

# AI Platform and Challenges

## Perspectives from AI Research Center (AIRC)

Junichi Tsujii  
Director  
AI Research Centre(AIRC)  
AIST, Japan

# AI Initiatives in Japan

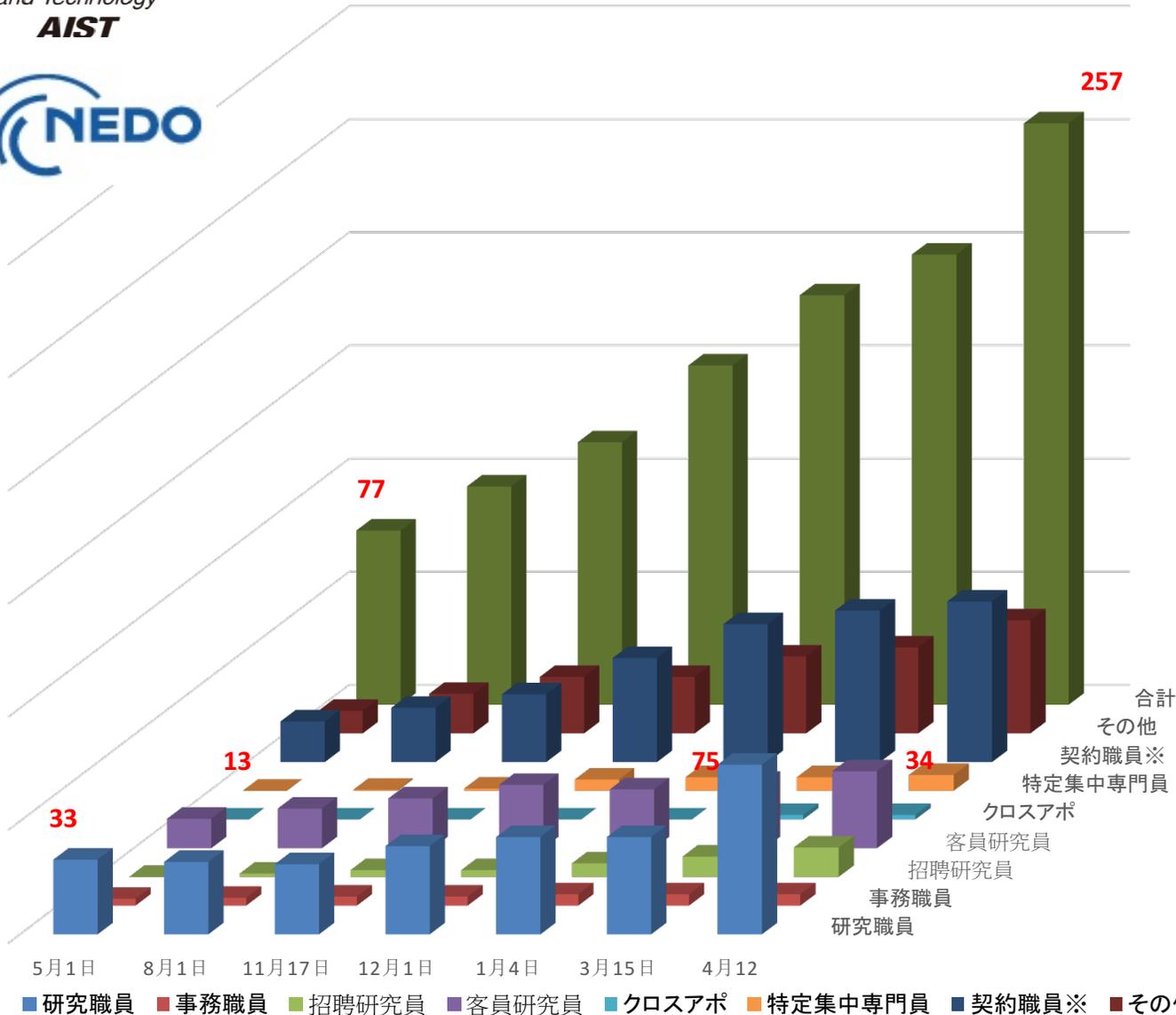
- Strategic Planning Committee of AI under Cabinet
  - Committee of Research Coordination
    - AIP Centre under MEXT --- Basic Research at Riken
    - AIRC under METI --- Translational Research at AIST
    - CiNet, NICT under MIA --- Communication/Brain Science
  - Committee of Industrial and Social Deployment of AI

# AI Embedded in the Real World

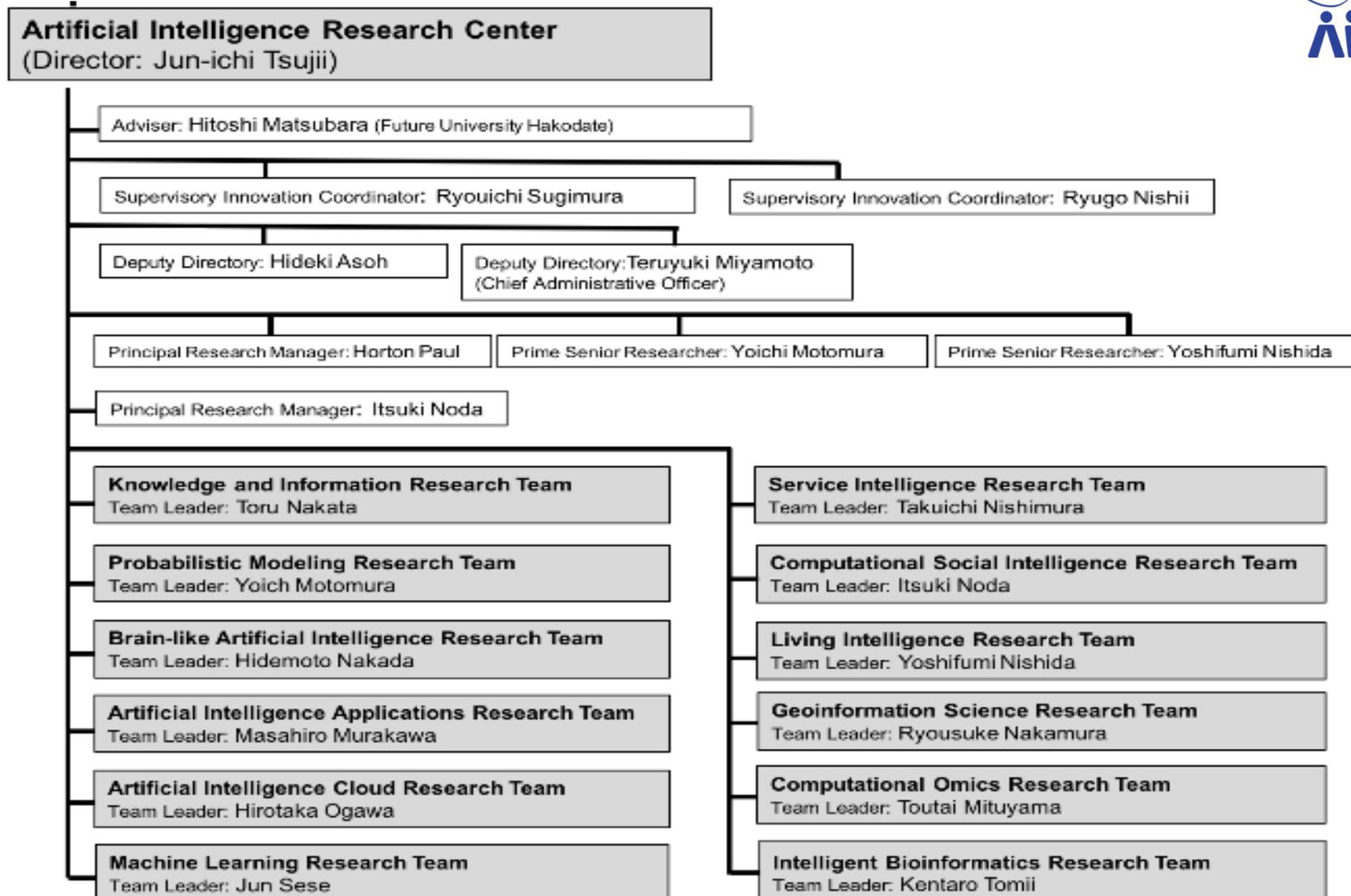


AI which cooperates with Human

# AI Research Center (AIRC) at AITST



	1/May 2015	12/April 2016
FT Researcher	33	75
Support Staff	3	5
PT Res. (Univ.)	0	13
Invited Res.	13	34
Cross Appoint.	0	2
PT Res. (Private Sector)	0	7
Support Staff (Eng., RA, PD)	18	71
Secretarial Staff	10	50
<b>合計</b>	<b>77</b>	<b>257</b>



# Outline

- AI Platforms
  - Background and the Rationale
- Examples of Research at AIRC
  - Self-navigation and Self-driving
  - AI for Big Sciences (Robot Scientist)
  - Processing of a large collection of Images
  - AI for Human Life (Elderly care, Medical care, Retailing, ..)
- Computational Infrastructures
- Conclusions and Comments

# Outline

- AI Platforms
  - Background and the Rationale
- Examples of Research at AIRC
  - Self-navigation and Self-driving
  - AI for Big Sciences (Robot Scientist)
  - Processing of a large collection of Images
  - AI for Human Life (Elderly care, Medical care, Retailing, ..)
- Computational Infrastructures
- Conclusions and Comments

# AI Platforms

- **Close interaction between research and deployment**
  - Speed of Deployment: If we have an idea that actually works, within a month it can be in front of 1.5 billion people (by LeCan, Facebook)
- **AI for AI → AI for Smart Society, Sciences, and Manufacturing**
  - Partnerships between research institutes and institutes using the technology
- **Technology and Data as Commodities**
  - AI Platforms
  - Tension between Commons and their Monetization

# AI Platforms

- **Close interaction between research and deployment**
  - Speed of Deployment: If we have an idea that actually works, within a month it can be in front of 1.5 billion people (by LeCan, Facebook)
- **AI for AI → AI for Smart Society, Sciences, and Manufacturing**
  - Partnerships between research institutes and institutes using the technology
- **Technology and Data as Commodities**
  - AI Platforms
  - Tension between Commons and their Monetization

# State of Affairs

## AI Research/Deployment

- **Giant IT Companies in the US**

- **Data, Seeds and Needs** in Single Organizations

- **Data** -- Data gathered by them
- **Seeds** -- Researchers/Engineers and Computational Resources
- **Needs** -- Clear Business Models

- **Closed eco-system**

- Transition from internal big data to scattered and/or more specialized data

- M&A of Start-ups

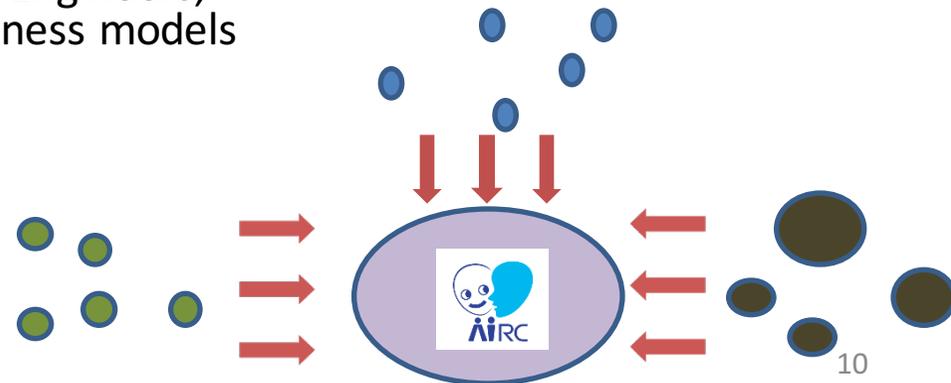
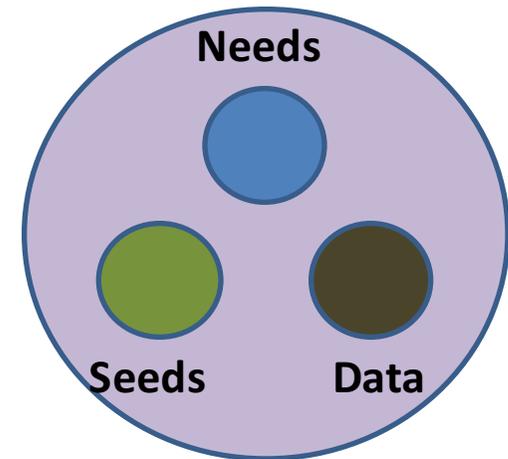
- **Japan, Europe and Academia**

- **Fragmentation** of Data owners, Researchers/Engineers, Smaller computational resources, lack of business models

- Open Alliance of stakeholders

- Cooperation with Start-ups and industries

Giant IT Companies (G,M,F,A)



# AI Platforms

- **Close interaction between research and deployment**
  - Speed of Deployment: If we have an idea that actually works, within a month it can be in front of 1.5 billion people (by LeCan, Facebook)
- **AI for AI → AI for Smart Society, Sciences, and Manufacturing**
  - Partnerships between research institutes and institutes using the technology
- **Technology and Data as Commodities**
  - AI Platforms
  - Tension between Commons and their Monetization

# State of Affairs

## AI Research/Deployment

### Giant IT Companies (G,M,F,A)

- **Giant IT Companies in the US**

- **Data, Seeds and Needs** in Single Organizations

- **Data** -- Data gathered by them
- **Seeds** -- Researchers/Engineers and Computational Resources
- **Needs** -- Clear Business Models

- **Closed eco-system**

- Transition from internal big data to scattered and/or more specialized data

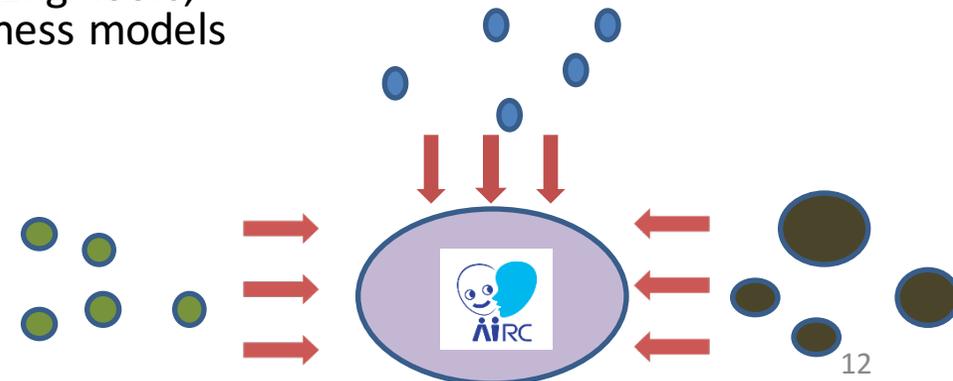
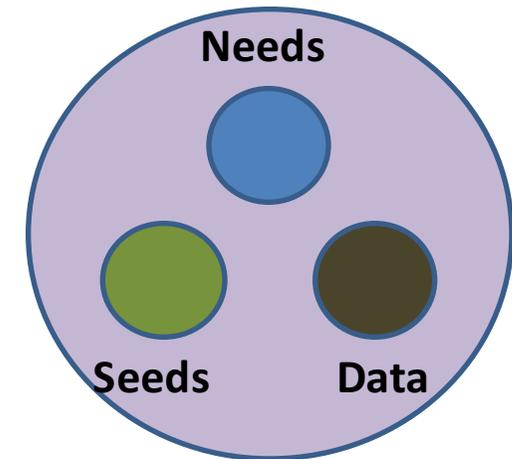
- **M&A** of Start-ups, **Lock-in** of users in their eco-systems

- **Japan, Europe and Academia**

- **Fragmentation** of Data owners, Researchers/Engineers, Smaller computational resources, lack of business models

- **Open Alliance of stakeholders**

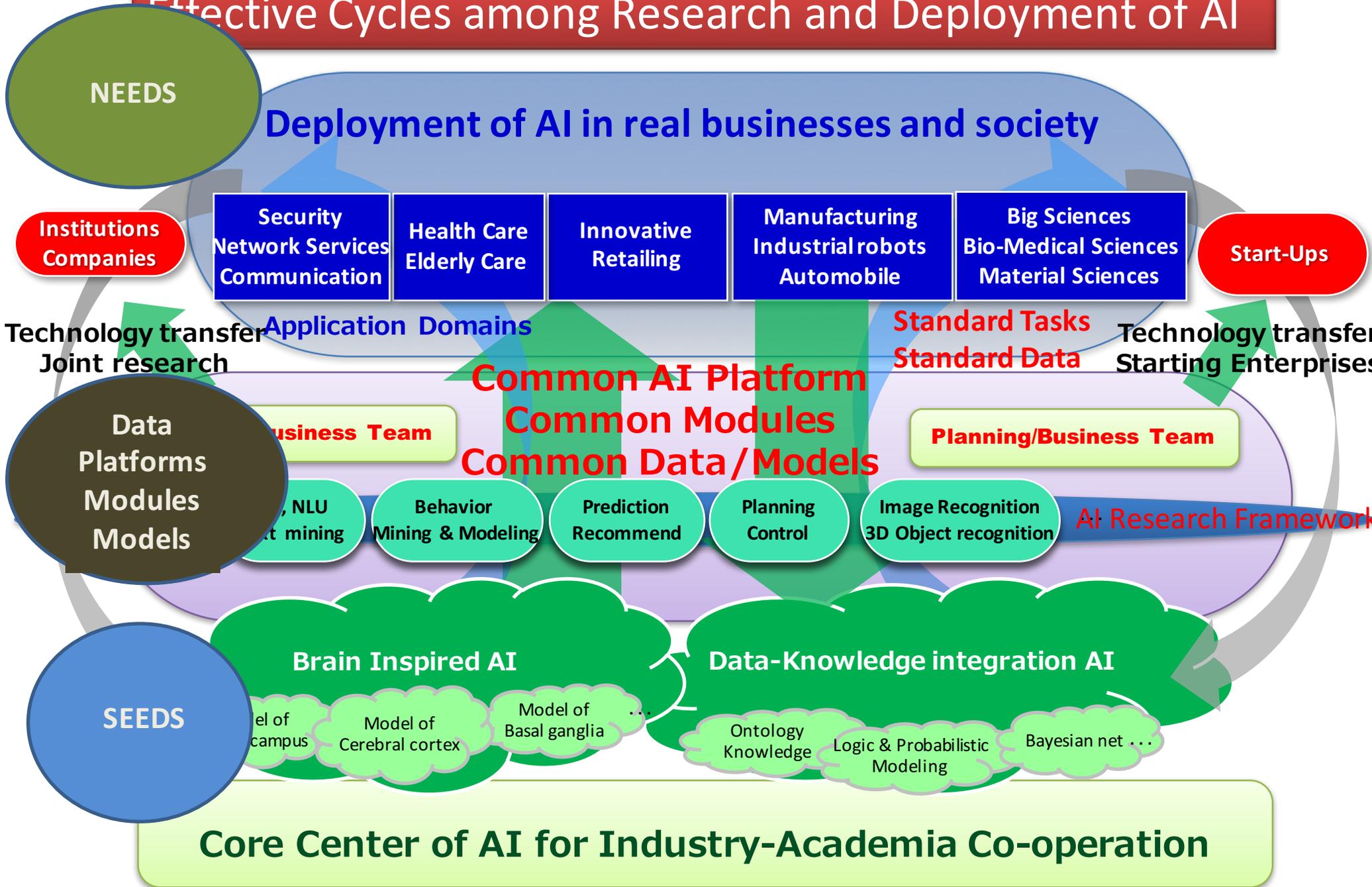
- **Cooperation** with Start-ups and industries



# AI Platforms

- **Close interaction between research and deployment**
  - Speed of Deployment: If we have an idea that actually works, within a month it can be in front of 1.5 billion people (by LeCan, Facebook)
- **AI for AI → AI for Smart Society, Sciences, and Manufacturing**
  - Partnerships between research institutes and institutes using the technology
- **Technology and Data as Commodities**
  - **AI Platforms**
  - Tension between Commons and their Monetization

# Effective Cycles among Research and Deployment of AI

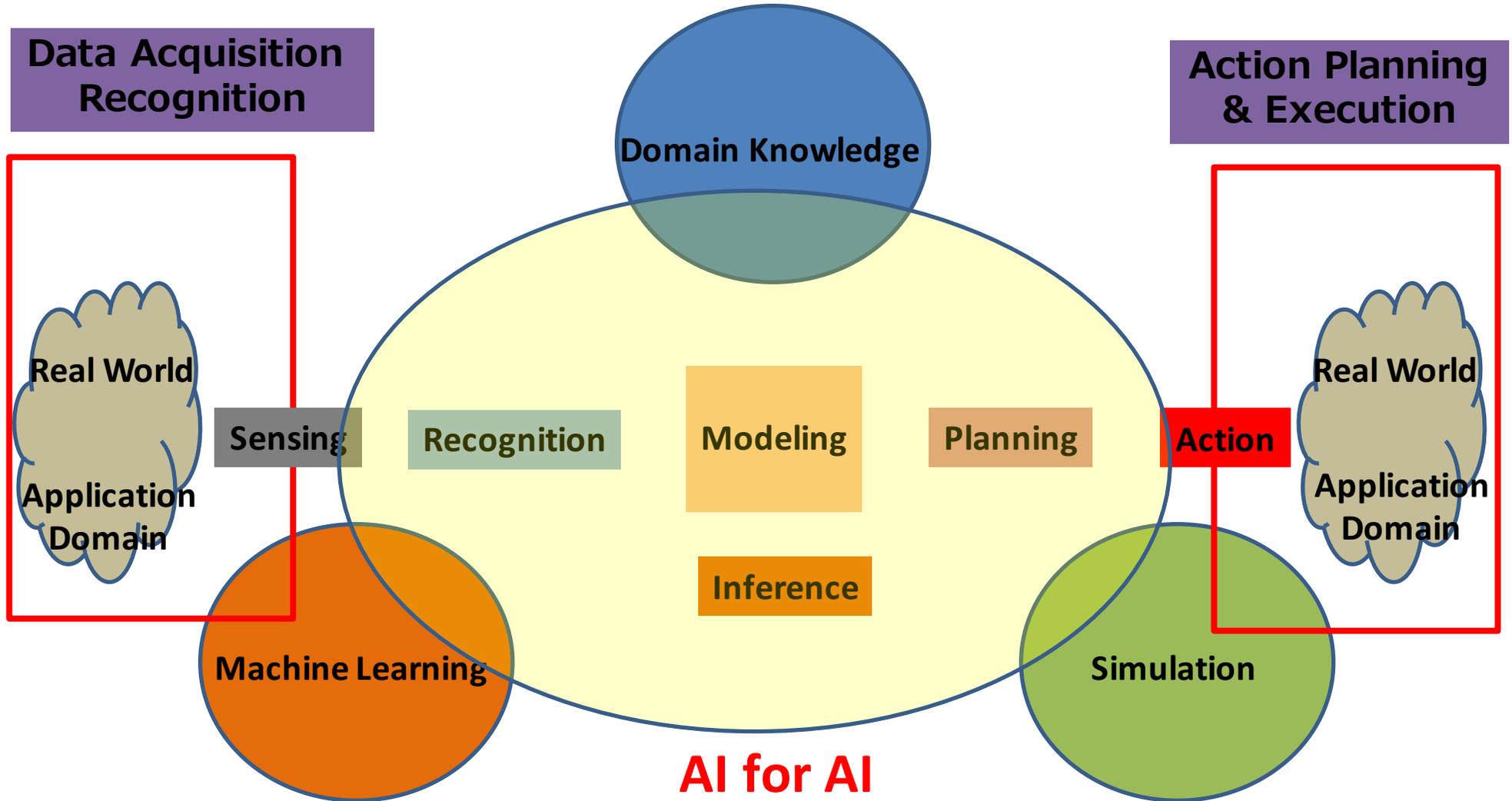


# AI Embedded in the Real World



AI which cooperates with Human

# Basic Components and Technologies in AI



**AI embedded in the real world—Partnerships**

# AlphaGo (2016)

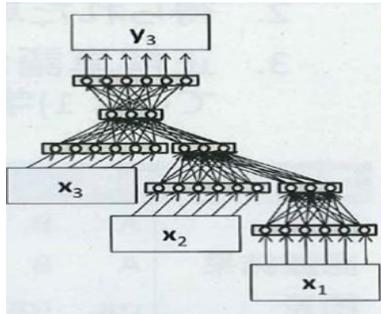
## Machine Learning and Simulation



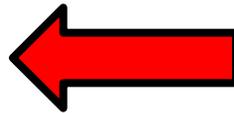
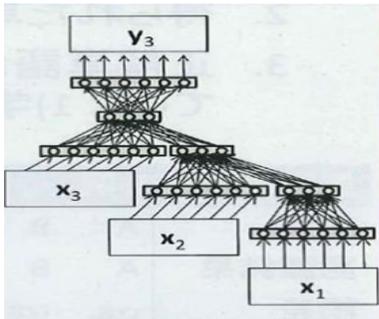
A game of perfect information

DNN

$v(s)$



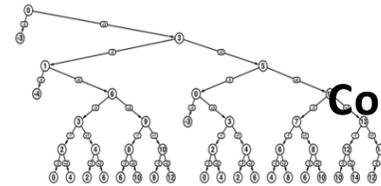
$p(a/s)$



Training Data



Database of  
Games in the past



Complete Simulation

# AI embedded in the real world

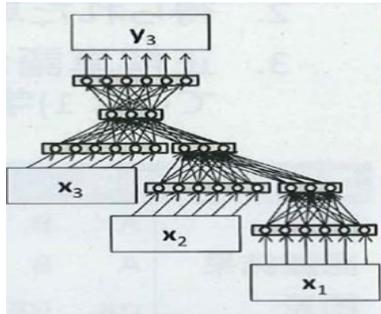
## Simulation and Machine Learning



Partially Known Information

DNN

$v(s)$

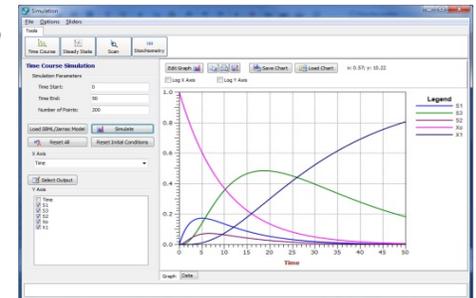
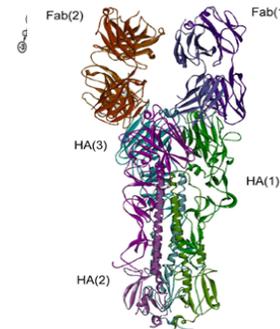
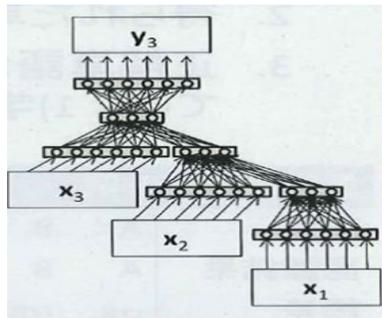


Train



Diverse databases  
Database of Games in the past  
Omits

$p(a/s)$



Incomplete and Partial Simulation

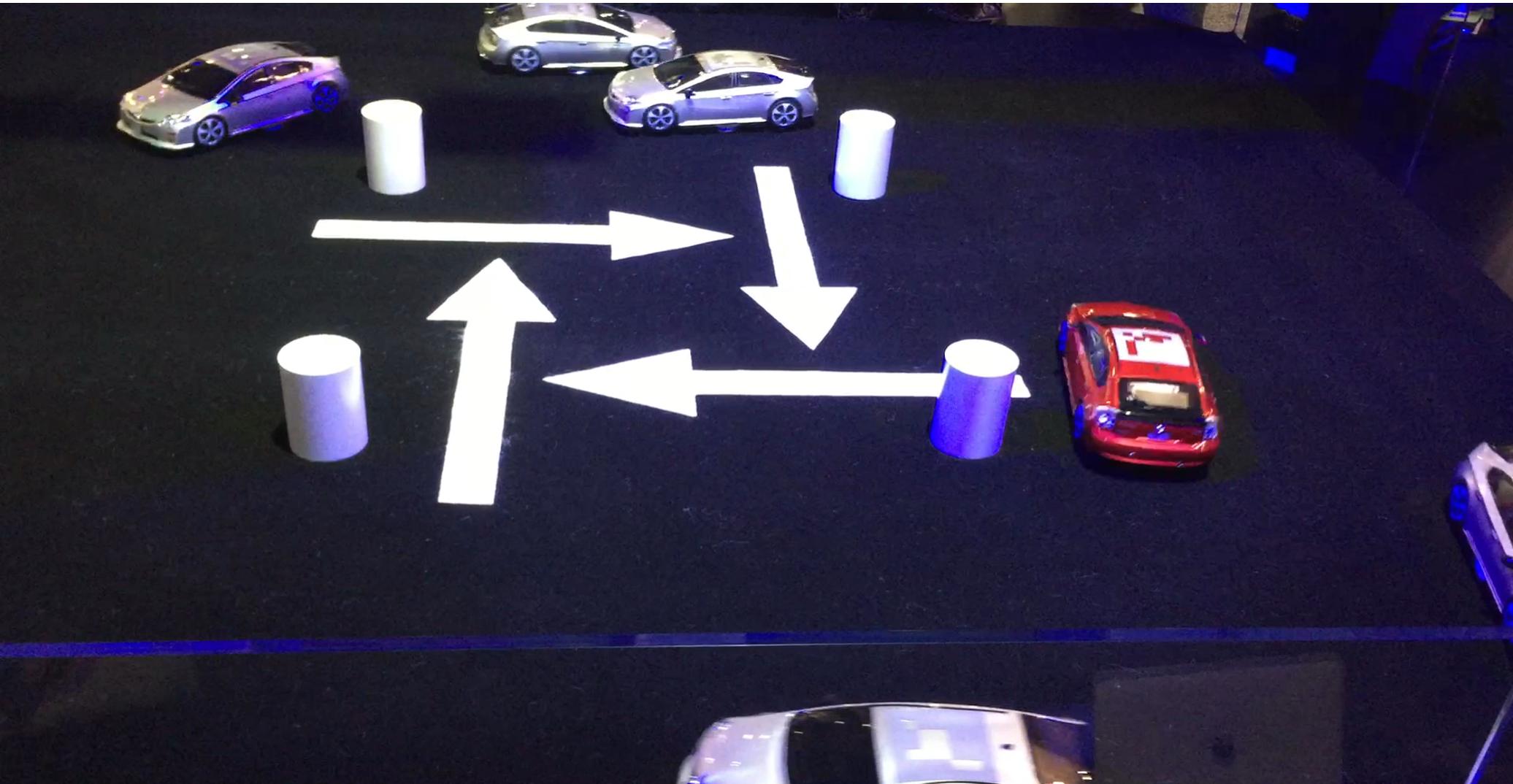
# Outline

- AI Platforms
  - Background and the Rationale
- Examples of Research at AIRC
  - Self-navigation and Self-driving
  - AI for Big Sciences (Robot Scientist)
  - Processing of a large collection of Images
  - AI for Human Life (Elderly care, Medical care, Retailing, ..)
- Computational Infrastructures
- Conclusions and Comments

# Recognition, Model of the World, and Action

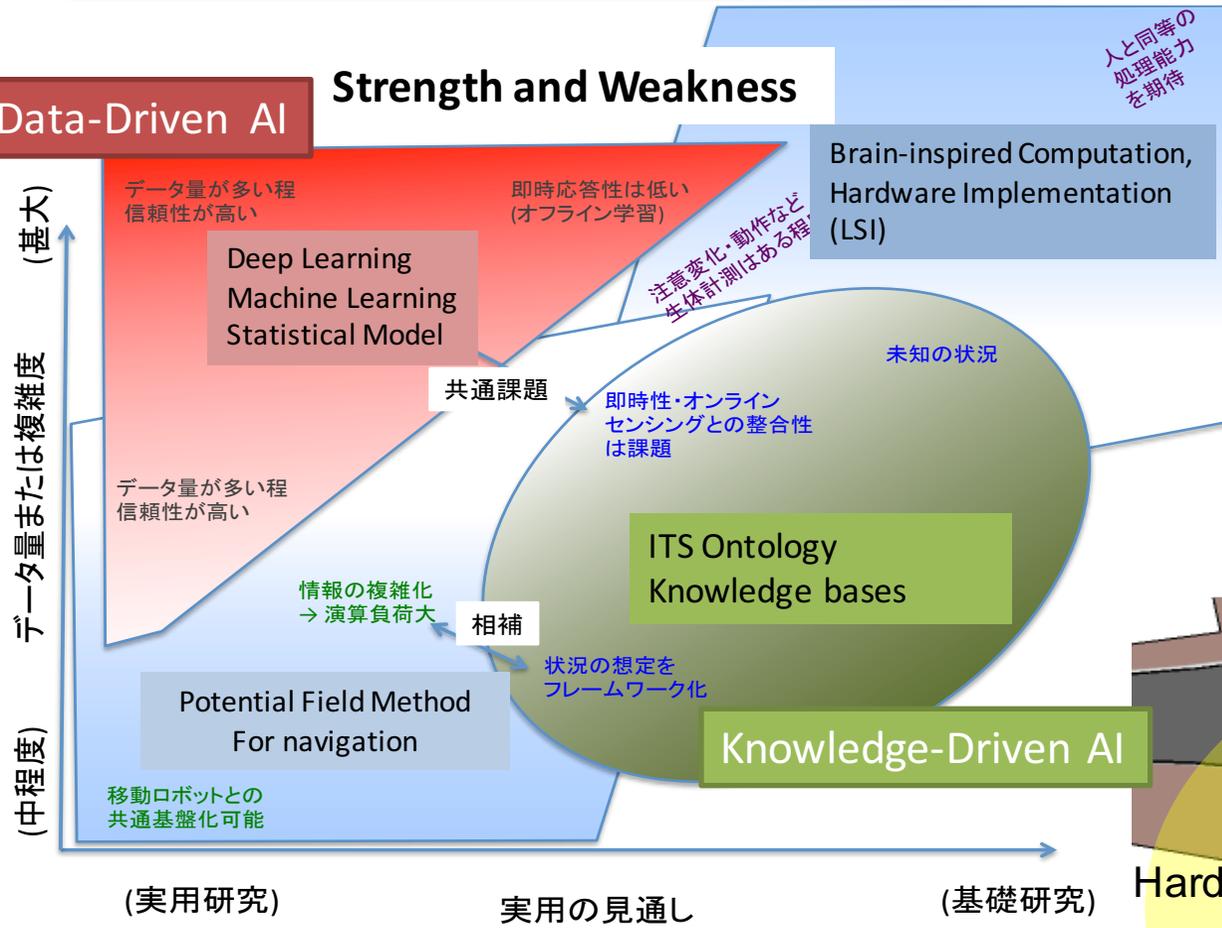
## Self-Driving in the Real World

# Reinforcement Learning by Simulation (PFN, Toyota)



# Ogatsuma and Ichise

KIT, NII, Waseda

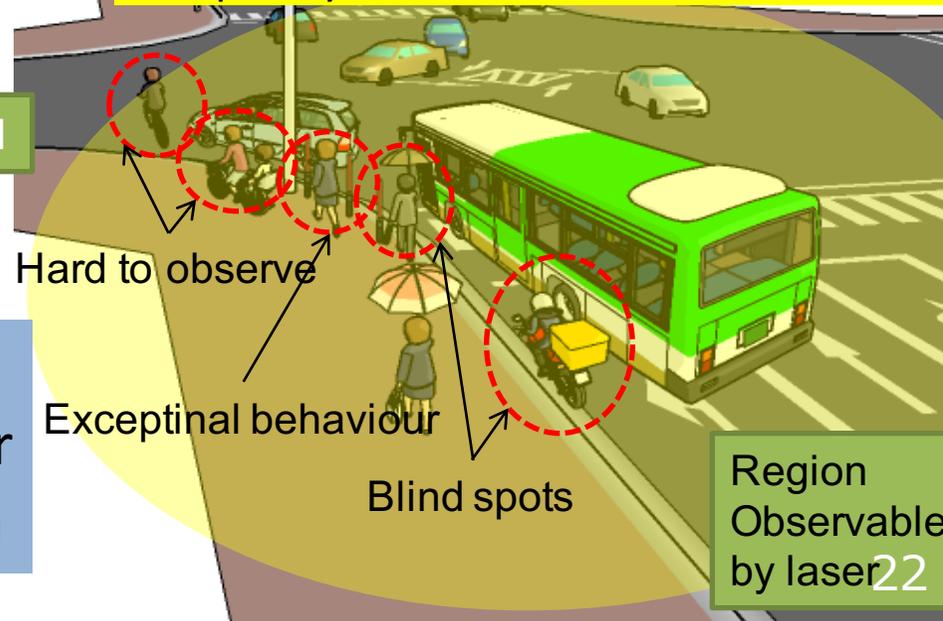


**Objective:** Self-Driving with context understanding, Risk avoidance with prediction

**Method:** Hybrid system of data-driven AI and knowledge-driven AI

**Implementation:** Close cooperation on field tests with auto-makers

- Use of diverse sensors, efficient hardware specialized for self-driving, real-time response (ADAS10ms-500ms)
- Use of Ontology to measure the complexity index of context



Data-driven AI and Knowledge-driven AI which can explain, Efficient inferences for being feasible in the practical application

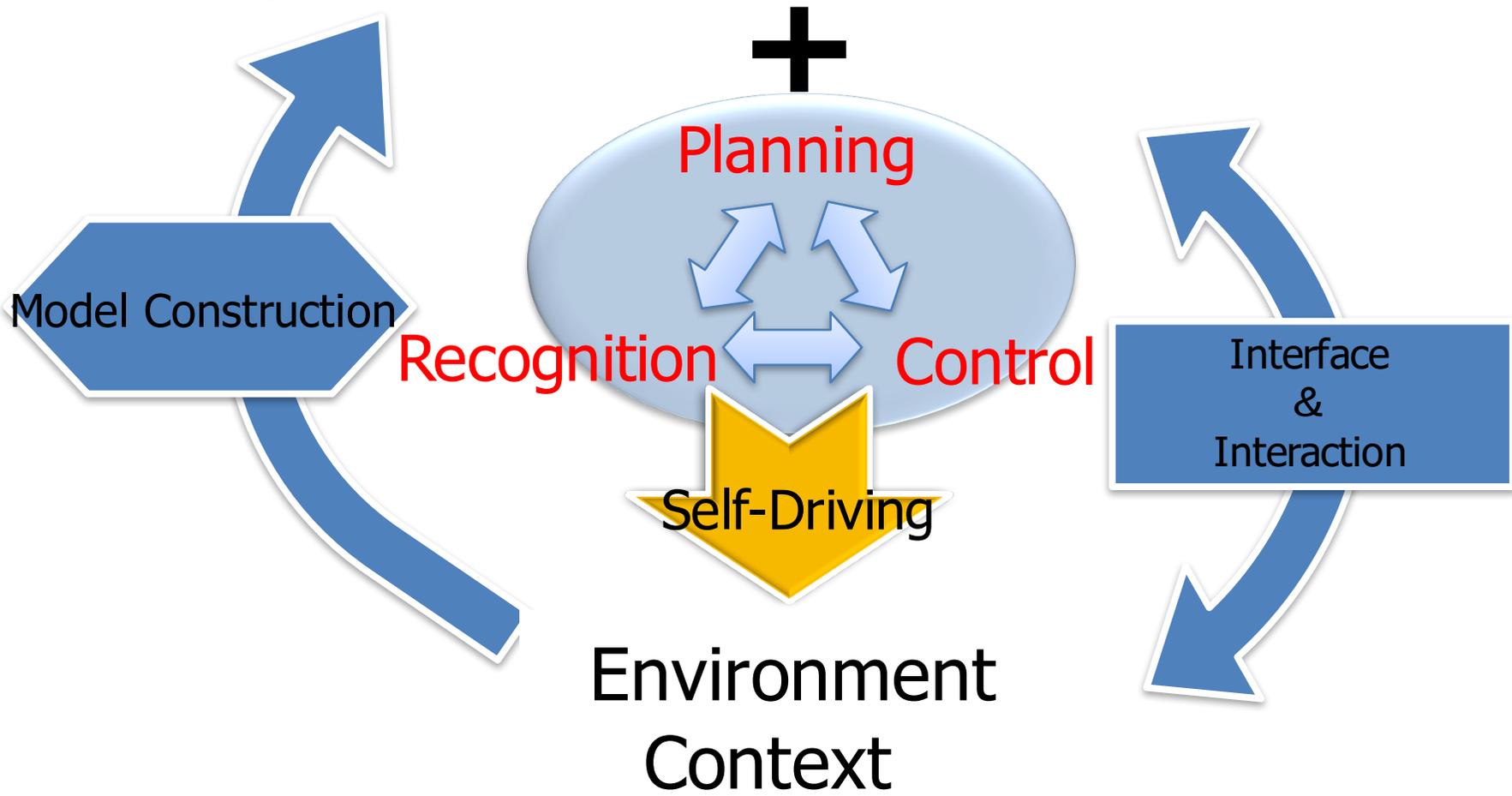
Region Observable by laser22

# Intelligent agent which cohabits and cooperates with human

Human Model (faces, bodies, movements, location)

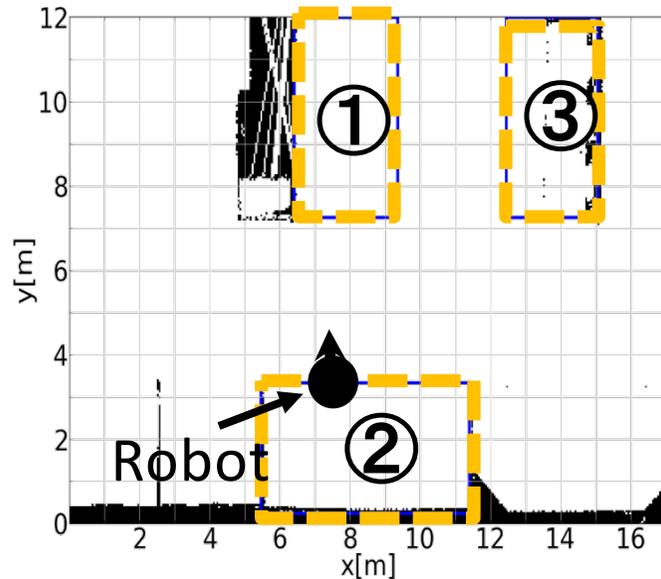
Model of Movements (map, paths, changes)

Model of Environment (shapes, locations, changes )



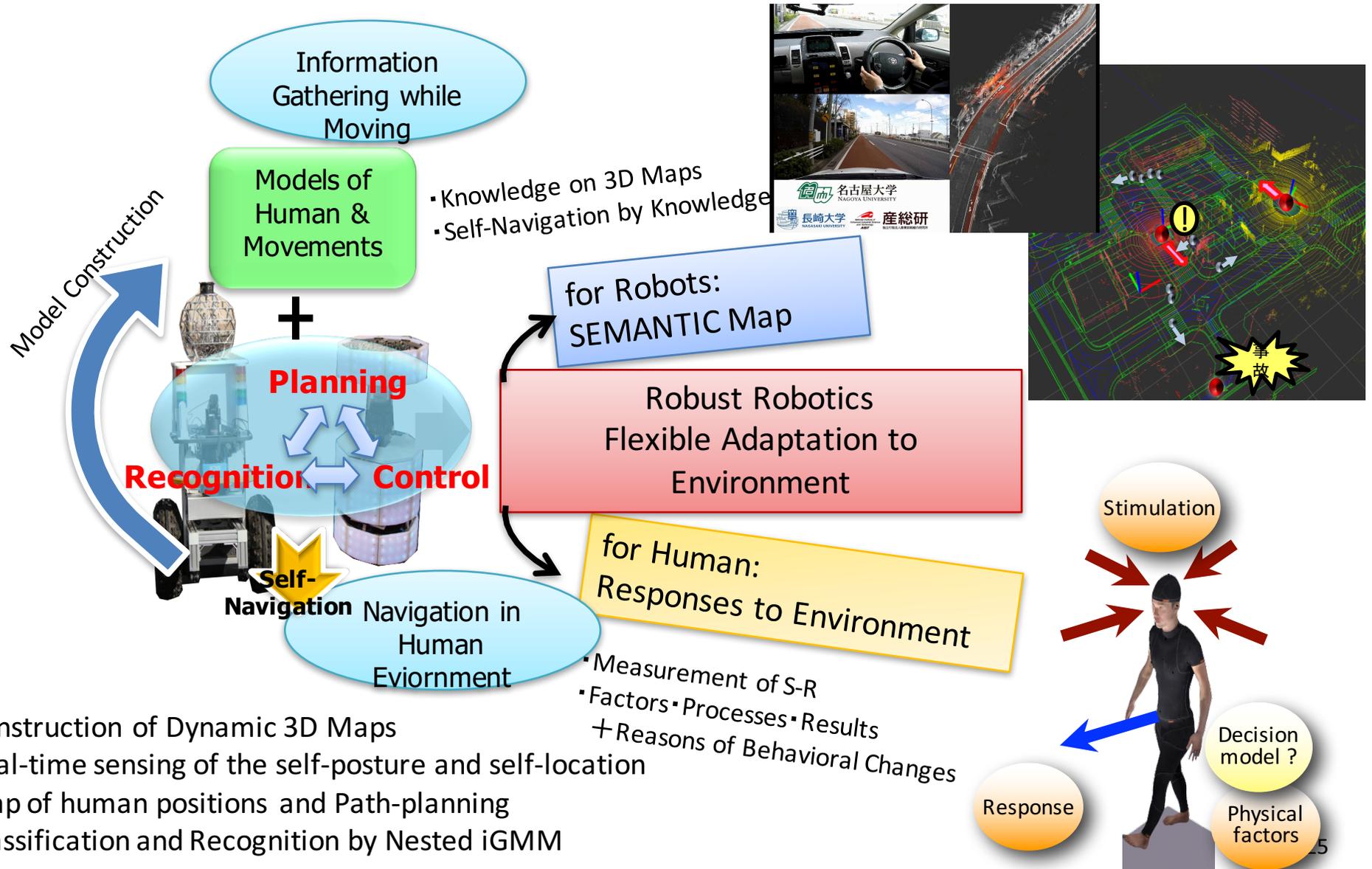
# Path-Planning based on **dynamic maps of human movements**

## ■ Self-Location + Environment Sensing → **Semantic Map**



- **Acquisition of Plausible paths through observation of Human movements**
  - Conventional Methods -- Path-planning based on maps of shapes of rooms and empty spaces
  - Avoiding other booths, selection of plausible paths used by human

# Self-Navigating Robot (Elderly Care) and Self-Driving



- Construction of Dynamic 3D Maps
- Real-time sensing of the self-posture and self-location
- Map of human positions and Path-planning
- Classification and Recognition by Nested iGMM

# AI for Big Sciences

## Robot Scientist in Biology



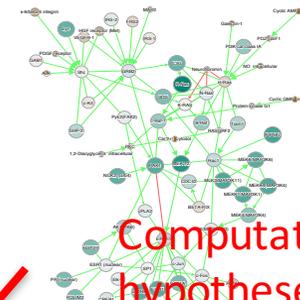
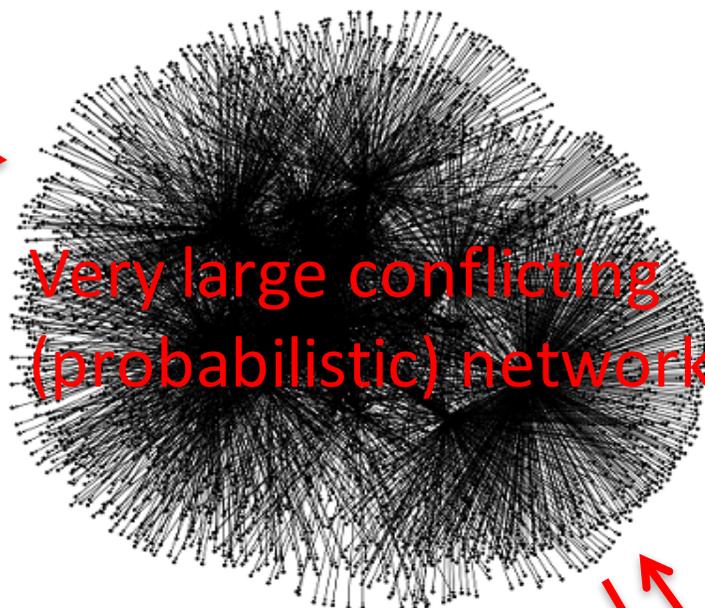
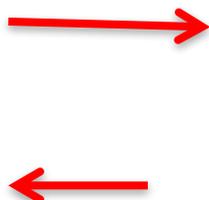
# Big Mechanism: Reading-Assembly-Explanation



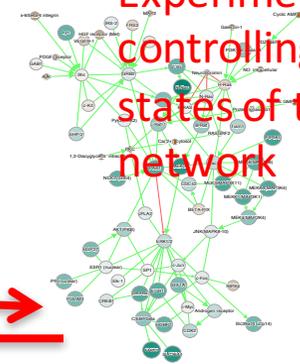
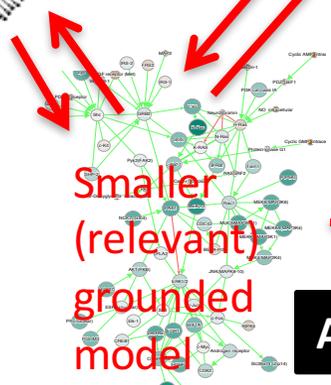
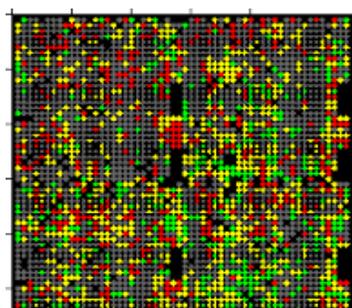
Reading

Assembly

Explanation



Computational hypotheses/  
wet lab  
Experiments  
controlling  
states of the  
network



A.Rzhetsky(U.Chicago)

# Open Science-The Need for Text Mining

## Types of documents

- Full papers
- Abstracts
- Reports, discharge summaries
- EMR
- Textbooks, monographs
- Grey content, online discussion forums

## MEDLINE

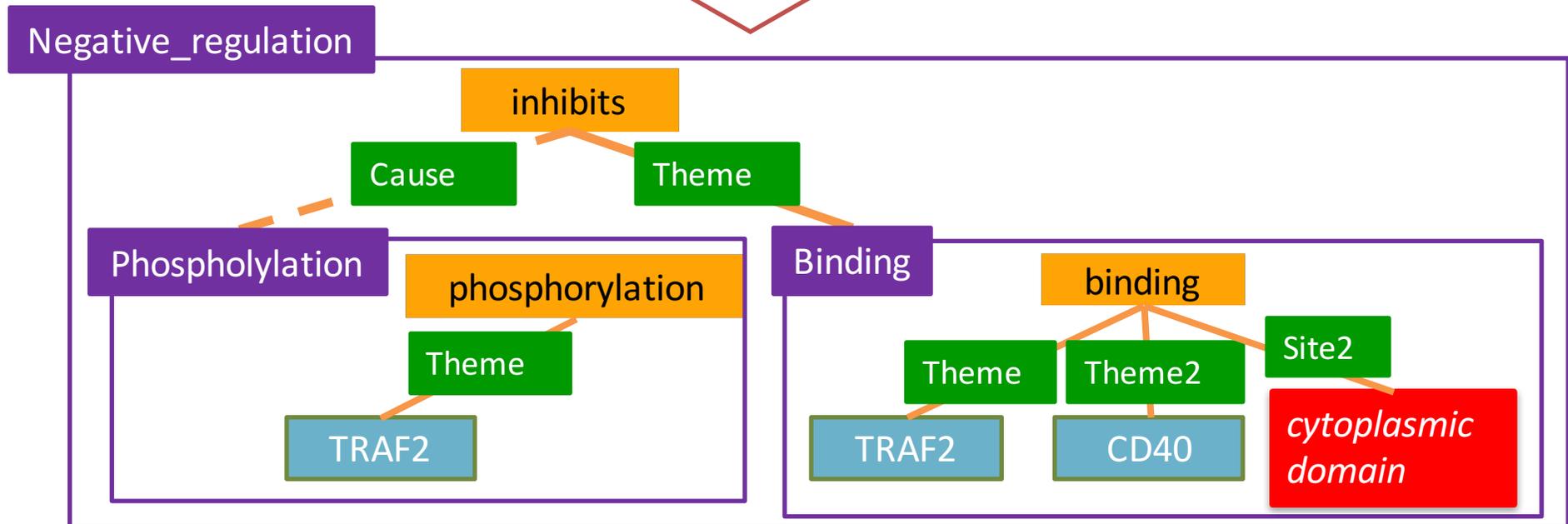
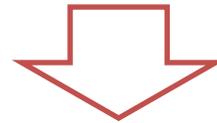
- 2005: ~14M
- 2009: ~18M
- 2013: ~22M
- 2015: ~26M

Overwhelming information in textual, unstructured format

# Event E <http://www.nactem.ac.uk/EventMine/>

Finding events ( trigger mentions , event types , and typed arguments including locations ) involving genes or gene products

... In this study we hypothesized that the phosphorylation of TRAF2 inhibits binding to the CD40 cytoplasmic domain. ...



# Finding Evidence -EuropePubMed Central

- Currently: runs on **2,550, 328 full texts**
- **82,198,474 facts** in 38,411,661 sentences
- Full parsing used a version of Enju (Mogura)
- **Parsing pipeline run on 60 machines at EBI ~30 days**



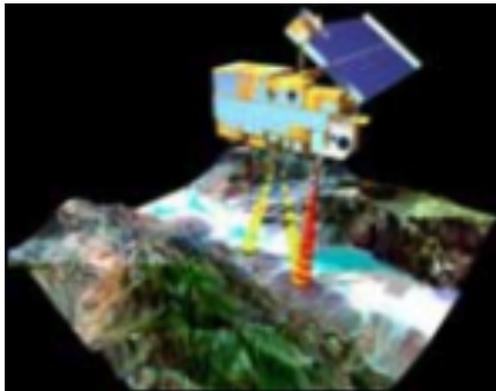
<http://labs.europepmc.org/evf>

# Understanding of Satellite Images

## Geology and Environment Science

# *Archives of Satellite Images*

- ASTER@AIST ~ **1PB**
  - **Images of all regions in the world from 1999**
- Landsat-8@Amazon S3 ~ 500 TB
  - **Images fro 20132013年**
  - Open data

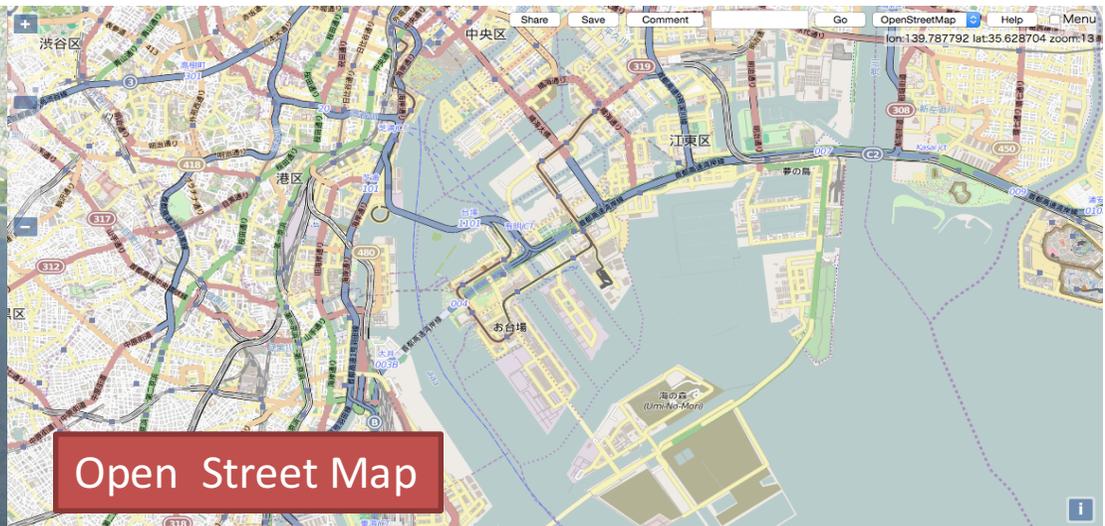
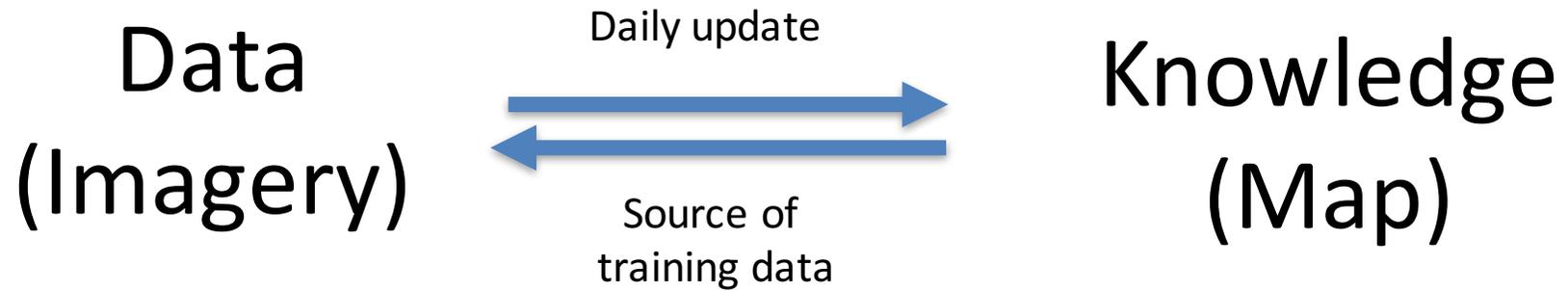


ASTER on NASA' TERRA



OLI on Landsat-8

# Mapping from Images to Maps



# Images by Landsat 8

<http://landbrowser.geogrid.org/landbrowser/index.html>



By sliding the bar, we can jump to the past

# Comparison of the image with the past image of the same region



Red Circles - Significant changes

# Mega solar detection by DNN





Database constructed by human specialist

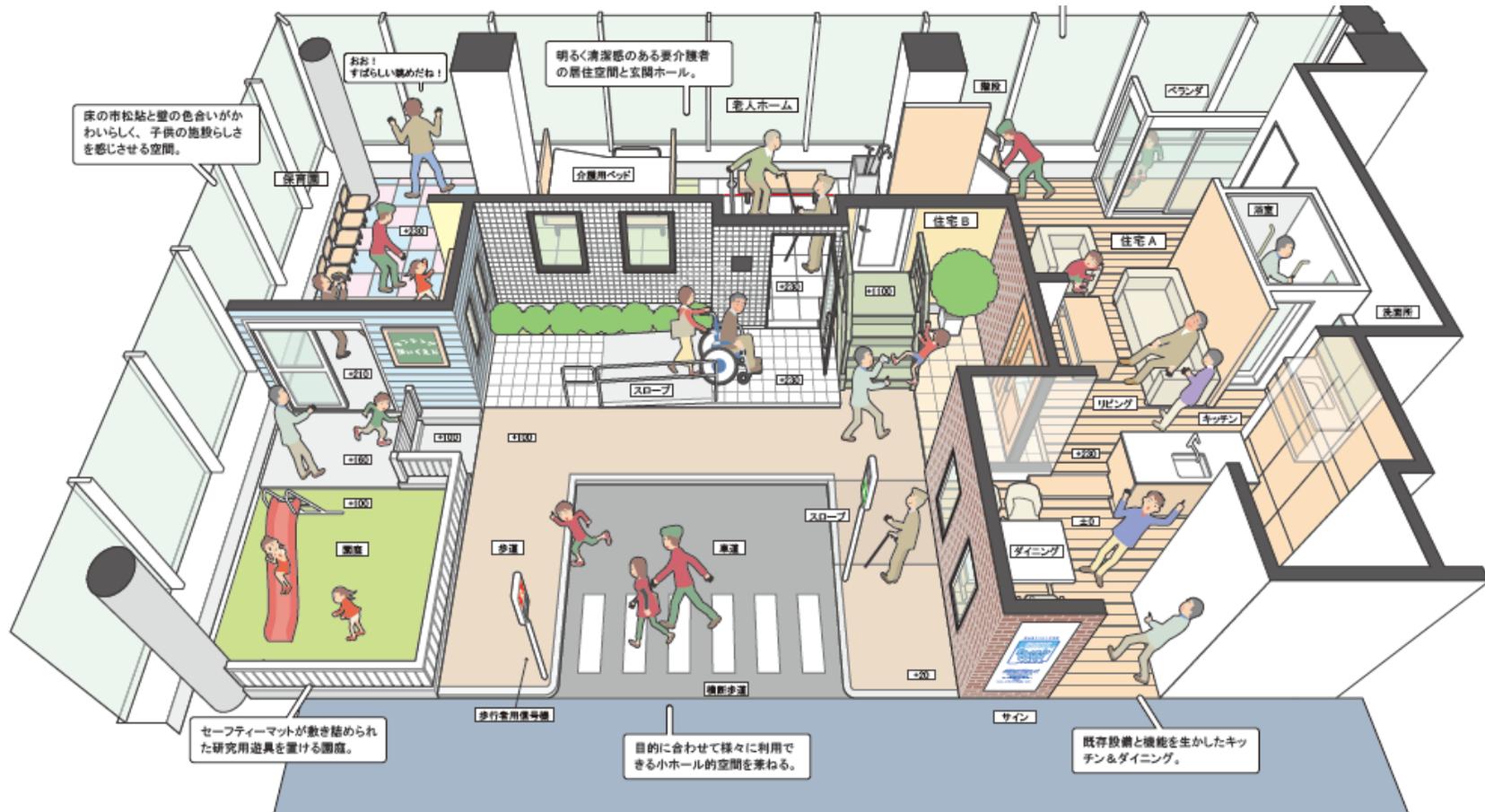


Classification by DNN



# AI for Human Life

## Living Intelligence Complex



# Outline

- AI Platforms
  - Background and the Rationale
- Examples of Research at AIRC
  - Self-navigation and Self-driving
  - AI for Manufacturing (Intelligent Robots, Plant Maintenance,..)
  - AI for Human Life (Elderly care, Medical care, Retailing, ..)
  - AI for Big Sciences (Bio-Medical Sciences, Environment Science, ..)
- **Computational Infrastructures**
- Conclusions and Comments

# Computation Infrastructures



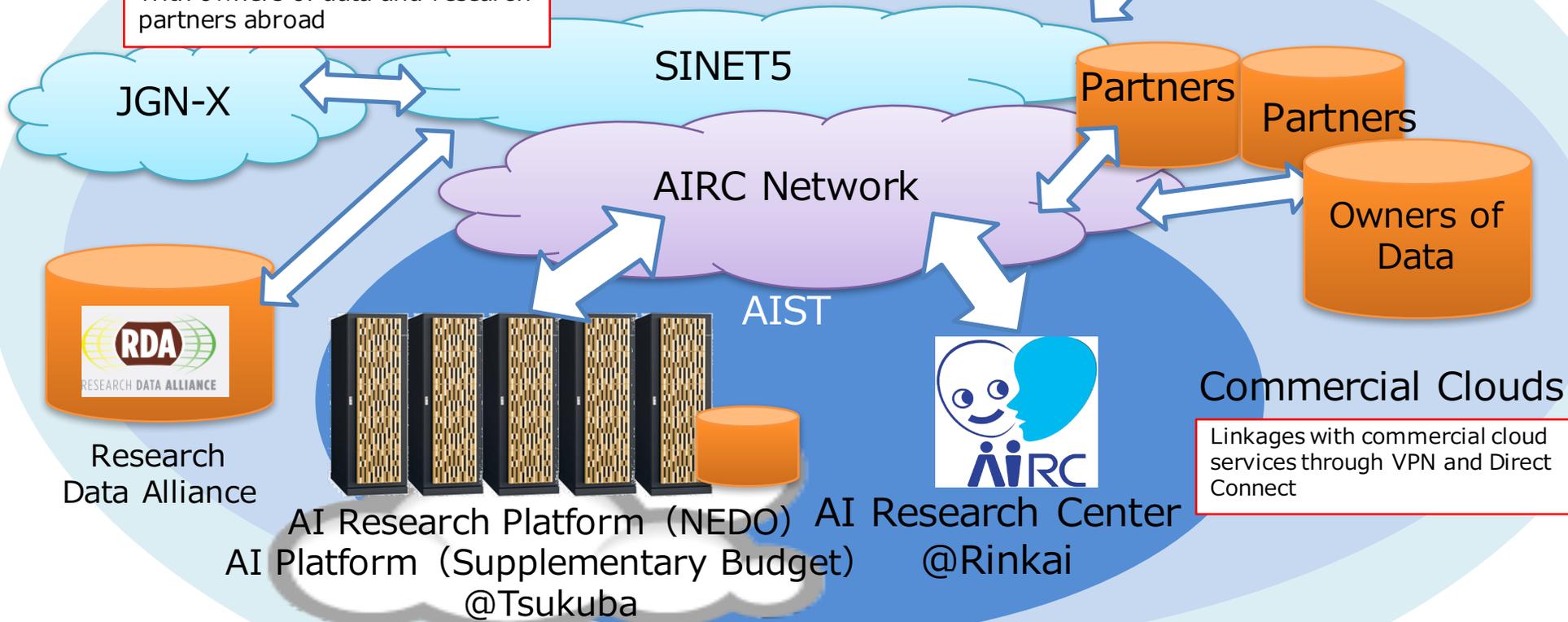
**Partners Abroad**

Stable and secure linkages  
With owners of data and research  
partners abroad

**Domestic  
Partners**

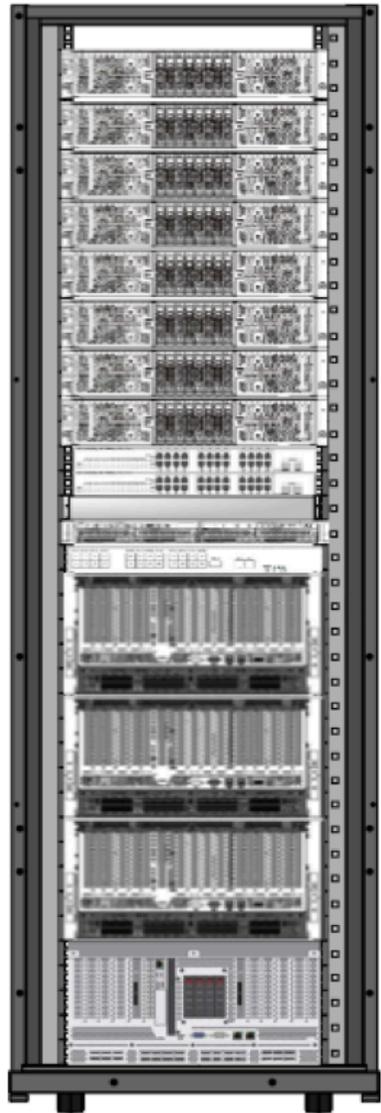


powered by  
**amazon  
web services**



High-Speed Data Transmission (100Gbps) and TV conference facilities through dark fibres

# AI Research Clusters



## GPGPU Server × 8

- 2 Sockets, 28 Cores
- 512GB Memory
- GPGPU accelerator × 4
  - 3,072 CUDA Cores
  - 12GB GDDR5 Memory
  - 7TFlops(Single Precision)

## 10GB Ethernet Switch

## Large Scale Memory Server

- 16 Sockets, 256 Core
- Symmetric Multiprocessing
- 12TB Memory Space

- Members of the NEDO Project use at AIRC
- Highly Efficient Deep Learning by using 32 GPU accelerators
- 16TB main memory for real time processing of large data and large scale simulation
- At service from June, 2016

# Outline

- AI Platforms
  - Background and the Rationale
- Examples of Research at AIRC
  - Self-navigation and Self-driving
  - AI for Manufacturing (Intelligent Robots, Plant Maintenance,..)
  - AI for Human Life (Elderly care, Medical care, Retailing, ..)
  - AI for Big Sciences (Bio-Medical Sciences, Environment Science, ..)
- Computational Infrastructures
- **Conclusions and Comments**

# Conclusions and Comments



- Concerted Efforts, Beyond individual researches
- Public vs. Private, Medical care, Elderly care, etc.
- Integration of Computation-Intensive and Data-Intensive Architectures
- Ownership of Data, Software, and Learned ML models
- Incentives for providing data and software