

Web-based Machine Learning for Gaming

Josh Meyer
&
Lindy Rauchenstein
{ Lead Scientists }

Artie[™]

Artie, Inc.

Instant Entertainment
Conversational
Web-based
Mobile-first



ML: {Implementation Trends}

Traditional

- Small models → On-device
- Big models → Server-side

New

- Shrink the big models
- Everything → on-device

ML: {Implementation Realities}

Web-based + On-device ==



Small model == small functionality

Slow download == high churn

ML: {Artie's Implementation}

Artie = { Language → Server
 Vision → Device }

Instant + quality + private

Artie: {Speech + NLP}

- **Artie's DeepServe**
 - **Built on Mozilla Voice STT**
 - **Custom NLP Solutions**

Artie: {Vision Constraints}

- Privacy
- Low latency
- Compatible with Project Tiny
- No dynamic memory
- Total download size 4 MB

Artie: {Vision Implementation}

TF Lite for Microcontrollers

- **Compatible with Project Tiny**
- **No dynamic memory**
- **Tiny Runtime (20kbs!)**

{Summary}

- Working Web Solution == Client + Server
- Model Shrinking
 - Vision == 😎
 - Language == 🤕
- TF Lite Micro vs. Barracuda?

Web-based Machine Learning for Gaming

Josh Meyer
&
Lindy Rauchenstein
{ Lead Scientists }

Artie[™]