ONNX and IBM Z

Elpida Tzortzatos, IBM Fellow, IBM Z AI Strategy and Architecture, CTO z/OS
elpida@us.ibm.com

Andrew M. Sica , STSM IBM Z AI and Analytics
andrewsi@us.ibm.com
You probably have used IBM Z today!

- 67 of the Fortune 100
- 45 of the world’s top 50 banks
- 8 of the top 10 insurers
- 4 of the top 5 airlines
- 7 of the top 10 global retailers
- 8 out of the top 10 telcos
Deploy on IBM Z
Deploy on IBM Z and seamlessly exploit innovations across the stack to infuse AI in every single transaction.

Organize Data
Import data from different applications and sources

Train anywhere
Public clouds, private clouds, on-premises, and hyperconverged systems.

Import Data → Model & Data Prep → Model Training → Deploy → Predict

Business Applications
IBM Z: Fully enabled platform for Business Intelligence

Business Applications
Build and Train Anywhere

IBM Z AI & Analytics / © 2021 IBM Corporation
Seamlessly embed AI into existing z/OS applications

IBM Watson® Machine Learning for z/OS
Helps data scientists and developers accelerate the use of AI and ML

- Simplify and deploy AI at scale
- Enable retraining of models
- Generate APIs for AI-powered applications
- Streamline model management

Delivers simplified AI model deployment and lifecycle management

Simplifies and streamlines calls to model scoring services directly from transactional workloads

Invoke through standard APIs

*http://spark.apache.org/
Value of ONNX to an IBM Z client

• Ensemble of technologies that enable tight integration of AI into business applications.

• Adaptability to different architectures and operating environments.

• Optimize for deployment; reuse of assets for low latency environments.

• Simplify and standardize deployment assets.
Thank you!

Any questions?

• What is a mainframe?

• IBM Telum announcement blog

• IBM Z AI Strategy