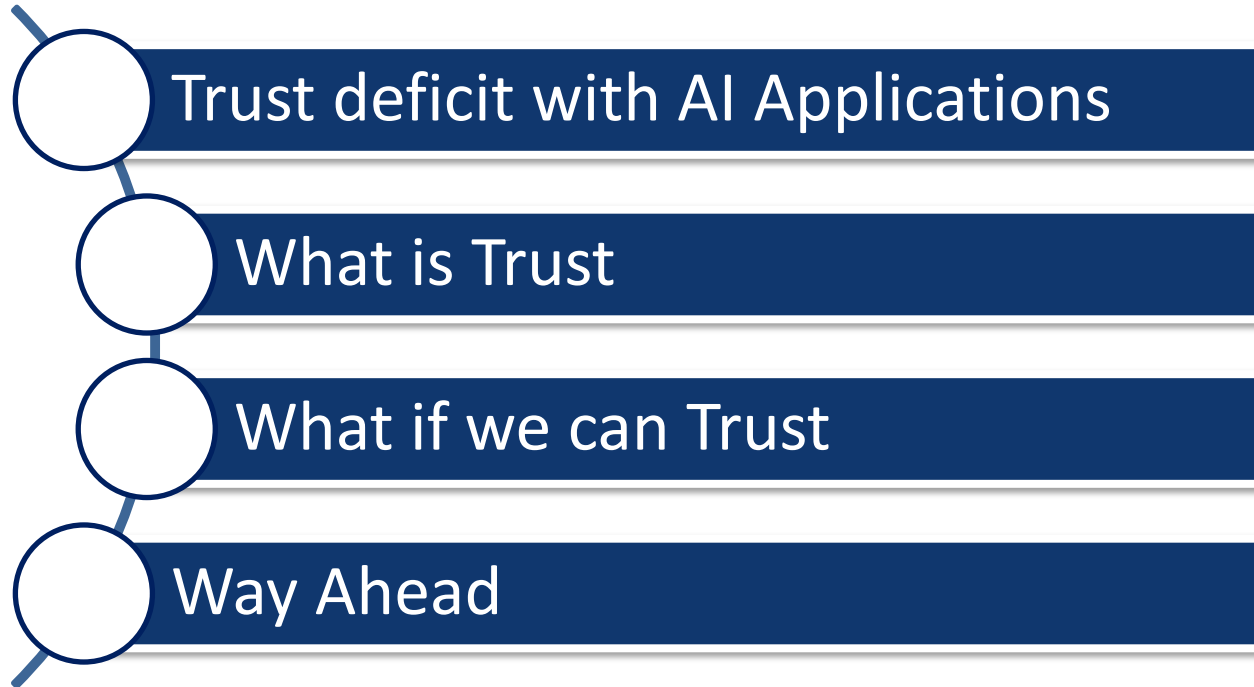


Trust & Explainability for Artificial Intelligence

- Lt Col Ashutosh Verma



Scope



Trust

- Google autonomous car crash.
- Tesla autopilot FATAL car crash.
- Risk assessment for committing crime racially discriminates subjects - COMPASS.
- Predictive policing – racially profiles the area.

Our fear and mistrust is based on such failures.



Trust in Military Applications



future of life

Technology is giving life the potential to flourish like never before...

News: AI Biotech Nuclear Climate Partner Orgs

AUTONOMOUS WEAPONS: AN OPEN LETTER FROM AI & ROBOTICS RESEARCHERS

Autonomous weapons select and engage targets without human intervention. They might include, for example, armed quadcopters that can search for and eliminate people meeting certain pre-defined criteria, but do not include cruise missiles or remotely piloted drones for which humans make all targeting decisions. Artificial intelligence (AI) technology has reached a point where the deployment of such systems is — practically if not legally — feasible within years, not decades, and the stakes are high: autonomous weapons have been described as the third revolution in warfare, after gunpowder and nuclear arms.

Many arguments have been made for and against autonomous weapons, for example that replacing human soldiers by machines is good by reducing casualties for the owner but bad by thereby lowering the threshold for going to battle. The key question for humanity today is whether to start a global AI arms race or to prevent it from starting. If any major military power pushes ahead with AI weapon development, a global arms race is virtually inevitable, and the endpoint of this technological trajectory is obvious: autonomous weapons will become the Kalashnikovs of tomorrow. Unlike nuclear weapons, they require no costly or hard-to-obtain raw materials, so they will become ubiquitous and cheap for all significant military powers to mass-produce. It will only be a matter of time until they appear on the black market and in the hands of terrorists, dictators wishing to better control their populace, warlords wishing to perpetrate ethnic cleansing, etc. Autonomous weapons are ideal for tasks such as assassinations, destabilizing nations, subduing populations and controlling killing a nation or ethnic group. We therefore believe that a military AI arms race would not be beneficial for humanity. There



The New York Times | <http://nyti.ms/2eApcwz>

The Pentagon's 'Terminator Conundrum': Robots That Could Kill on Their Own

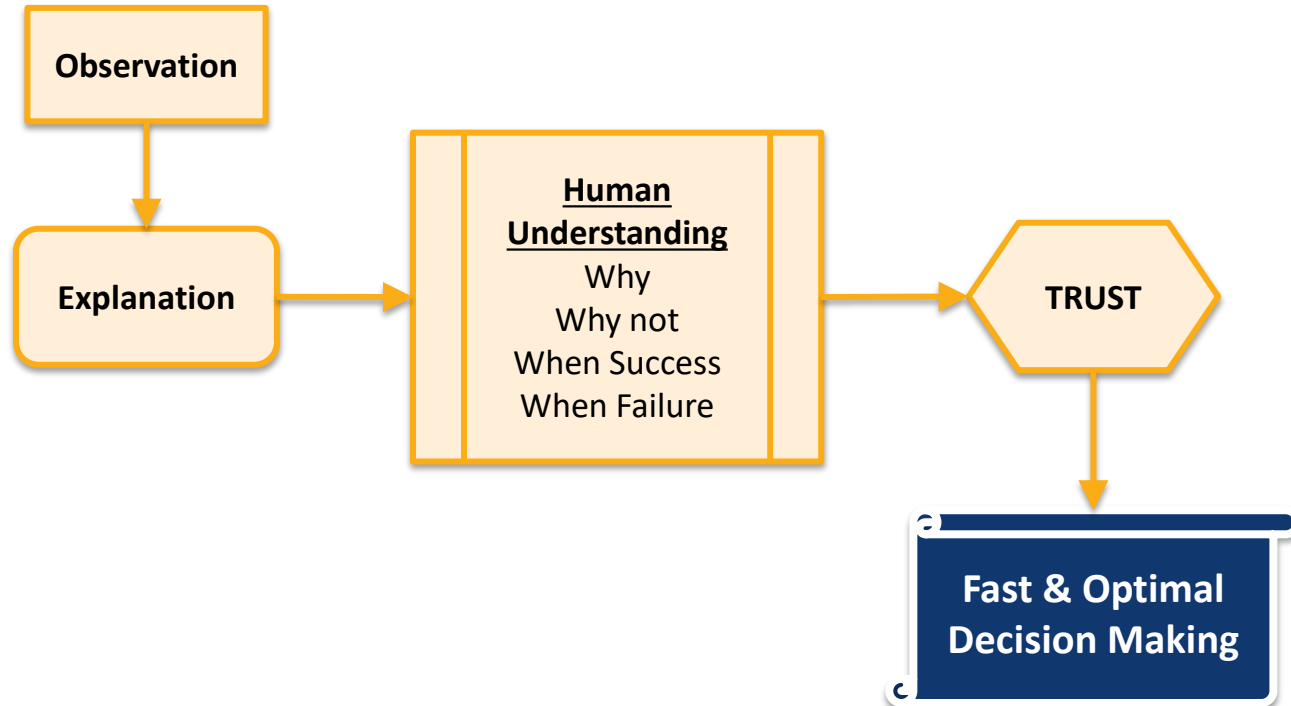
The United States has put artificial intelligence at the center of its defense strategy, with weapons that can identify targets and make decisions.

By MATTHEW ROSENBERG and JOHN MARKOFF | OCT. 25, 2016

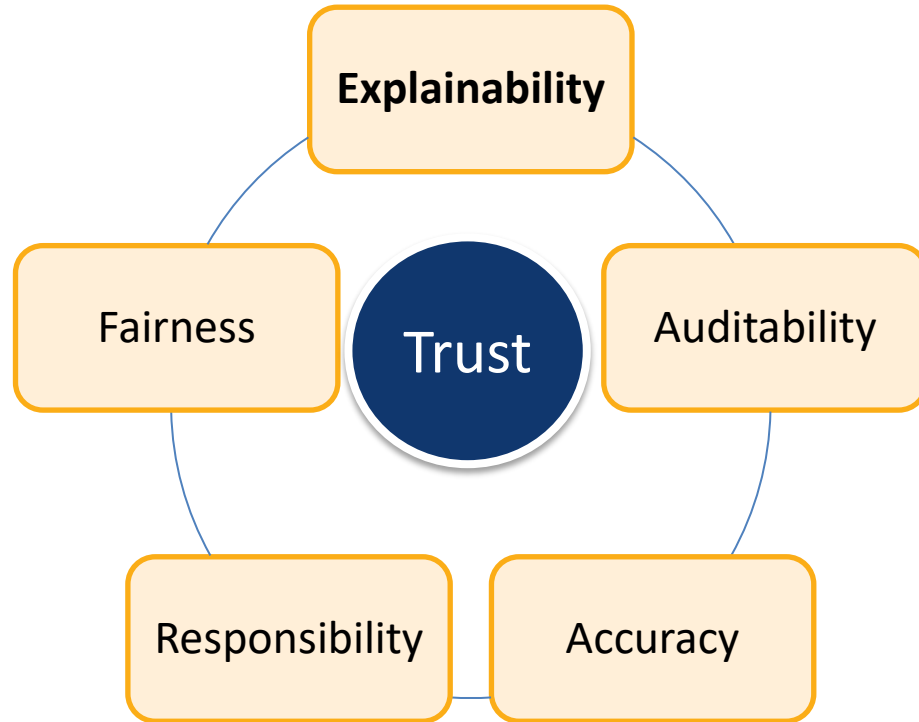
Signatories:-

Stephen Hawking, Steve Wozniak, Stuart Russell, Nils Nilsson and 20,000+

Necessity of Explanation

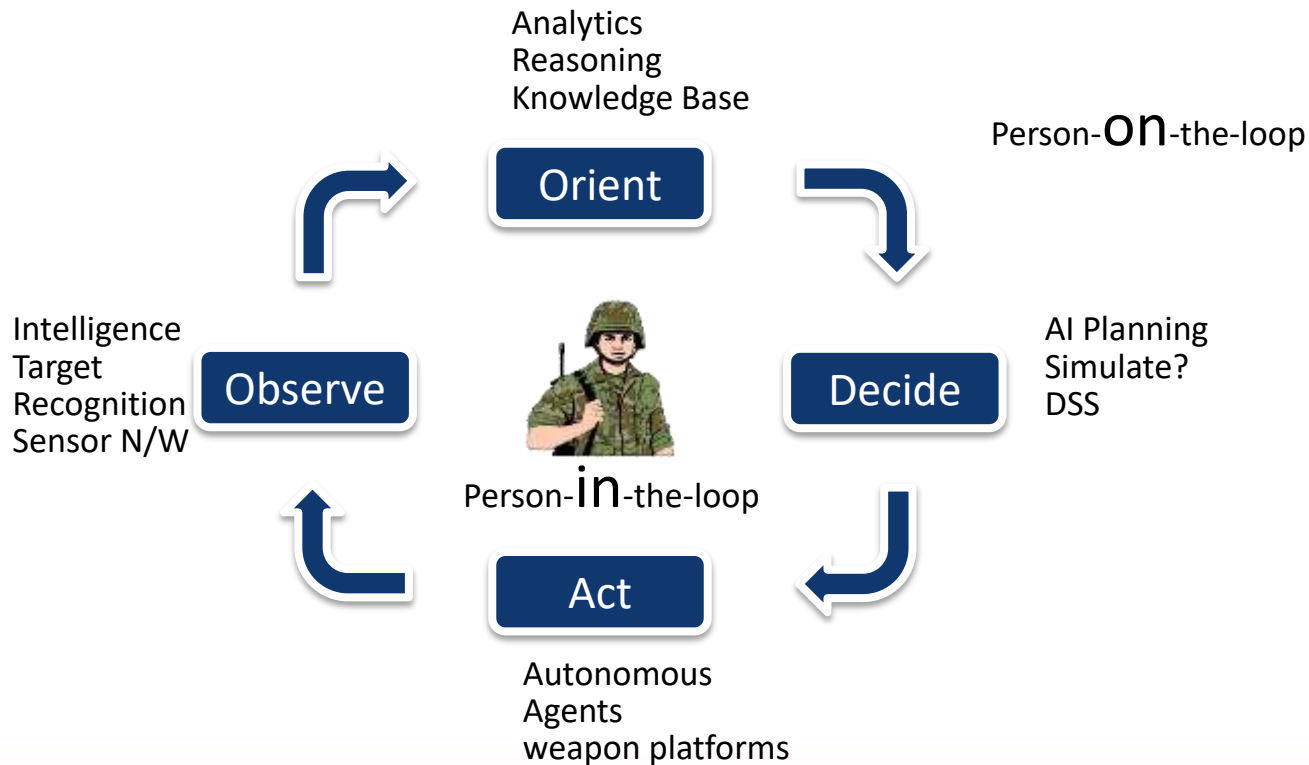


Trust in AI Algorithms



Military Decision Making

The OODA Loop



China embracing AI

Baidu welcomes China's military to join China Brain project on AI systems

Posted in [Software](#) on 9th Mar, 2015 08:16 PM by [AlexMuller](#)

Robin Li, founder and CEO of Chinese online search giant Baidu, has proposed a state-level project on artificial intelligence (AI) systems called "China Brain" and hopes to gain support from the military, reports our Chinese-language sister newspaper Want Daily.

RATE THIS

  +48

 Share

Is China Really Building Missiles With Artificial Intelligence?

There are real limits on the amount of AI acceptable to navy commanders.

By [Abhijit Singh](#)
September 21, 2016



With rising security challenges in the global commons, there is growing interest in the subject of "intelligent" weapons systems. This is especially so in

ASIA PACIFIC

China Announces Cuts of 300,000 Troops at Military Parade Showing Its Might

[点击查看英文中文](#) | [Read in Chinese](#)

By [EDWARD WONG](#), [JANE PERLEZ](#) and [CHRIS BUCKLEY](#) | SEPT 2, 2015



Imagenet Challenge 2016

- All category winners were teams from China.
- 33 of the 82 teams from China.

Ref:-

<http://thediplomat.com/2016/09/is-china-really-building-missiles-with-artificial-intelligence/>

<http://www.planettechnews.com/software/software/baidu-welcomes-china-s-military-to-join-china-brain-project-on-ai-systems>

http://www.nytimes.com/2015/09/03/world/asia/beijing-turns-into-ghost-town-as-it-gears-up-for-military-parade.html?_r=0

'Strength' of Chinese Academia



National University of
Defence Technology



Naval University of
Engineering



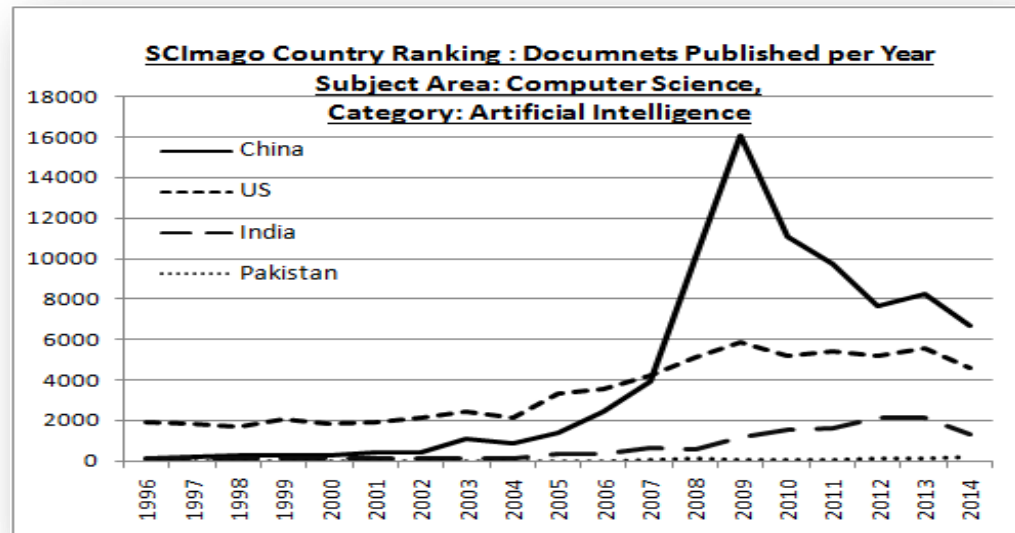
Armoured Engineering
Academy



Harbin Institute of
Technology



Chinese Association for
Artificial Intelligence

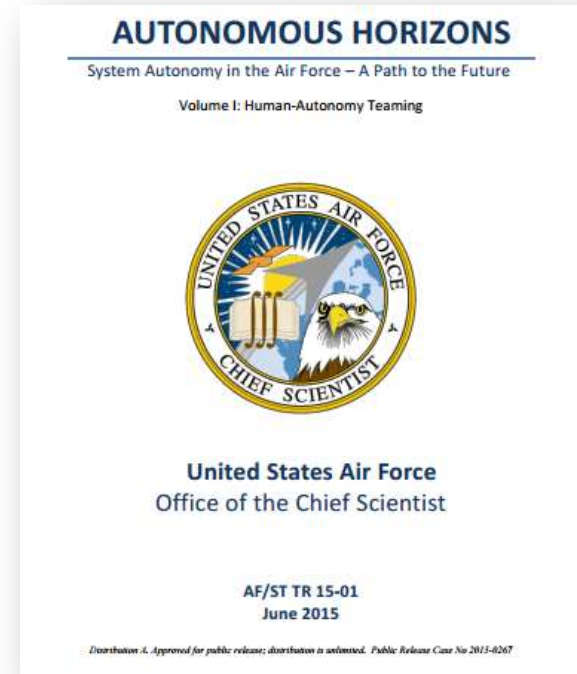


863 Programme - National High Tech R&D Programme

Project 985 to promote and develop the Chinese higher
education system

“Autonomous software working synergistically with the innovation of empowered airmen.”

“Authority and responsibility for warfare in the hands of airmen while creating tools that enhance their situation awareness and decision-making, speed effective actions, and bring needed extensions to their capabilities.”



Ref: <http://www.af.mil/Portals/1/documents/SECAF/AutonomousHorizons.pdf?timestamp=1435068339702>

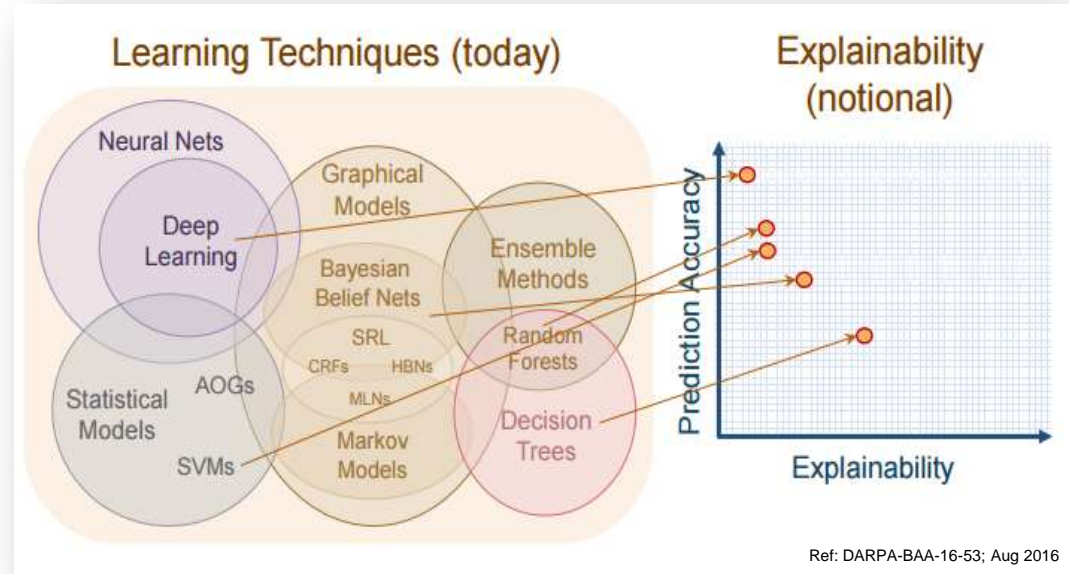
USA: DARPA

- DARPA Grand Challenges:-
 - Autonomous Driving.
 - Autonomous Cyber Defence.
 - Autonomous Spectrum Collaboration.
- Fast Lightweight autonomy.
- BOLT – Broad operational language translator.
- XAI – Explainable AI.



Challenges

- Complexity of algorithms.
- Input data quality.
- Human introduced bias, e.g. racial profiling.



- **XAI Research - DARPA-BAA-16-53**
- Possible approaches:-
 - AI reasoning, planning and learning.
 - IoT to improve data gathering to remove human induced bias.

Trust and explainability necessary *ab initio* for military application.

“Trust is the glue that holds an organisation together & the lubricant that moves it forward.”

– Gen Colin Powell

Thank You

