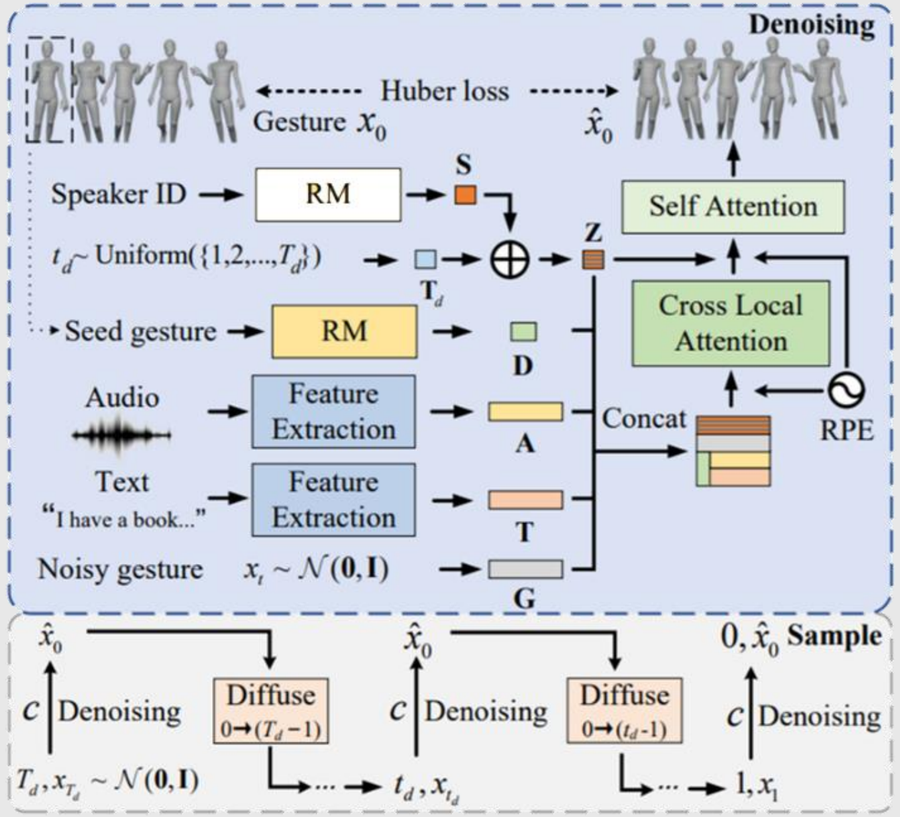
ML – DIFFUSESTYLEGESTURE+

Diffusion Process

- At each step, a noisy gesture sequence xt conditioning on c (including seed gesture, audio, ID and text) are fed into this model

Feature Extraction

- **Gesture:** Position, velocity, acceleration, rotation matrix, rotational angular velocity, and rotation angular acceleration of each joint in each frame.
- **Audio:** combination of e MFCC, Mel Spectrum, Pitch, Energy, WavLM, and Onsets.
- **Speaker ID:** one-hot vectors
- **Text:** The 300-D word embeddings using Fast Text



source: https://arxiv.org/abs/2308.13879

ML – COMMON EVALUATION METRICS

Qualitative Metrics

- Example: crowdsourcing subjective evaluation of the generated gestures through user studies (e.g. asking people to rate an animation based on different factors and calculate the mean opinion score)

Quantitative Metrics

Examples:

- FID/FGD: Fréchet Inception/Gesture Distance
- R-Precision
- Diversity
- Multimodality
- Multi-Modal Distance

ML – EVALUATING OUR DATASET

What have we done so far?

- Train DSG+ using our initial motion capture (in TWH format) + TWH vs. only TWH

What next?

- Correct the errors in the TWH data (realistic joint rotation, noise removal, foot sliding and weight correction) and combine them with our initial mocap data. We expect that by retaining the model on this data, we will get Better quality animations.
- Training DSG+ using our new dataset and perform qualitative and quantitative comparisons. Again, since our data benefits from the above error corrections, we expect to get higher quality animation generated by the model.
- Verifying if the model is trained well using skeletal data with accurate proportions (ie. Bone lengths differ from person to person)

M-body.ai
Software Tools

TOOL DEVELOPER CHALLENGES

Why is it difficult for developers to create animation tools leveraging ML?

- Requires **expertise** in multiple domains: tool development and ML
- Need to implement **custom** data type conversion
- Mixture of programming **languages/environments** (C#, C++, Python)
- Complexity of setting up **dependencies** for ML models
- Different features exist in different tools, leading to **complex workflows**

SOFTWARE GOALS

What are the goals for our software tools?

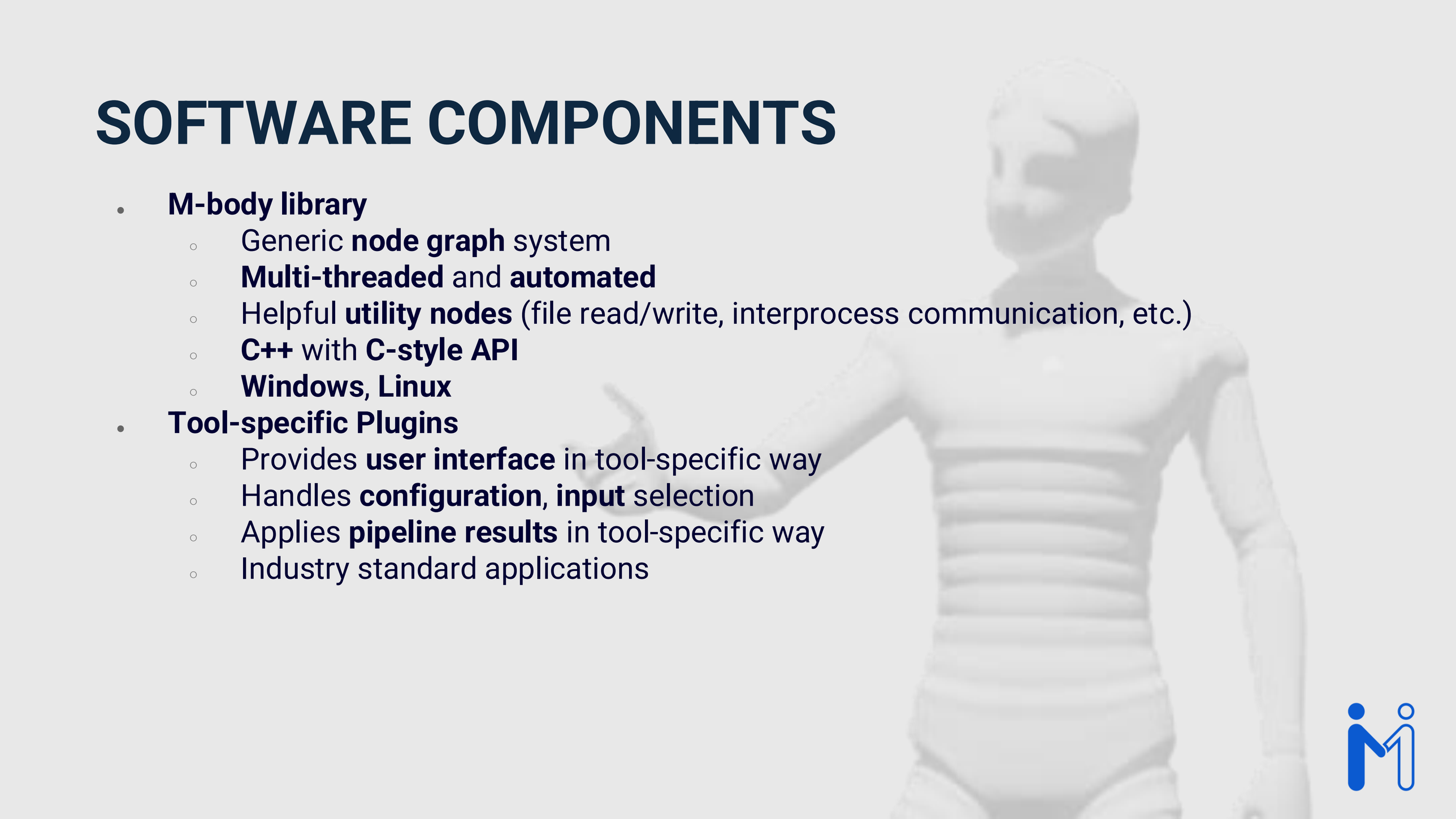
- Ease integration of generative animation models into tools/pipelines
  - Unity, Unreal, Maya, etc.
- Access to generative models from **industry-standard** tools
- Open-source, free, **commercially usable**

ML INTEGRATION APPROACHES

- **ONNX - Open Neural Network Exchange**
  - Converts ML models into common set of operators
  - GOOD: Reduces external dependencies
  - BAD: Issues converting models and data pre-processing to ONNX
- **Embedded Python - Pybind**
  - Python runtime integrated into application
  - GOOD: Reduces some external dependencies
  - BAD: Requires Python wrapper, execution may be inconsistent
- **OS Command Line Execution**
  - Execute command line Python from code
  - GOOD: Use Python model and data pre-processing as-is
  - BAD: External Dependencies (Python, libraries)

SOFTWARE COMPONENTS

- **M-body library**
  - Generic **node graph** system
  - **Multi-threaded** and **automated**
  - Helpful **utility nodes** (file read/write, interprocess communication, etc.)
  - **C++** with **C-style API**
  - **Windows**, **Linux**
- **Tool-specific Plugins**
  - Provides **user interface** in tool-specific way
  - Handles **configuration**, **input** selection
  - Applies **pipeline results** in tool-specific way
  - Industry standard applications



M-body.ai Collaboration

OPEN SOURCE HUB

GitHub

Collaborative CODE management

m-body

Web portal and CRM

Discord

Communication between developers and the community

OPEN SOURCE - LICENSES

- **Four licenses** that allows for **commercial use**
- Well-known in the industry and for research purposes
- Easy to use
- Two types of licenses :
  - **Code** and **software systems** licenses: MIT (expat)
  - **Dataset** and **documentation**: CC BY
- Dataset ethical challenges: human body and what’s taking into accounts
  - Moral rights
  - Personal image integrity
  - Personality rights

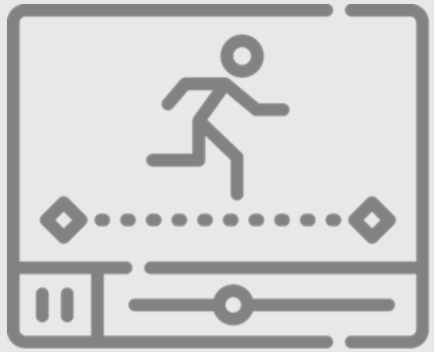
OPEN SOURCE - COLLABORATION

Stay tuned - Keep up to date

Experiment - Register as one of the partners interested in taking part in demonstration sessions, interviews and user tests to inform research with your needs and requirement.

Co-development - Co-developer community once the open source coding infrastructure will be in place.

M-body.ai
Work Conducted



WORK CONDUCTED

What work has been conducted already?

- Conducted **background research** into existing literature, tools, and datasets related to generative animation
- Consulted with **industry** and **academia** to identify pain-points
- Captured preliminary multi-agent **animation dataset**
- Set up existing **generative** animation **models** to validate our captured data
- Implemented **prototype** software to help integrate ML models into commonly used tools (Unity, Unreal, Maya)

TIMELINE

2024

Q1

Interviews with industry partners to refine requirements

Q2

Release of open source project for first prototype generative character animation tool

Release of sample dataset of multi-agent conversational animation

Q3

Demonstration of prototype tool and dataset to gather feedback

Testing and validation of prototypes with industry to steer project direction

Q4

Release of open source project for beta version of generative character animation tool

Release of open dataset of multi-agent conversational animations

2025

Q1

Work with partners to integ rate into workflows and additional improvements and features. Finalize maintenance plans from open source projects and outputs

Q2

Release of open source project for production generative character animation tool

Release of final dataset of multi-agent conversational animations

QUESTIONS?
BE PART OF THE PROJECT!

Contact us to experiment or co-develop

mbody@sheridancollege.ca

Follow what’s going on!

m-body.ai
www.linkedin.com/company/m-body-ai/
www.facebook.com/mbodyAI