AI4ME

AI for Personalised Media Experiences

A (quick) Overview of the AI4ME project and compute requirements for the COIN Research Group

9 November 2023

Rajiv Ramdhany (BBC)
Nick Race (Lancaster University)
Daniel King (Lancaster University)

University of Surrey · BBC · Lancaster University

AI4ME
AI for Personalised Media Experiences

Who are the BBC?

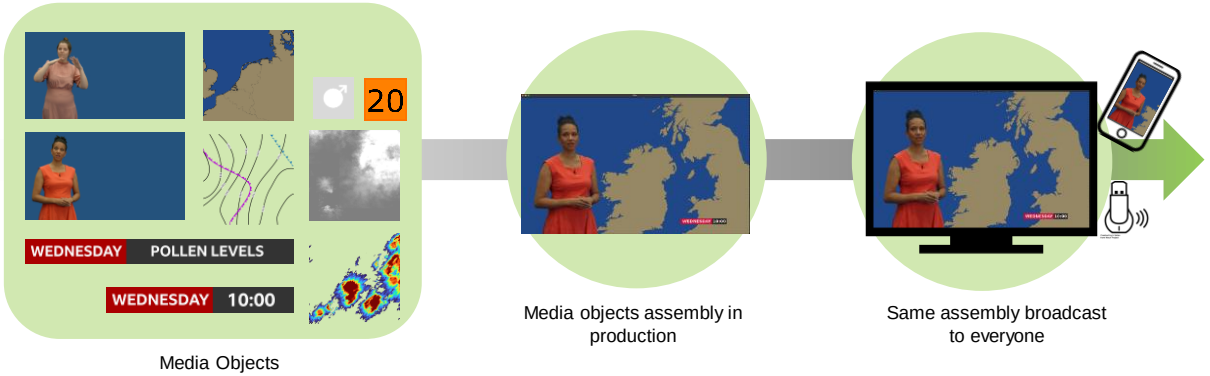
- The British Broadcasting Corporation, is the United Kingdom's public service broadcaster.
  - Founded in 1922; one of the oldest and largest broadcasters in the world.
  - **BBC R&D**: research and innovation in on every aspect of the broadcast chain, from Audiences, Production and Distribution
- Some of the major technology innovations and contributions by the BBC:
  - **First Public Television Broadcast** in the world (1936).
  - **The first regular color television broadcasts** in the world (1967).
  - **BBC Microcomputer** for early computer education of many British children and adults (1980s)
  - First **Digital Satellite Television** service in the world **(1990)**
  - **First High-Definition Broadcasting**
  - **BBC iPlayer**: First major **catch-up TV service** in the world **(**2007); set the standard for similar platforms.
  - Leading testing and adoption of **4K and Ultra-High Definition broadcast**
  - Producing immersive **VR and Immersive Content/Experiences**
  - **Binaural Audio** for immersive spatial audio experiences for headphone listeners
  - **BBC micro:bit** - entry-level programmable computer for children
  - Low-latency dynamic adaptive streaming over HTTP
  - Dynamic adaptive streaming over QUIC
  - **Object-based Media** trials
  - Hybrid Log Gamma (HLG) system for **High Dynamic Range** (HDR),

BBC

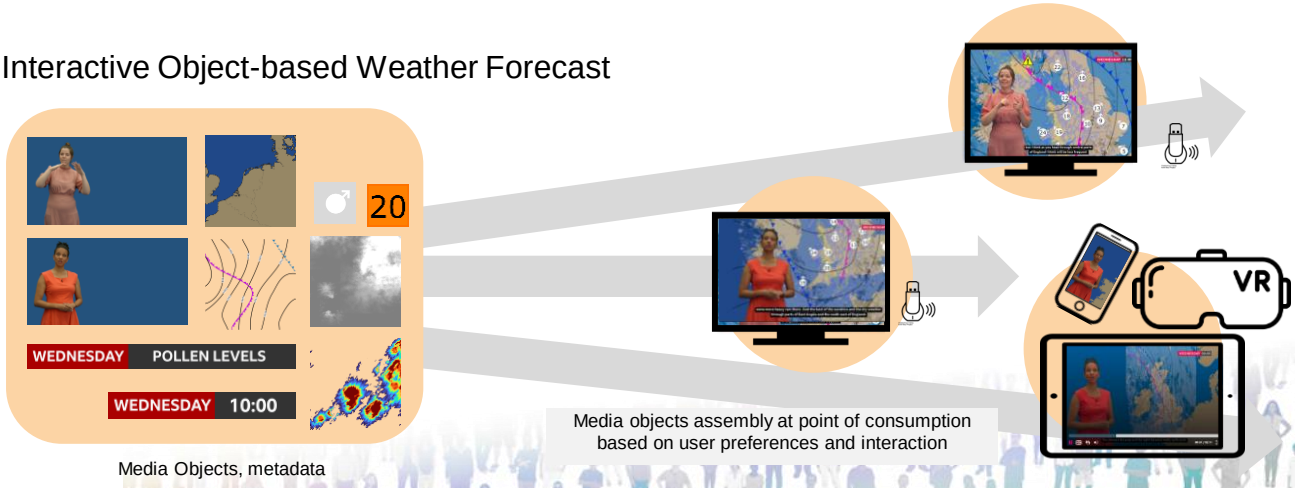
Object-Based Media Primer

- Object-Based Media allows the content of programmes to change according to the requirements of each individual audience member.
- Media Objects – assets that make the content
  - Audio, video, graphics, captions, 3D models, data streams, textures
  - Code: shaders

Traditional linear broadcast



Interactive Object-based Weather Forecast



AI4ME Project

Future Personalised Object-Based Media Experiences Delivered at Scale Anywhere

Audio/
Video

Stream 2

Object Capture and Representation

Objects
+ Scene
Metadata

Stream 3

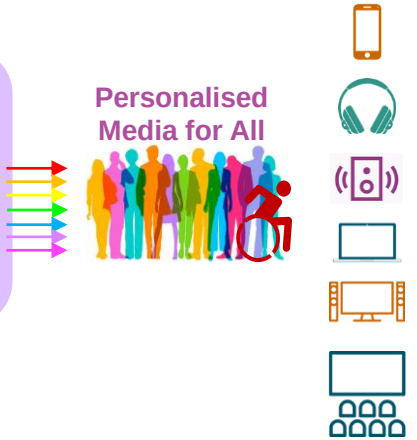
Object-Based Production

Objects
+ Production
Metadata

Stream 4

Compute-Aware Network Delivery

Personalised Media for All



Intelligent Personalised Media Delivery at Scale

- AI-powered technologies to deliver personalised media content cost-effectively and efficiently to mass audiences
- Technologies for distributed execution and intelligent placement of media applications and rendering services

AI4ME Service Resources

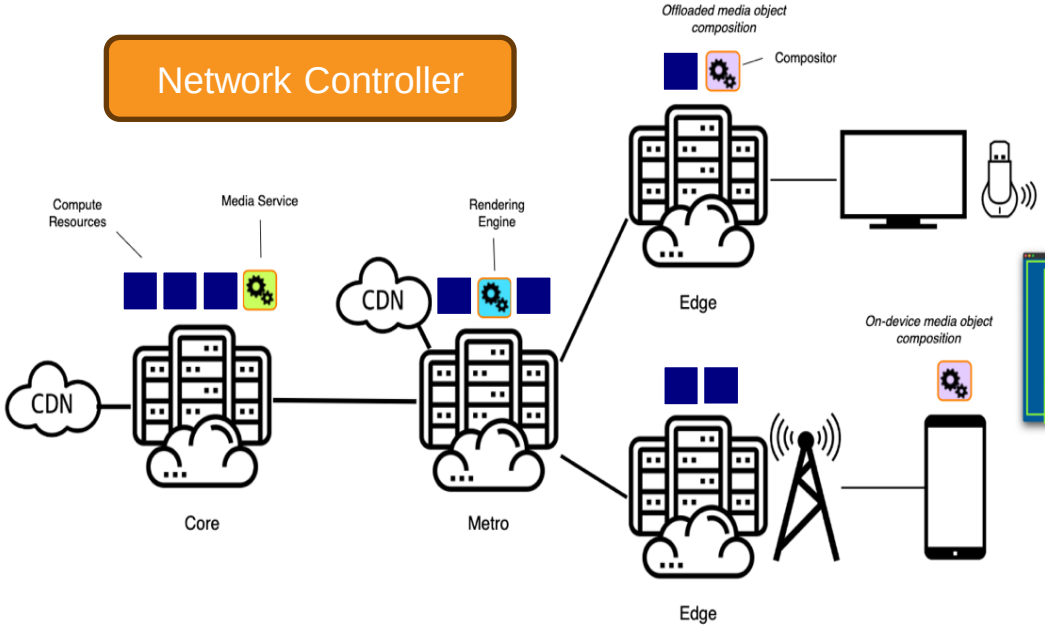
- Streaming AI4ME OBM to many users requires a combination of hardware and software resources to ensure a smooth and high-quality streaming experience
- The key types of software resources that are used for a service, include:
  - Content Servers (Cloud Containers): for OBM processing, generating additional content (titles, AR) and steaming
  - Content Storage: High-capacity storage to store UHD OBM video files – over a distributed file system
  - Content Management System (CMS): A CMS to organise and manage OBM content, including metadata, thumbnails, and access controls
  - CDN Components: Schedulers, load balancers, cataloguers, caching nodes
  - GPU/CPU Resources: Rendering engine and transcoding
  - Memory: Significant RAM is necessary to buffer and process video streams efficiently, at scale
  - CDN Components: Schedulers, load balancers, caching nodes
- <u>The computational and network optimisation challenges are legion.</u>

Network Architecture and Functions

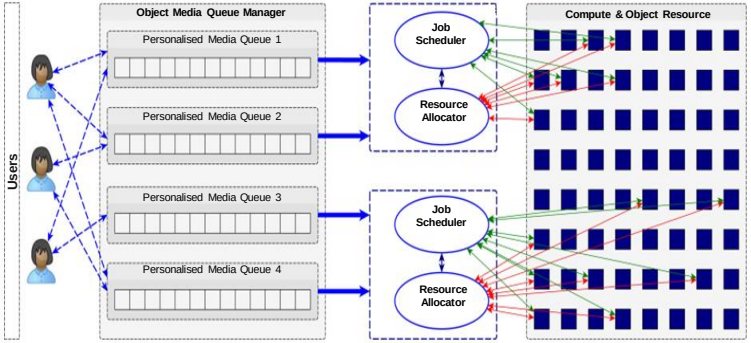
AI4ME High-level Network Architecture

- As jobs are requested, the computational demands are also identified, and users are assigned to AIM4ME service instances
- Cloud orchestrator will direct traffic to an available Service Instance, optimising OBM content delivery and reducing user latency

Network Controller



AI4ME Cloud Orchestrator



Applicability of AI/ML to AI4ME

- **Investigating two AI Use Cases**
  - Combining Cloud Infrastructure Orchestration and Network Optimisation for AI4ME Services
  - AI4ME Service and Network Fault Detection, Prediction and Resolution
- **However, AI Challenges for solving complex network problems include:**
  - Data scarcity and quality. AI algorithms require large amounts of high-quality data to train effectively
    - Network data can be scarce, especially for complex networks with many nodes and edges
  - Developing AI models that can learn from sparse and noisy network data
  - Using AI models that are interpretable and explainable
    - Can we explain why they have made a particular prediction about network performance?
  - Developing AI models that are computationally efficient and scalable
    - The AI models would need to run in real time on the large complex AI4ME infrastructure
  - Developing AI models that are secure and privacy-preserving.
    - Identifying AI models that can be trained on sensitive network data without revealing any confidential information of AI4ME users

Questions/Discussion

Rajiv Ramdhany (BBC) - Rajiv.Ramdhany@bbc.co.uk

Nick Race (Lancaster University) - n.race@lancaster.ac.uk

Daniel King (Lancaster University) – d.king@lancaster.ac.uk

Tomorrow (Nov 10, 2023) Rajiv will present a more detailed overview of AI4ME network challenges in the IETF CATS WG.

ai4me.surrey.ac.uk

AI4ME
AI for Personalised Media Experiences