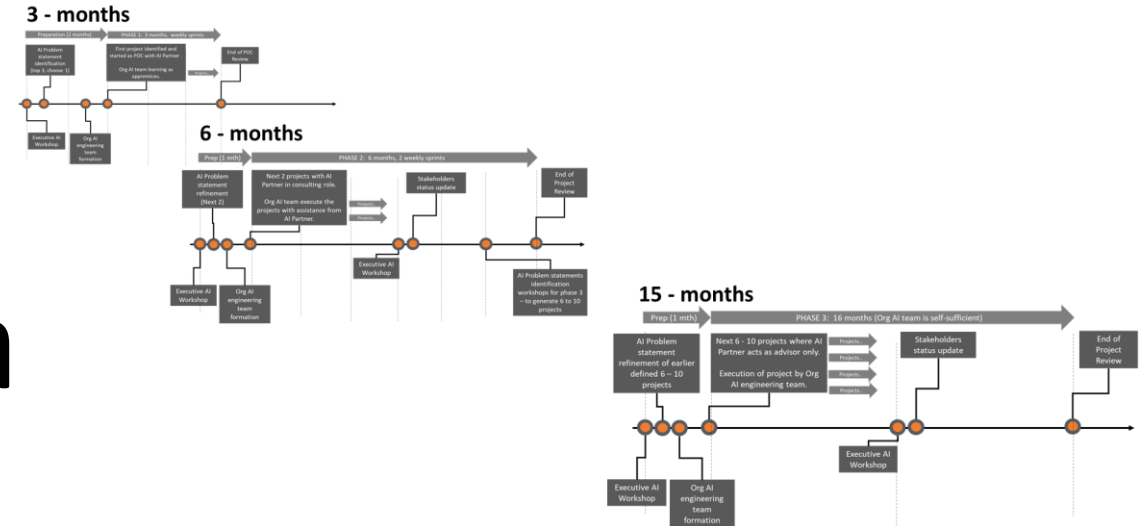


Centre of Excellence for Data Science and AI

Build or outsource? How do you go about building your team?

Laurence LIEW
Director, AI Industry Innovation
AI Singapore



Beyond the data science and AI team

- **Structure**

- Data science and AI teams needs to work closely with business, learn their domain, speak their language
 - Else disjointed, far away, not trusted by the business users
- However, if too deeply involved, biases sets in
 - Predictions may “make the boss look good”

- **Data-driven culture**

- Too much of a good thing
 - Trust the data and model too much and upset the sales (“gut feel”) folks
- Too much gut-instinct
 - When data shows otherwise – ignore or push back

- **Avoid complexity (KISS)**

- Too complicated (long development time, expensive, difficult to change) models not necessary better
- Models must be deployable! Eg Netflix \$1M prize 1st runner-up was deployed and not the 1st prize winner
- Too simple also an issue if the model cannot capture the nuances of the underlying data.

Beyond the data science and AI team

- Core Data Science and AI Team
- Have a Chief Data / Analytics / AI Officer at the C-Suite
- Start small
 - Short 3- 6 months MVP
 - Use this as the champion to gain mindshare and support from stakeholders
- Data, Data and Data
 - Cleaned?
 - Unbiased?
 - Labeled?
 - Enough for Deep Learning? Else do standard machine learning

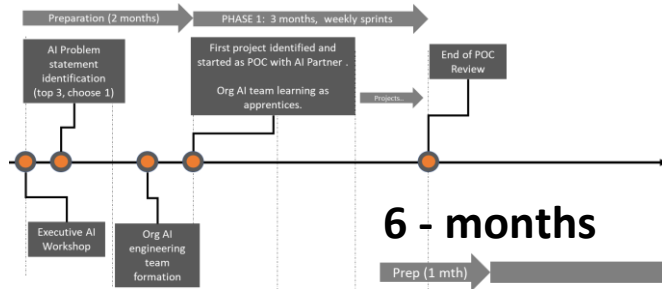
COE for Data Science and AI

- 3 Phase (24-months) ramp-up
 - Phase 1: 3 months, small initial POC
 - Phase 2: 6 months, 2 bigger projects, setup of data science/analytics/AI Stack
 - Phase 3: 15 months, 6 projects, full AI/HPC Stack up and running
- Ideas and best practices from experience in building Data Science and DevOps team
 - Revolution APAC office
 - Customers

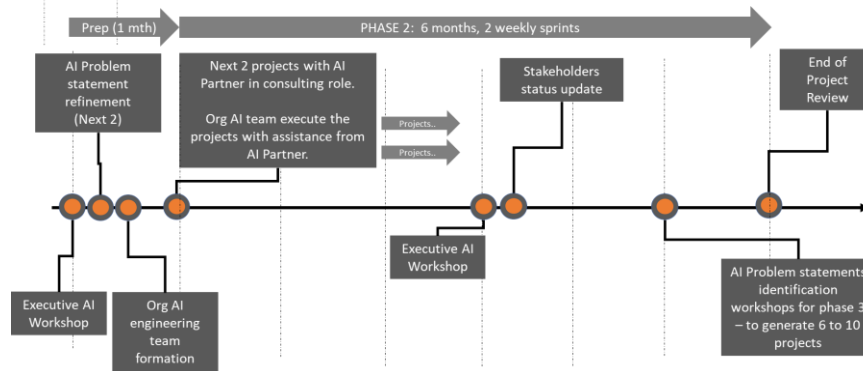
Revolution Analytics Singapore: 2012-2016 (Revolution Analytics was acquired by Microsoft in 2015)

24-months Coe Plan

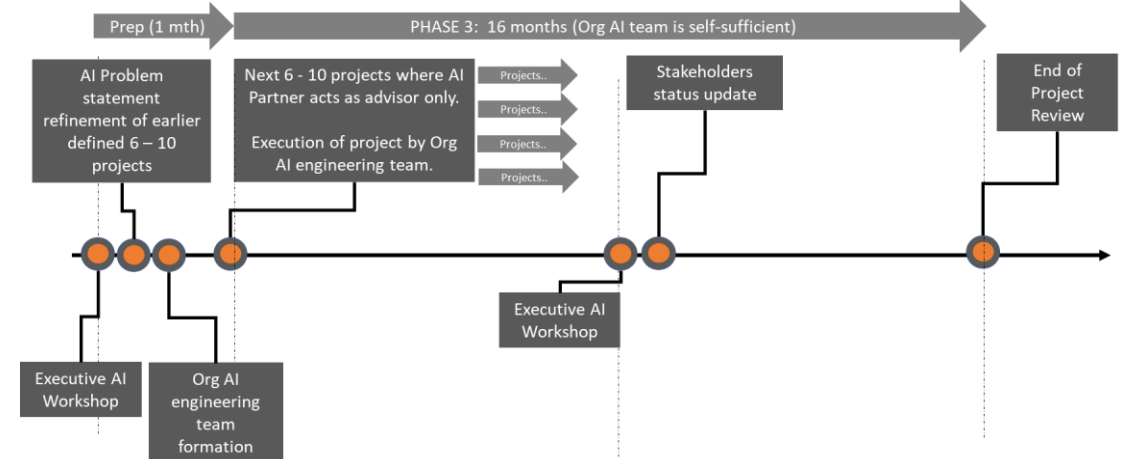
3 - months



6 - months



15 - months



Phase 1

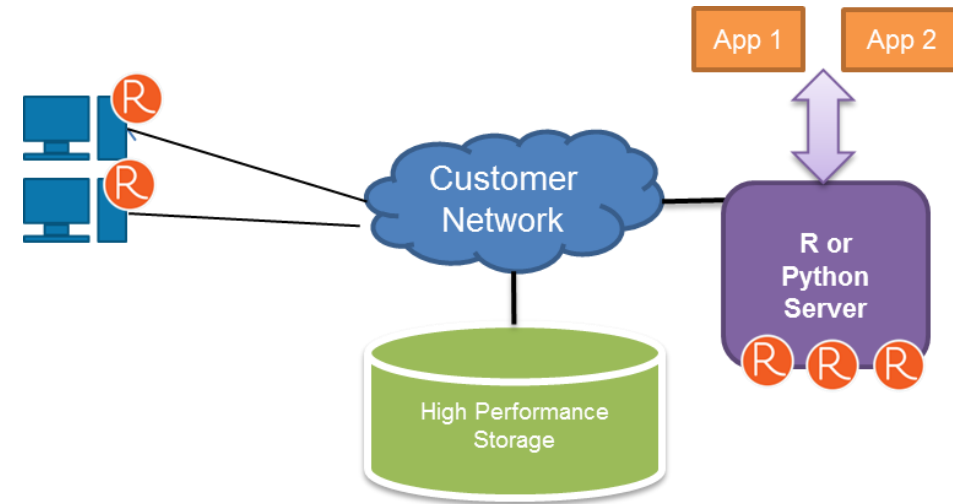
People and Setup

- **AI Partner**

- 1 x AI Scientist (25%)
- 1 x AI Engineer (100%)
- 1 x AI DevOps Engineer (100%)

- **Customer**

- 1 x AI project
- 2 software engineer to be trained in AI/ML with R/Python
- 1 x IT engineer to be trained on AI/HPC Stack



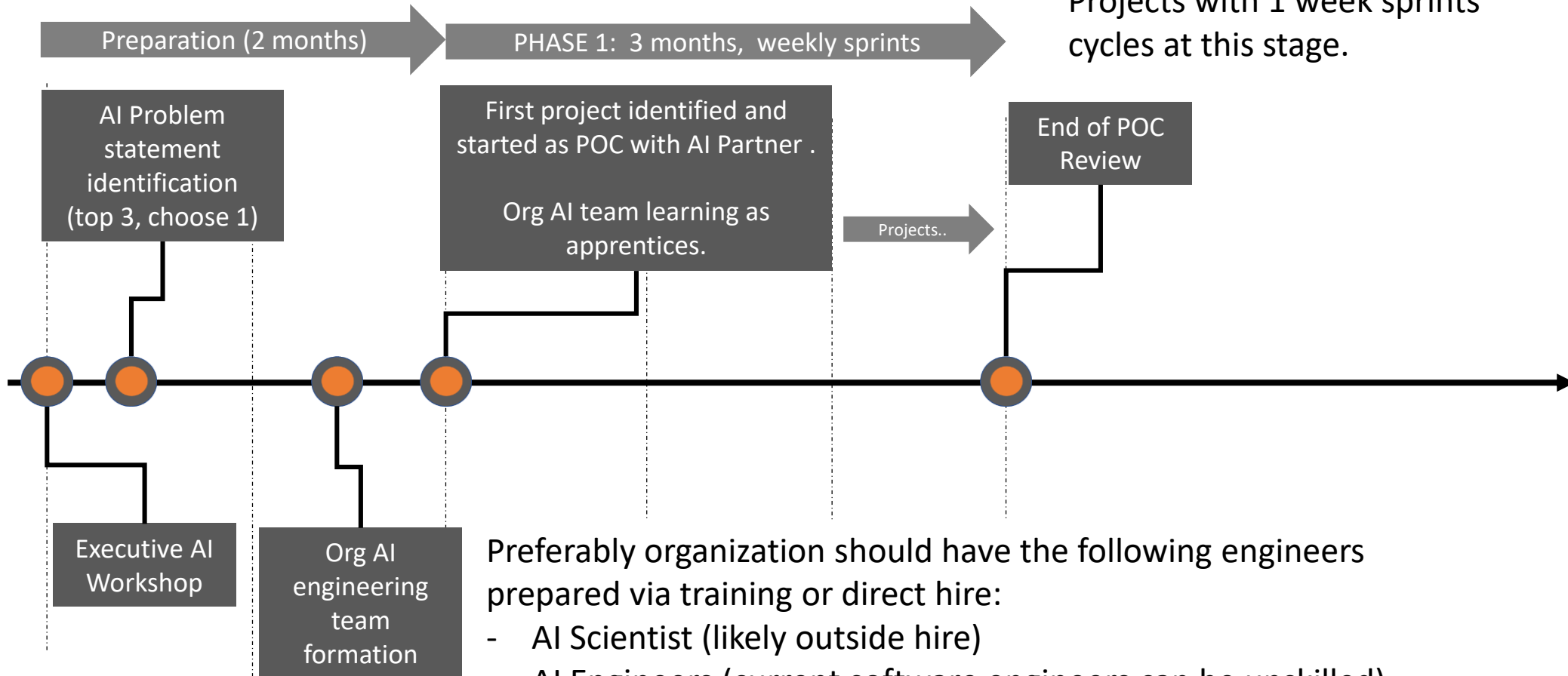
Usage model:

1. Build initial models on workstations with GPUs
2. If Deep Learning, explore use of transfer learning to accelerate model building
3. Run bigger models on GPU servers/cluster
 - Each user to limit himself/herself to 4-cores max per run so as not to starve others of CPU processing
 - So max 8-concurrent jobs of “4-cores 64GB ram”
4. IT will build end-user facing applications using the models built
5. WEB or MOBILE applications as frontend, executing R/Python models or making use of output from models built.

Phase 1

3 months – Keep it simple!

We recommend an agile methodology for the AI Projects with 1 week sprints cycles at this stage.



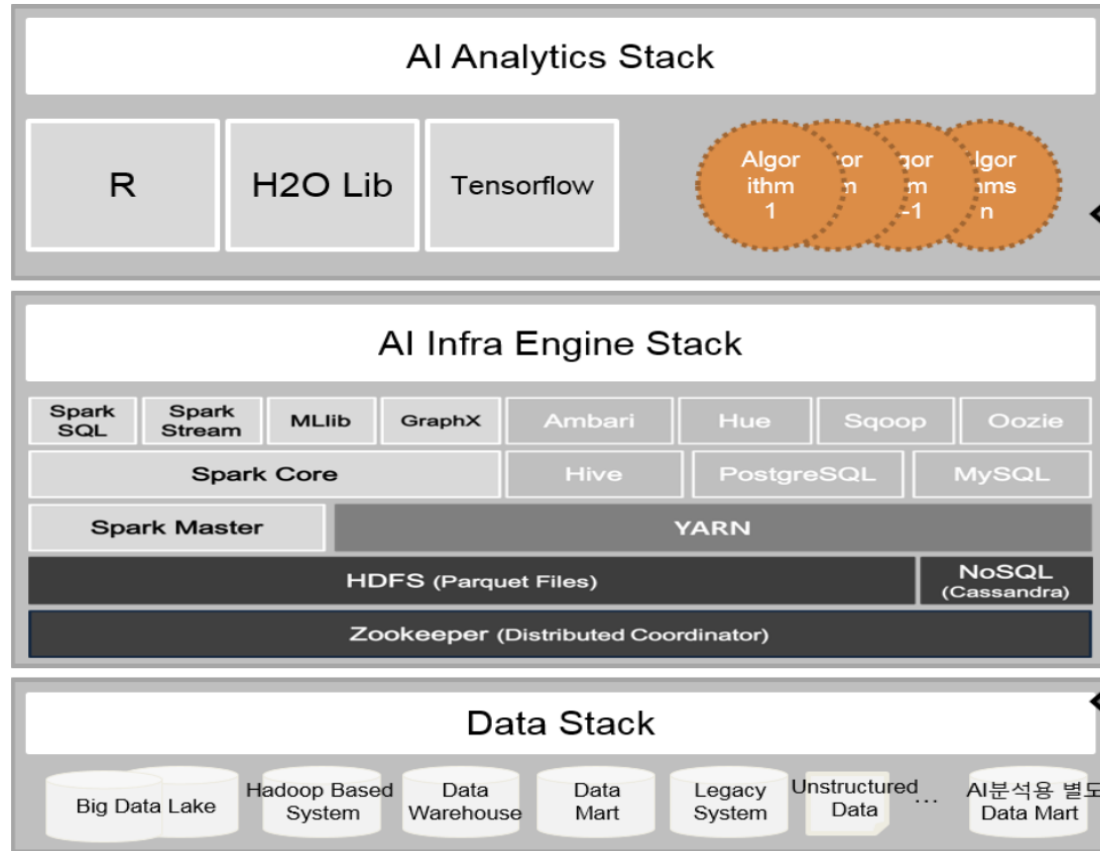
Preferably organization should have the following engineers prepared via training or direct hire:

- AI Scientist (likely outside hire)
- AI Engineers (current software engineers can be upskilled)
- AI Devops engineers (current IT engineers can be upskilled)

Phase 2

Phase 2: 5 months CoE Ramp-up

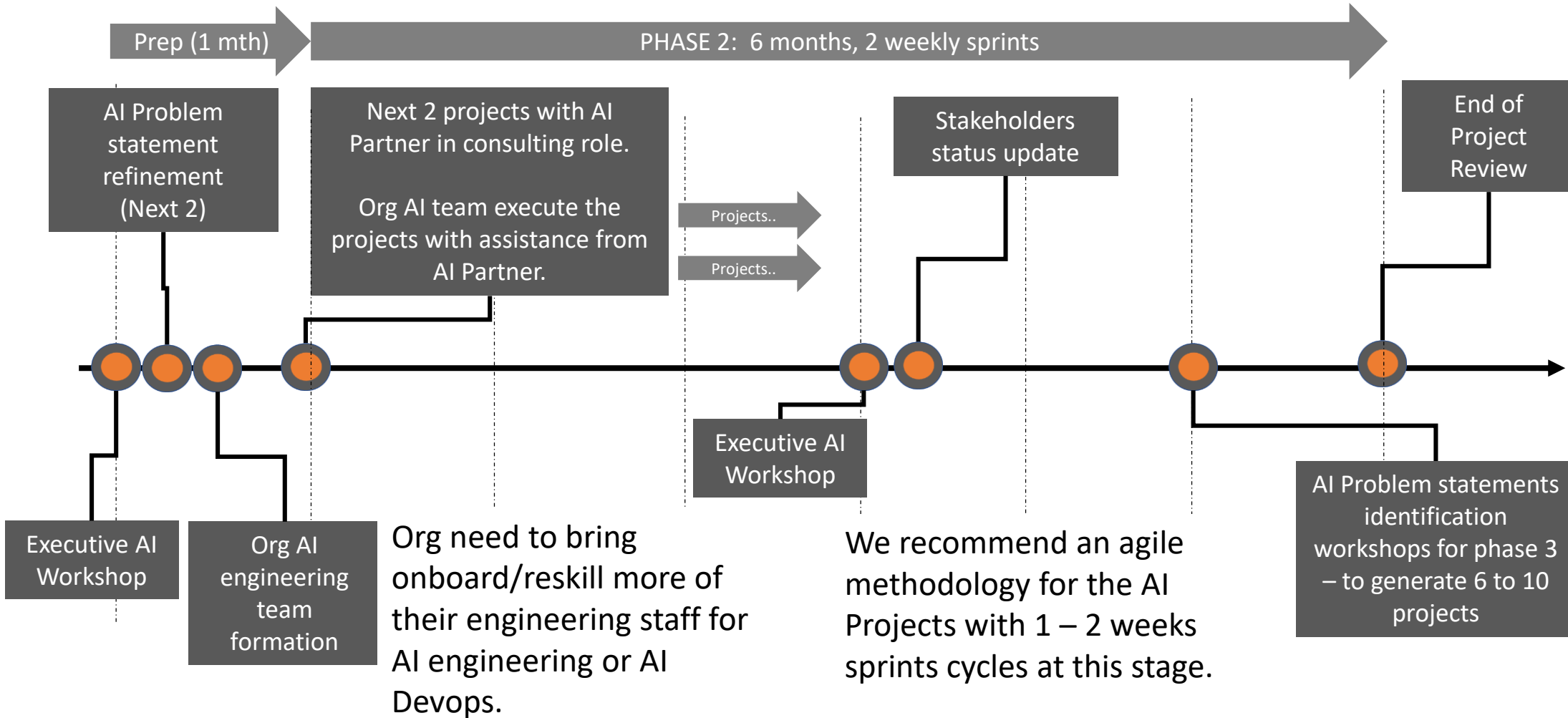
- **AI Partner**
 - 1 x AI Scientist (25%)
 - 1 x AI DevOps Architect (25%)
 - 1 x AI Engineer (100%)
 - 1 x AI DevOps Engineer (100%)
- **Customer**
 - 2 x AI project
 - Additional 2-4 software engineer to be trained in AI/ML with R/Python
 - Additional 1-2 x IT engineer to be trained AI/HPC Stack



- **Build bigger/more complex models**
- **Build up your Analytics Stack**

Phase 2

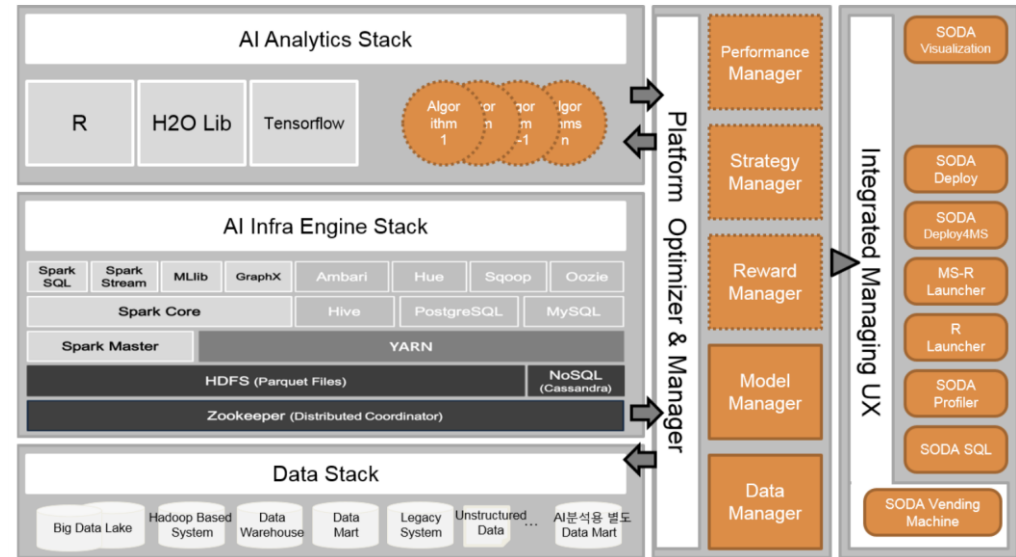
Bigger projects + AI Stack



Phase 3:

16 months CoE Capability Development

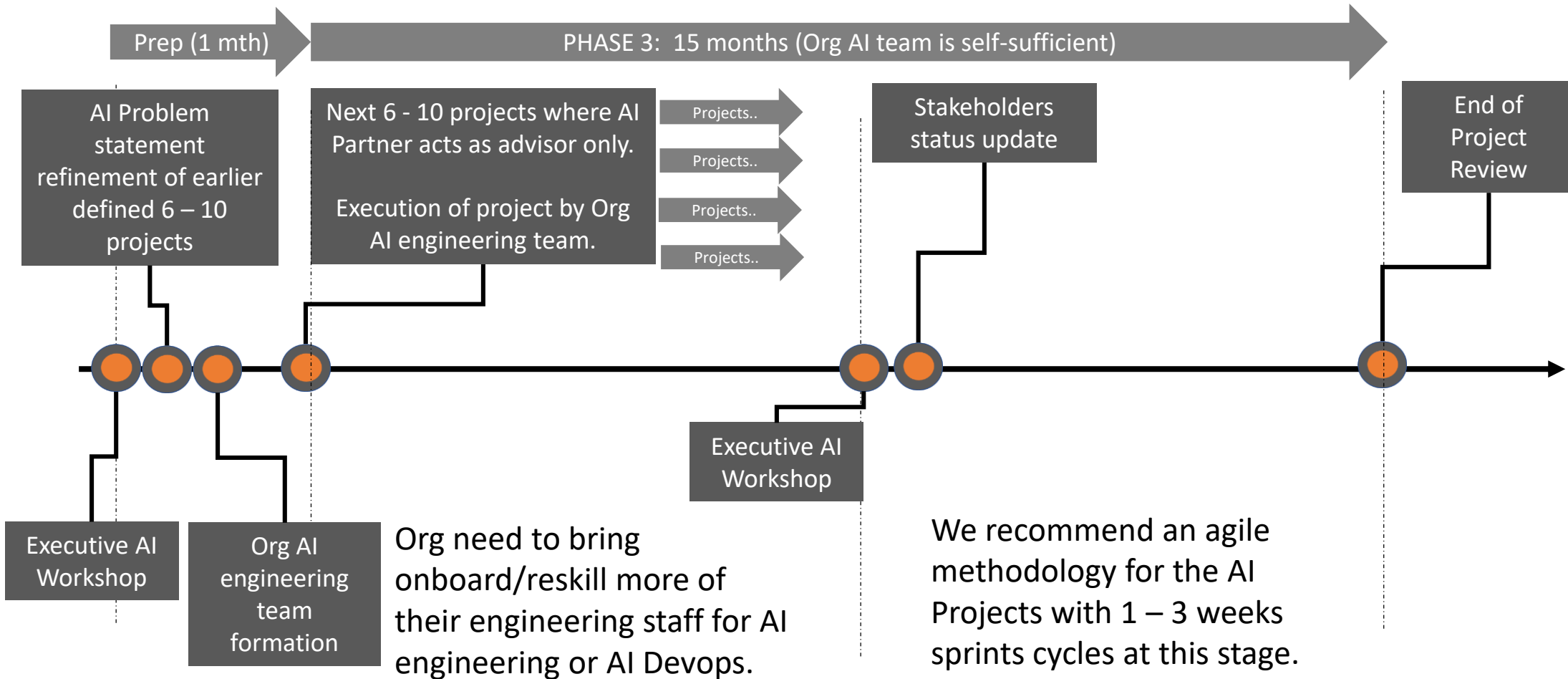
- **AI Partner (start to scale down – more as advisory role)**
 - 1 x AI Scientist (25%)
 - 1 x AI DevOps Architect (25%)
 - 1 x AI Engineer (25%)
 - 1 x AI DevOps Engineer (25%)
- **Customer (should be operational)**
 - 8 – 10 AI projects pipeline
 - Additional software engineer to be trained in AI/ML with R/Python
 - Additional IT engineer to be trained AI/HPC stack



- **Build FASTER, bigger/more complex models**
- **Full Analytics Stack up and running**

Phase 3

Full Analytics Stack Achieved and Team in Place



Phase 3

organization-wide innovation with analytics

- **Customer Organization**

- Core Data Science team formed
- Core Devops team formed
- A scalable analytics platform built to undertake the various analytics workload
- Target a sustainable 5 - 6 projects per year on average (depends on team size of course)
- Agile/Scrum methodology recommended to ensure success

- **AI Partner**

- Step down and provide on-going advisory if required

- **End of 24 months**

- Review next steps