

ML/AI in Web apps- examples, sample challenges and network collaboration

Göran AP Eriksson

Ericsson Research

Web5G workshop, London May 2018

Goals



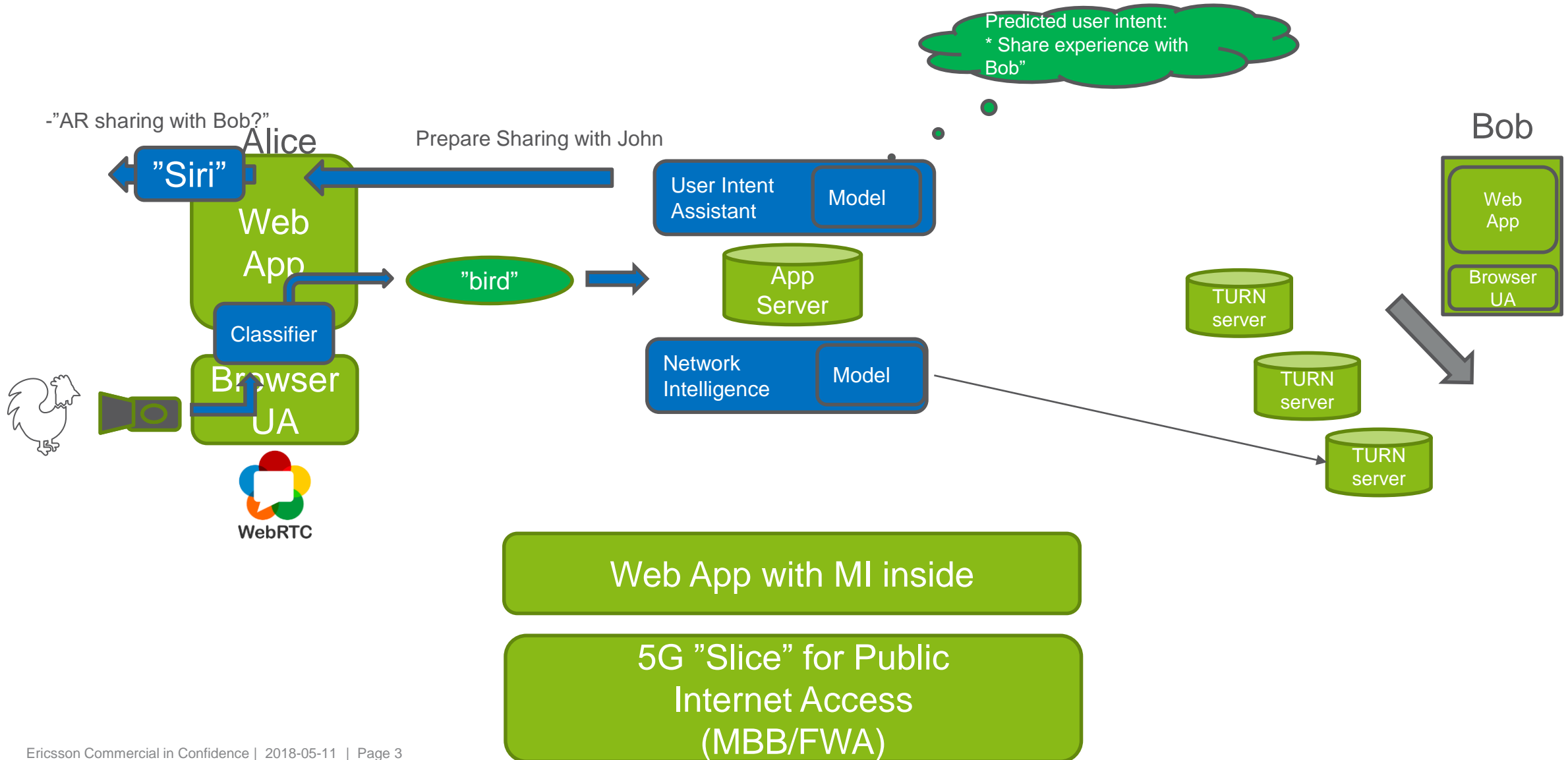
2022'ish



- A picture how Web Applications tomorrow could use ML/AI ("MI")
- Ideas about technologies we can expect from "Web Cloud Native"
- Thoughts about "Collaborations" between Web Applications and the "5G"



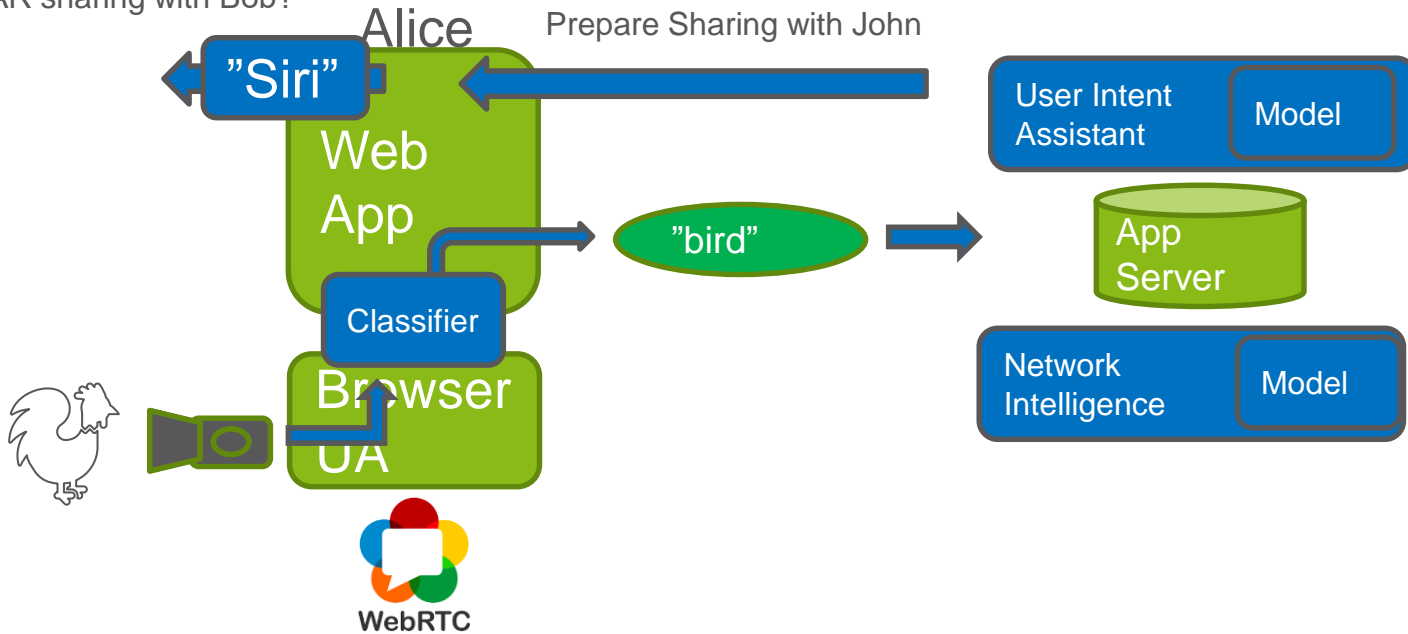
An example web app with MI inside...



An example web app with MI inside...



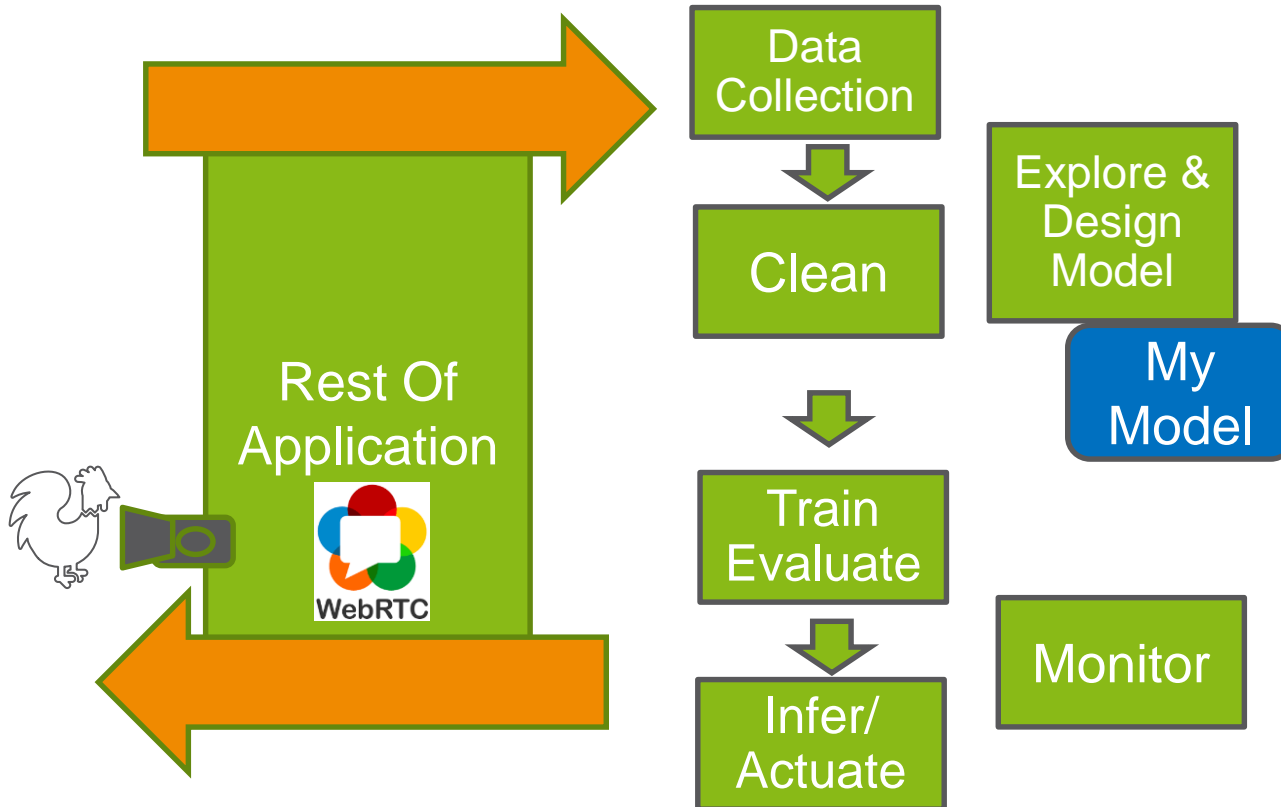
- "AR sharing with Bob?"



- 4 models in operation
- A lot of devices and users
- Google search, Assistant, YouTube, GCP, ...
- Many models in operation...
- What does it take to have an MI model "in operation"?

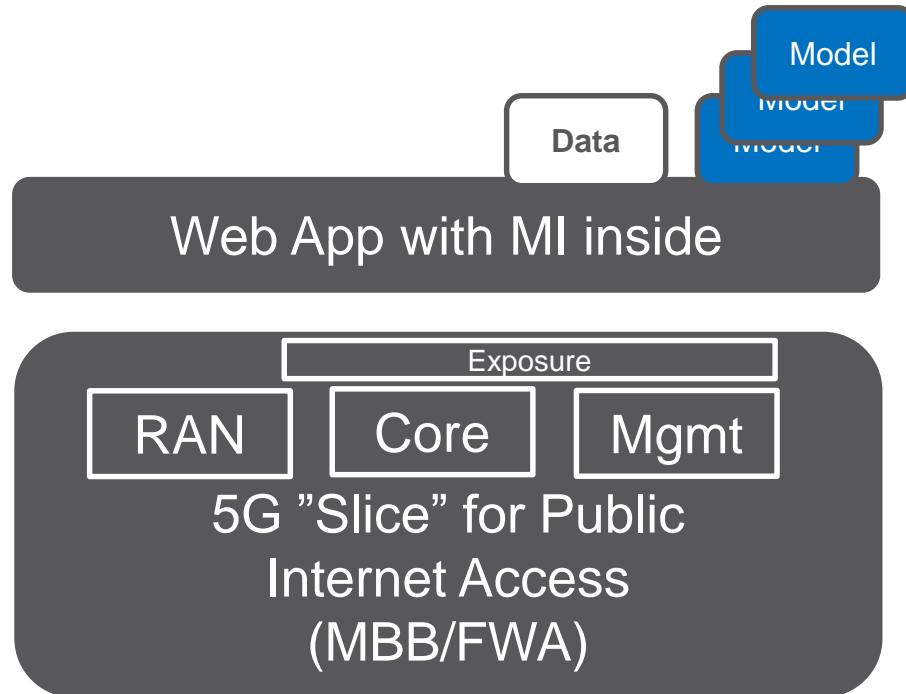


Simple ML application anatomy



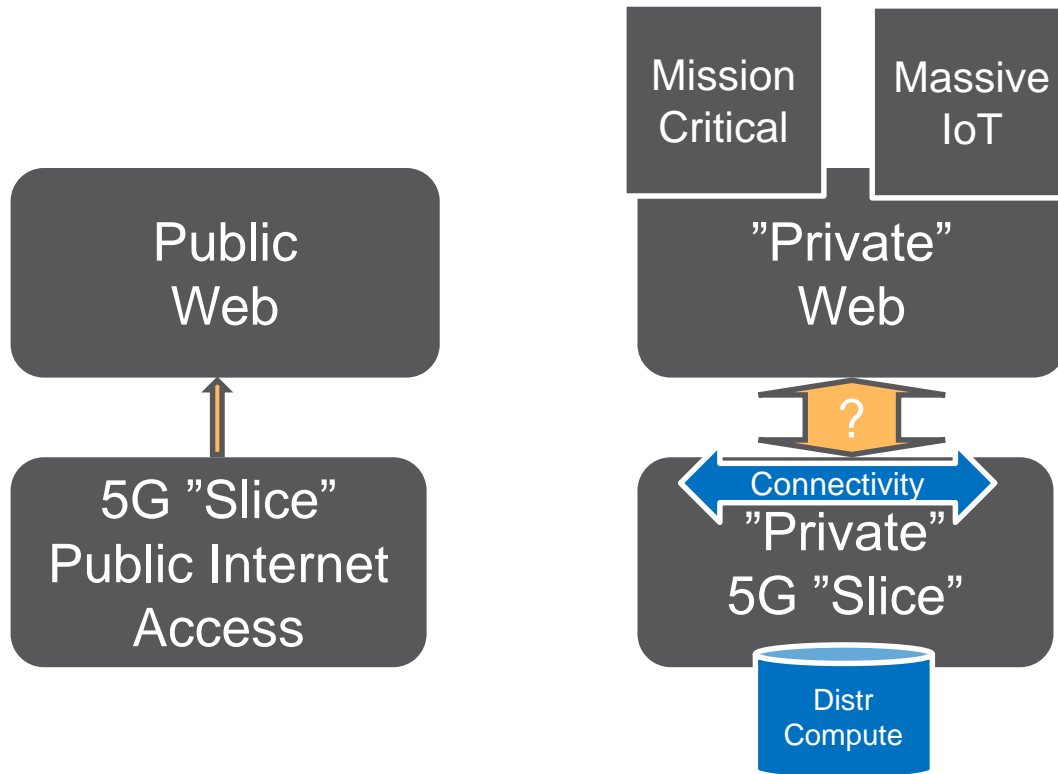
- › Two processes
 - Service delivery
 - Model devOps
- › Model provides predictions or classifications
- › Model devOps specifics:
 - Deep Learning training compute intensive
 - Prediction Accuracy vary
 - Production cost/prediction?
 - Robustness, Security and privacy

Challenges and some "co-optimizing" questions



- > Model devops
 - Training
 - KPI's
 - Performance: speed, accuracy, robustness, security
 - Production cost
- > 4 models, B's of devices, Model devOps- a large scale software system.
- > Food for thought:
 - Web Cloud native MI technologies impact on 5G systems?
 - 5G enabling Media intensive and "MI native" Apps?
 - Cross layer collaboration and exposure?

Co-optimization? App-network collaboration?



- › 1st- Connectivity and distributed computing
- › Cross Layer is Not a new question
- › Technically, there are possibilities and barriers
- › But barriers from business perspective in "MBB" area/Internet/Public Web
- › But in Private networks, "verticals", such as manufacturing, Smart City, etc, cross-layer may stand a better chance to succeed
- › And in MI, "DATA is King"- A driver for Shared data lake/pipeline and learnings cross layer?



End