DEEP LEARNING AND BIG DATA IN CYBERSECURITY

IARIA PANEL DISCUSSION

PANEL MEMBERS

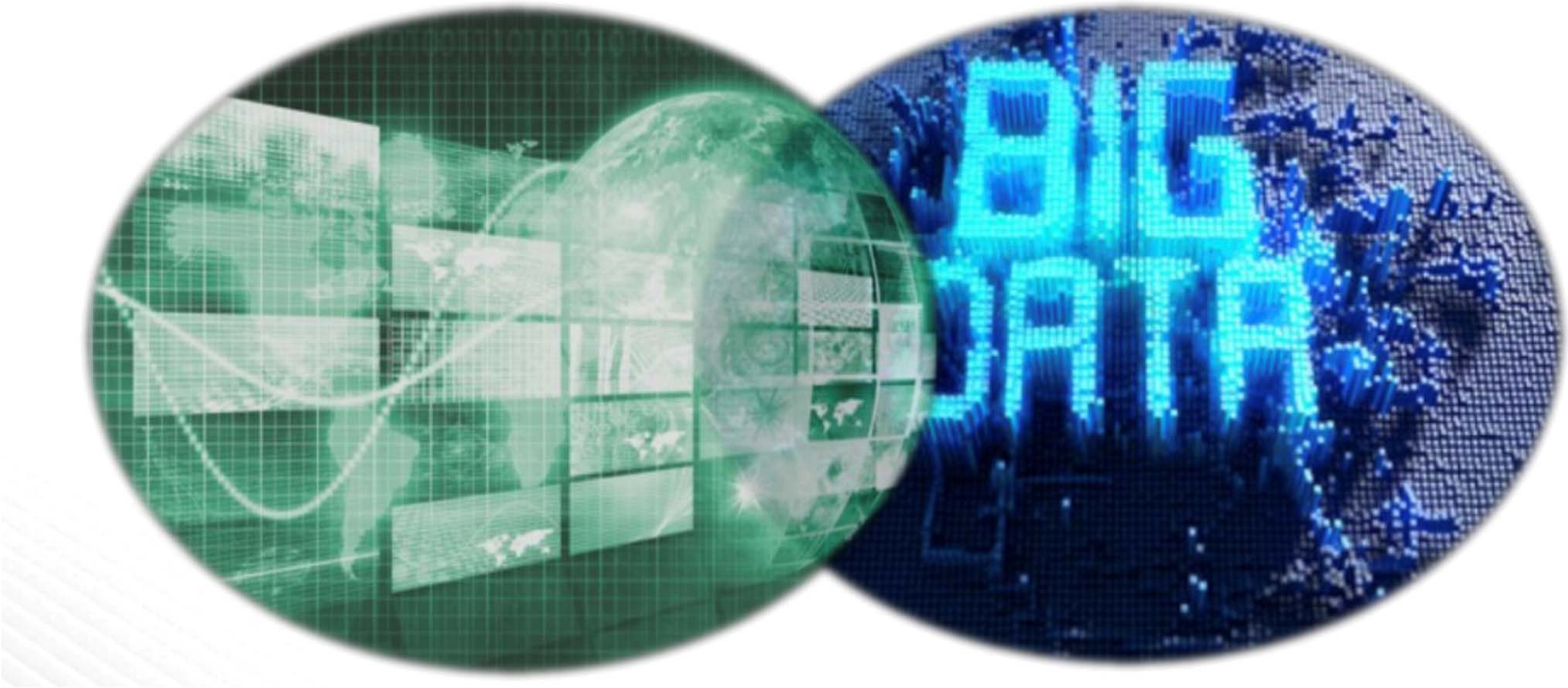
- FELIX W. BAUMANN
- SERAP ŞAHIN
- JIN CUI
- DIMITRIS KARDARAS

VU VRIJE UNIVERSITEIT AMSTERDAM

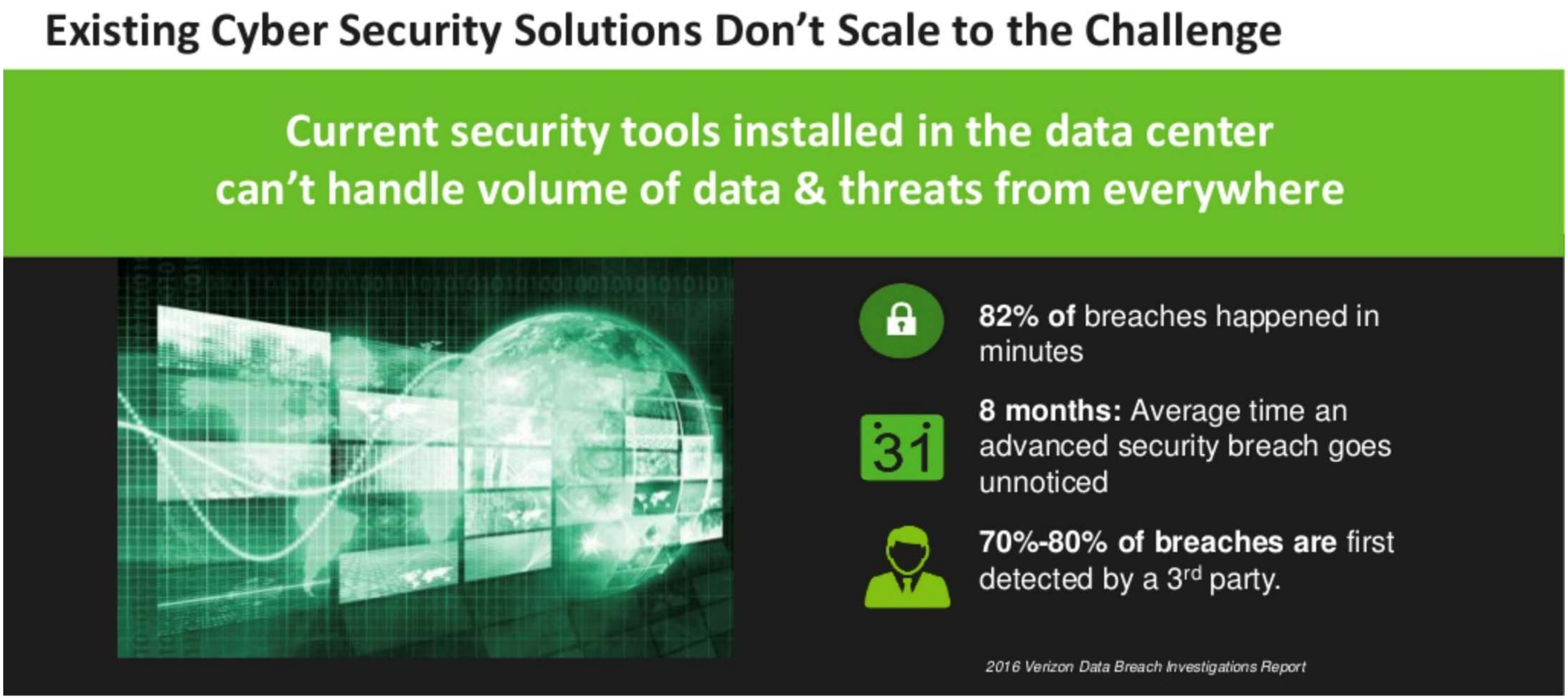
LOOKING FURTHER

DEEP LEARNING AND BIG DATA IN CYBERSECURITY

Digital World Generates Big Data That Security Teams Need to Process



DEEP LEARNING AND BIG DATA IN CYBERSECURITY



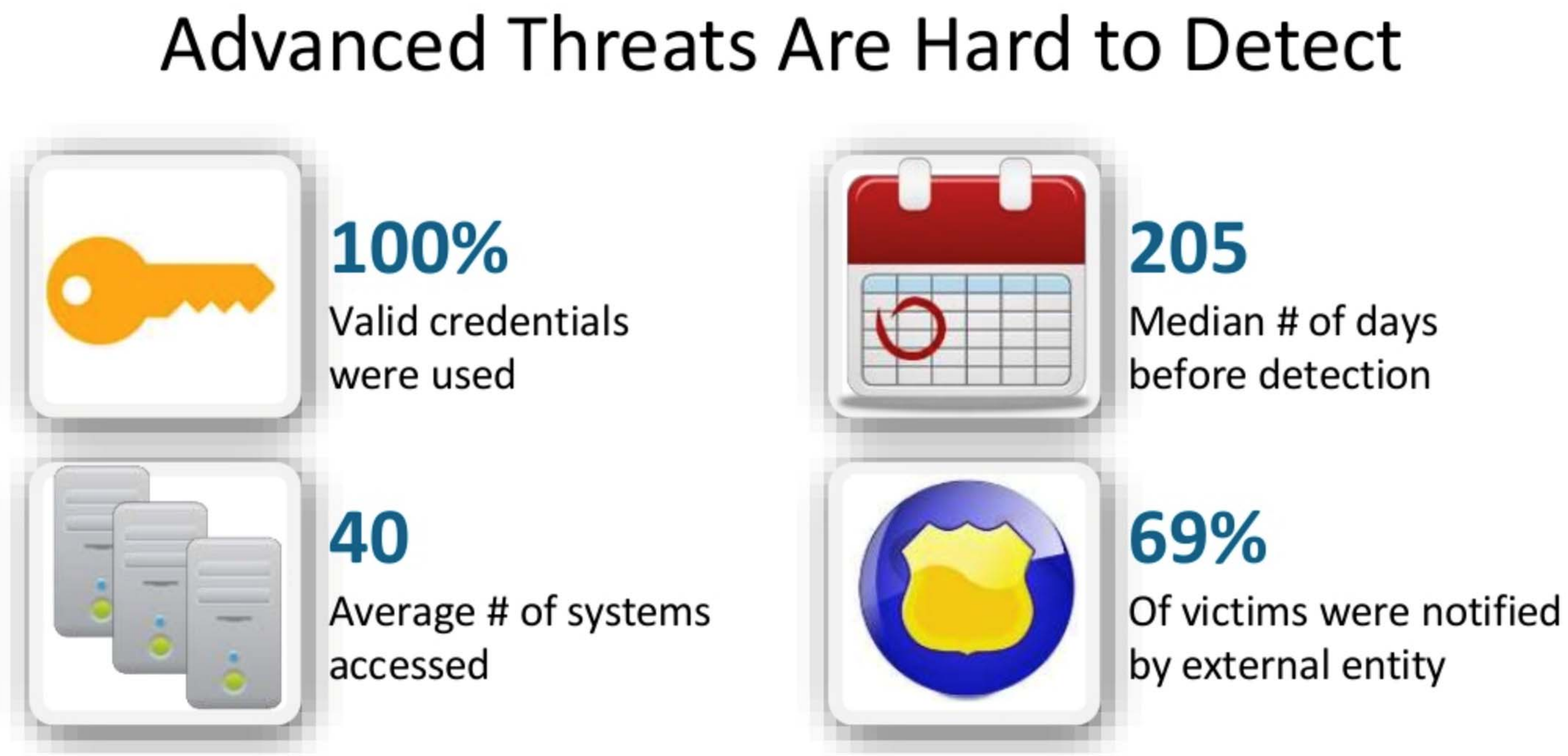
Existing Cyber Security Solutions Don’t Scale to the Challenge

Current security tools installed in the data center can’t handle volume of data & threats from everywhere

- **82% of** breaches happened in minutes
- **8 months:** Average time an advanced security breach goes unnoticed
- **70%-80% of breaches are** first detected by a 3rd party.

2016 Verizon Data Breach Investigations Report

DEEP LEARNING AND BIG DATA IN CYBERSECURITY



Advanced Threats Are Hard to Detect

100%

Valid credentials were used

205

Median # of days before detection

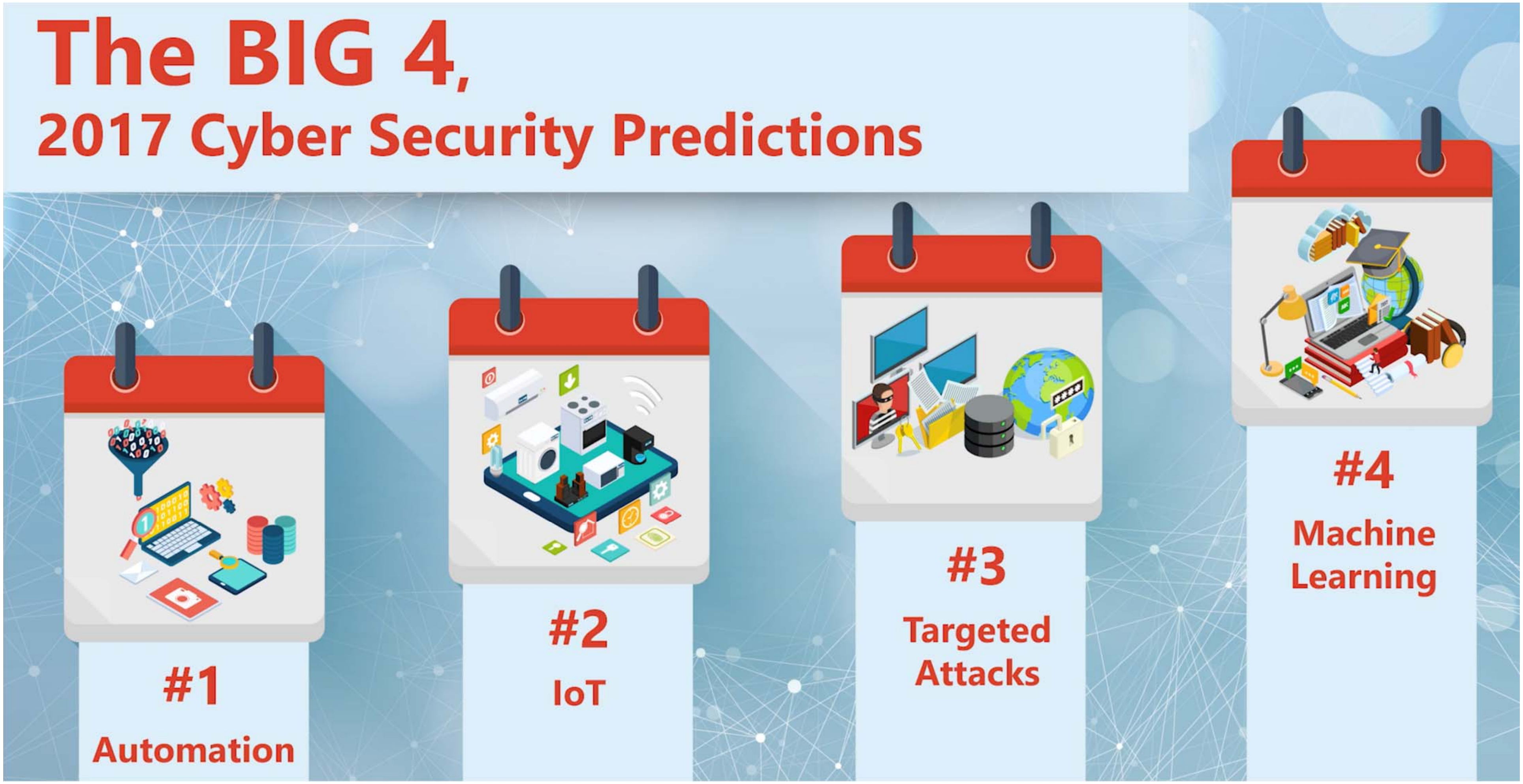
40

Average # of systems accessed

69%

Of victims were notified by external entity

DEEP LEARNING AND BIG DATA IN CYBERSECURITY



DEEP LEARNING AND BIG DATA IN CYBERSECURITY

Felix W. Baumann
TWT GmbH Science & Innovation
Germany

TWT GmbH
Science & Innovation

TWT GmbH
Science & Innovation

Panel on DATA ANALYTICS/CYBER - Deep Learning and Big Data in Cybersecurity

Felix W. Baumann

ADVCOMP, Barcelona, 15.11.2017

Stuttgart
München
Friedrichshafen
Ingolstadt

Ernsthaldenstraße 17
70565 Stuttgart
Telefon: +49.711.215777.0
info@twt-gmbh.de
www.twt-gmbh.de

Image Sources: TWT GmbH

TWT

Project: SePiA.Pro

Service platform for the intelligent optimization of production lines

- http://projekt-sepiapro.de
- BMWI (Germany) funded
- Three years, from 2016
- Five partners
- Smart Service World

Smart Services

Data in Production/Industrial Environment

- Optimised for individual stations
- Data only used for control
- Data integrity/security questionable

Smart Services?

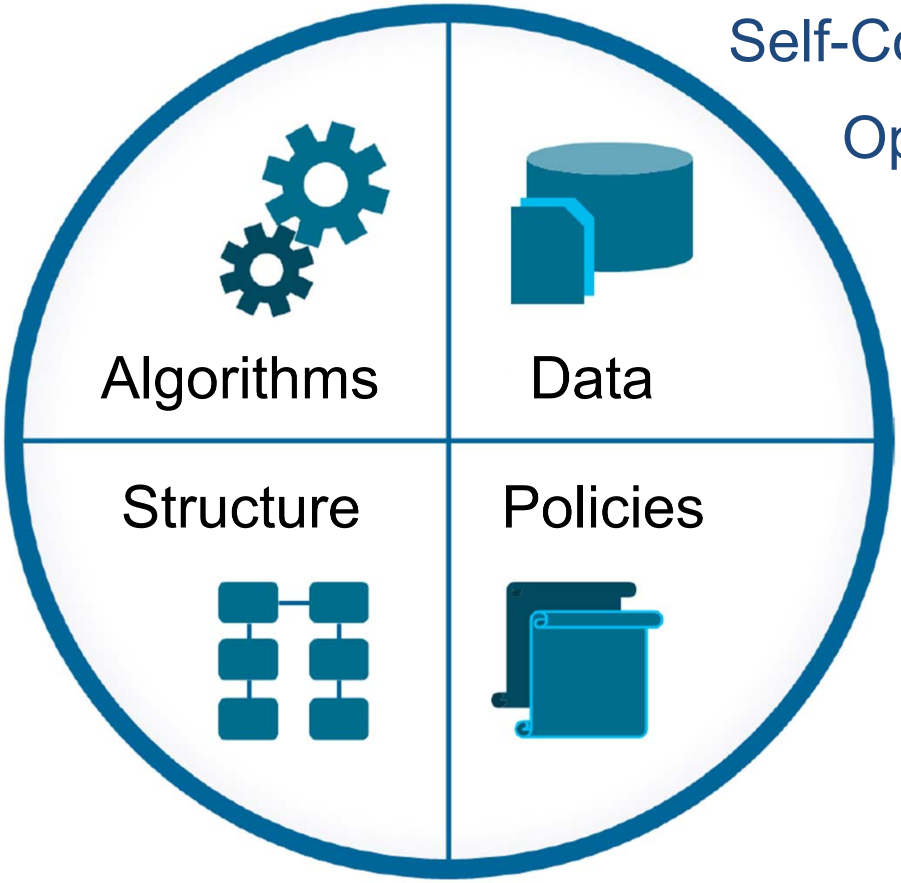
Feasibility of Data Transportation

- Function shipping
- Automatic deployment
- Secure data access

- Web-based, integration/aggregation of data
- Support for analysis and expert involvement

TWT

Deployment Archive



Self-Contained

OpenTOSCA based

Aggregate functionality

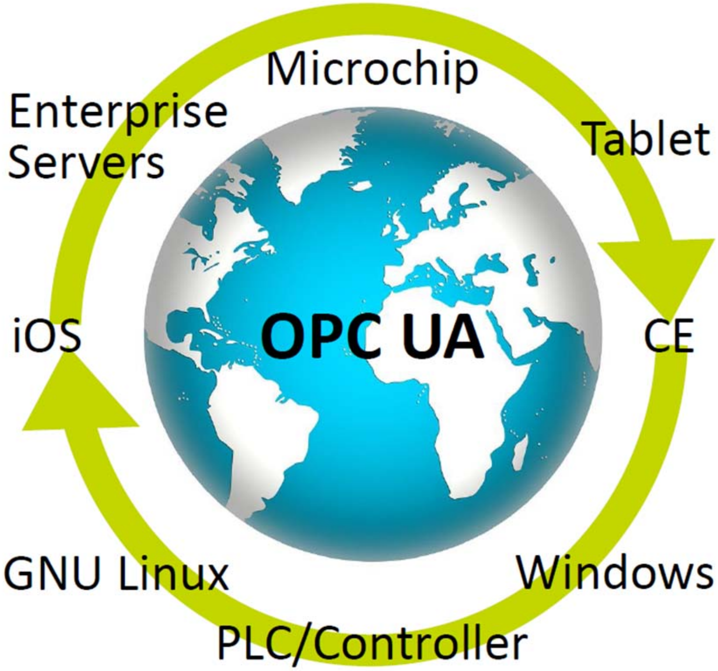
Safeguard data

Data+function shipping

Deploy runtime

TWT

OPC UA



Standardised

Flexible data-model

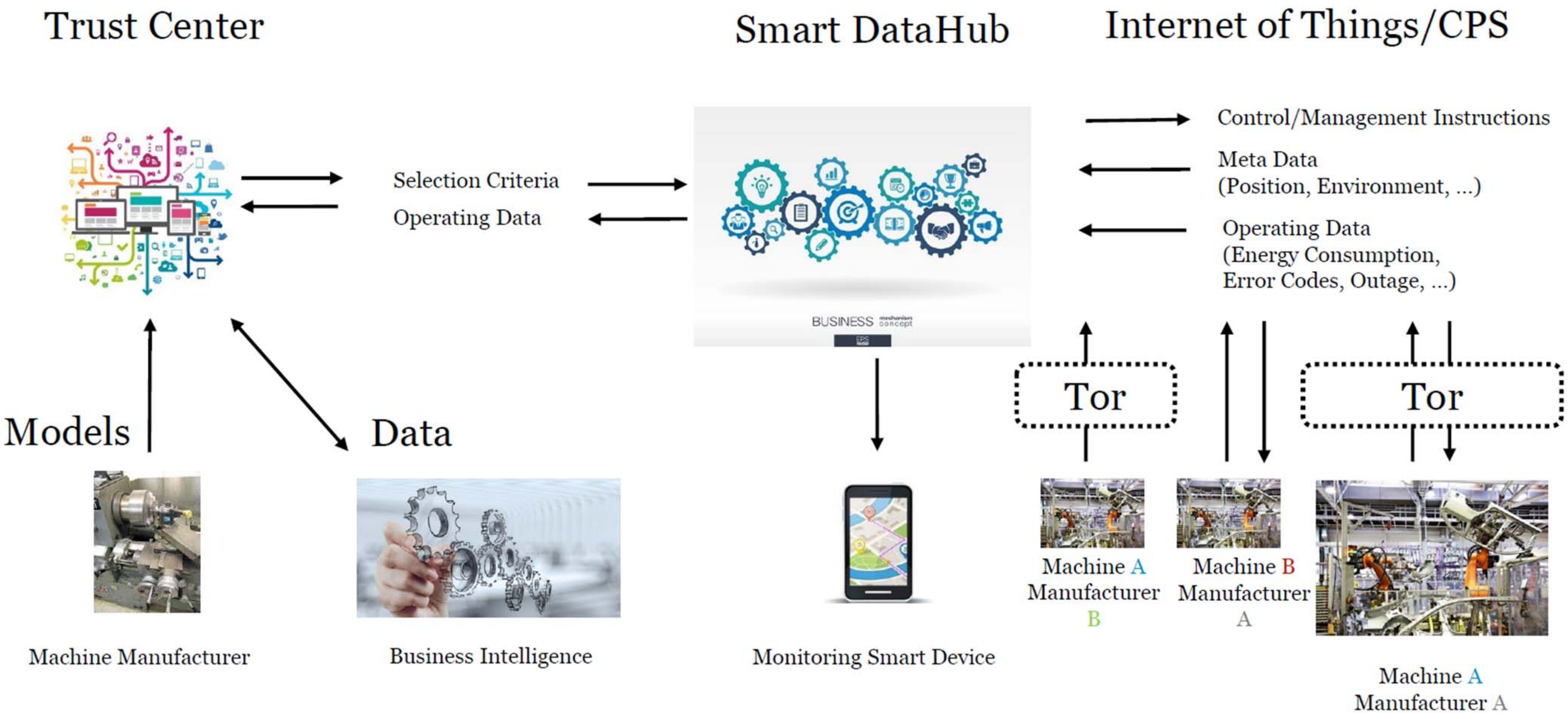
Data and analysis

Compatibility

Extensible

Legacy machinery

Overview



Trust Center

Smart DataHub

Internet of Things/CPS

Selection Criteria

Operating Data

Control/Management Instructions

Meta Data
(Position, Environment, ...)

Operating Data
(Energy Consumption,
Error Codes, Outage, ...)

Tor

Tor

Models

Data

Machine Manufacturer

Business Intelligence

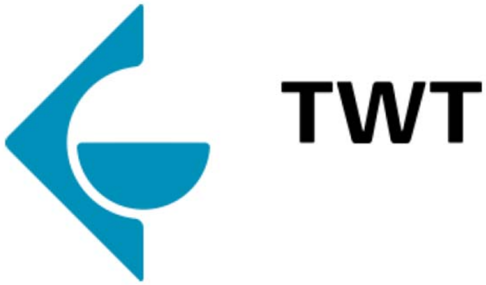
Monitoring Smart Device

Machine A
Manufacturer
B

Machine B
Manufacturer
A

Machine A
Manufacturer A

Project Partners



DEEP LEARNING AND BIG DATA IN CYBERSECURITY

Serap Şahin
Izmir Institute of Technology
Turkey

“A New Solution Direction?”

Panel : Deep Learning and Big Data in Cybersecurity

Asst. Prof. Dr. Serap ŞAHİN
İzmir Institute of Technology

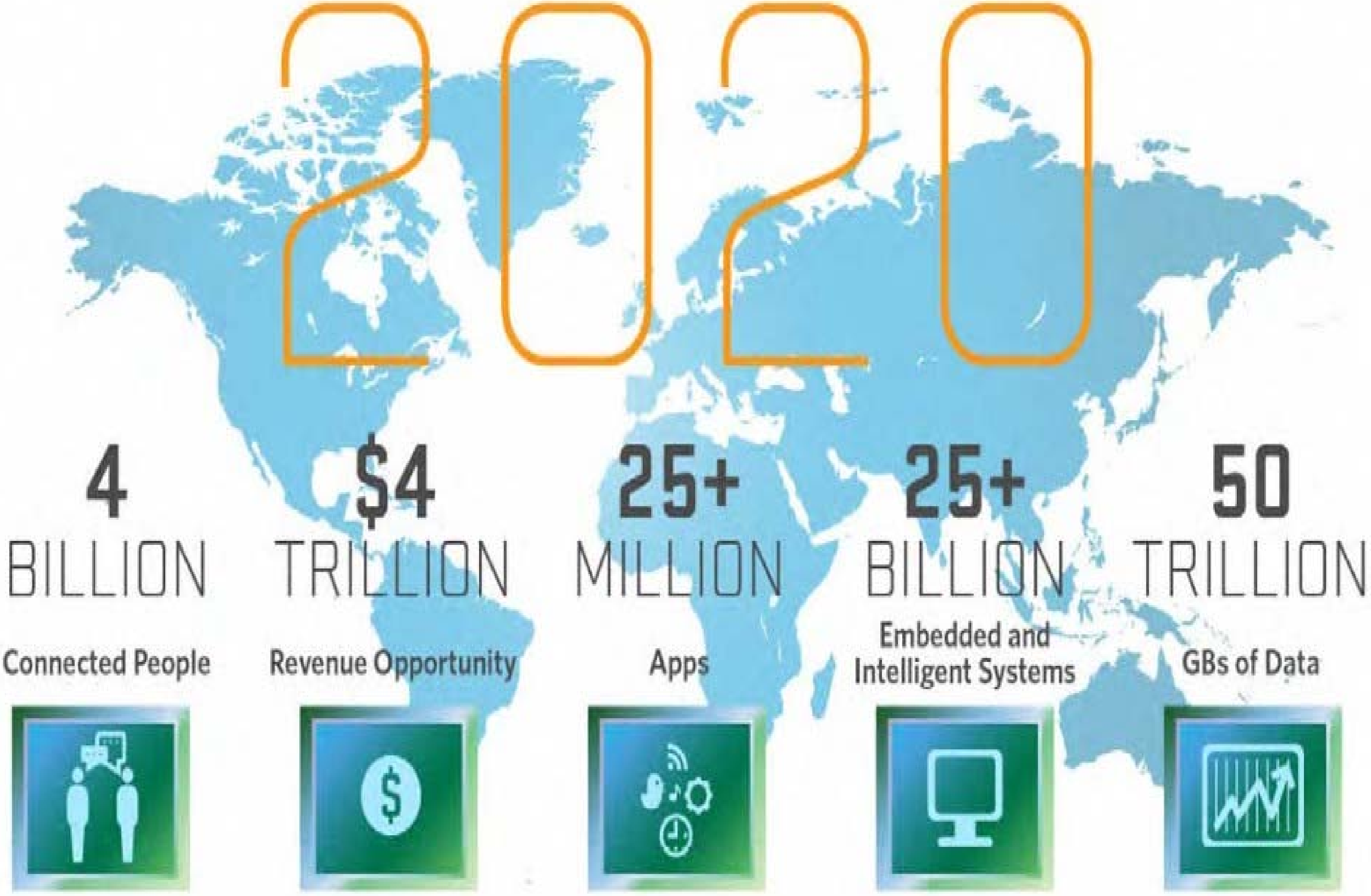
Data Analytics Conference, Nov. 15, 2017
Barcelona

Challenges

- The security is a continuous activity
  - Dynamically evolving according to the changes in architecture and service, and point of view of system owners and users.
- Systems, systems of systems has high heterogeneity; old and new Hw and Sw entities are integrated,
  - IoT, Industry 4.0, critical infrastructures etc.
- High connectivity is exist among entities, they are accessible from every where and every time.
- Current security problems are stochastic.
  - The vulnerabilities and attack types can have many unpredictable combinations.

Challenges

- New methods/tools coming with their own new vulnerabilities and risks?
  - **For instance**; AI is used to distinguish human behaviors from bot nets and to detect bots.
  - The attackers also use same AI abilities to create new equipped bots which behave likes human.



2020

4 BILLION

Connected People

$4 TRILLION

Revenue Opportunity

25+ MILLION

Apps

25+ BILLION

Embedded and Intelligent Systems

50 TRILLION

GBs of Data

Source: Mario Morales, IDC

Required Specs for the Solution

- Solution should match the nature of the security problems,
- It should be deployable and feasible for all components, either hardware or software.
- It should have capable of adapting the strategy to new kinds of threats/attacks and,
- It should generate solutions dynamically.

Required Specs for the Solution

- The responses of the security approach should be monitored and controlled by related entities.
- Under this solution:
  - The collected information should be analyzed to identify new vulnerabilities and attacks to improve the security level.
- This critical information collection and exchange;
  - should be organized and managed using **secure and efficient information sharing models**.

Answer can be a Decentralized Autonomous Solutions ?

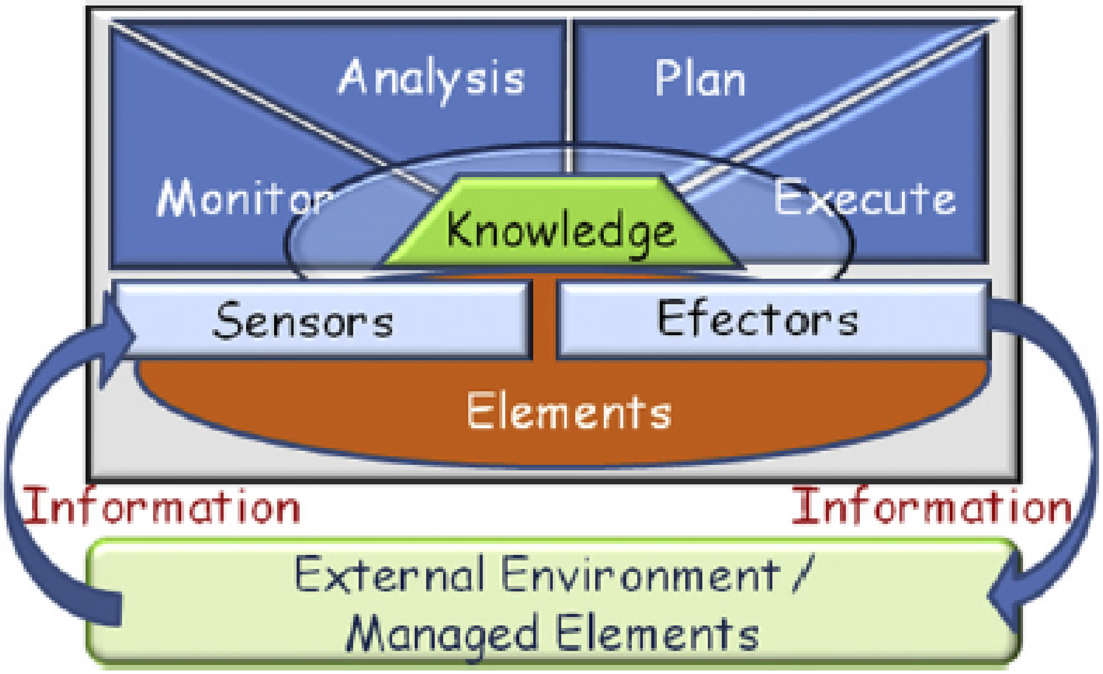
- To solve this problem with central solutions is too hard and not work properly.
- If we think from a decentralized perspective;
  - try to solve it within local scale boundaries
  - with intelligent, communicated and self awareness autonomous entities.

Q. What are the requirements and, strong and weak parts to realize this distributed autonomous solution?

Autonomic Application

- Which is a collection of autonomic elements, which implement intelligent control loops to monitor, analyze, plan and execute actions, using knowledge of the environment by hardware and software entities.

Autonomous Element.



Autonomic Application

- Detecting security problems in local Hw/Sw entities is similar to the behavior of biological systems **(Hariri and Parashar, 2005)**

Characteristics of Autonomic Apps and Systems

1. **Self Awareness**: It “knows itself” and is aware of its state and its behaviors.

2. **Self Configuring**: configure and reconfigure itself under varying and unpredictable conditions.

3**. Self Optimizing**: able to detect suboptimal behaviors and optimize itself to improve its execution.

4. **Self-Healing**: able to detect and recover from potential problems and continue to function smoothly.

Characteristics of Autonomic Apps and Systems

5. **Self Protecting**: capable of detecting and protecting its resources from both internal and external attacks and maintaining overall system security and integrity.
6**. Context Awareness**: be aware of its execution environment and be able to react to changes in it.
7. **Open**: It must function in a heterogeneous world and should be portable across multiple hardware and software architectures. Consequently it must be built on standard and open protocols and interfaces.
8. **Anticipatory**: be able to anticipate to the most possible extent, its needs and behaviors and those of its context, and be able to manage itself proactively

Current Situation (-)

- Data is big and dynamically changing.
- Data analysis techniques and tools has crucial role to analyze it, but have many obstacles .
  - **For instance**: Data representation structures and techniques are so important to reduce algorithmic complexities and process dynamic and big data flows.
  - Local data processing and analyzing; the importance of in-memory operations is increasing
- Identification of each entity requires new solutions

Current Situation +

- Communication infrastructure is more and more robust
- Processing power is increasing in each type of single and tiny entity.
- Huge server farms and data centers are exist.
- All kind of processing units are connected and accessible among them.
- Algorithms and software tools; with AI, ML etc., all these give new abilities to extract hidden knowledge and take decisions very fast and locally.

Proposed Solution Architecture

- Consists of **autonomic elements**, each performing a fixed function and interacting with other elements, possibly in very dynamic environments.
- An autonomic element;
  - Comprised of one or more managed elements **(functional units),**
    - each performing its operational function, with one **autonomic manager** (management unit
      - that controls the managed elements' configuration, inputs, and outputs.

Proposed Solution Architecture

- The Hw or Sw autonomous entities are able
  - to recognize the security problems (self-healing, -protection),
  - sharing information with other autonomic components
  - (context awareness), for then selecting the more appropriate reaction behavior and
  - implementing the necessary changes (self-optimizing and configuring) for the whole system.

Proposed Solution Architecture

- The self-adaptive applications should
  - monitor and organize the global reaction, such as the immune system of a living organism.
- In a self-adaptive system and/or network, services are
  - able to recognize the security problems,
  - sharing information with other autonomic components,
  - for then selecting the more appropriate reaction behavior and implementing the necessary changes.

Proposed Solution Architecture

- Those software systems must be informed by a trust model
  - **which resources are to be trusted.**
- In addition, system must be capable of detecting its own malfunctioning, diagnose the respective failure, and consequently repair itself.
  - **For example**, a system might notice through self-monitoring that it is running much slower than expected (Shrobe, November 4, 2002).

Proposed Solution Architecture

- The results of the technical report from MIT showed that (Shrobe et al., April 10, 2007) :
  - self-awareness and self-adaptivity can be successfully applied to monitoring the behavior of systems,
  - diagnose failures, and
  - adapt and recover from both insider and external attackers.

Conclusion

- I believe that we can develop local self-awareness entities with local immune or reaction systems for each created entity in virtual world.
- The questions are
  - "how is it possible for a simple entity",
  - then for "systems" and
  - "systems of systems" etc..

Conclusion

- The solution is dependable for technological capabilities, communication infrastructure and design success.
- Current processing and communication capacity move us to process much more data.
- AI and its new generation facilities reach us to extract many hidden information from any kind of data.
- Why we do not combine all of these for the cyber health of Industrial Systems?"

Thank you for your attention

References

Hariri S, Parashar M. Handbook of bioinspired algorithms and applications, chapter the foundations of autonomic computing. CRC Press LLC; 2005.

Shrobe H, Laddaga R, Balzer R, Goldman N, Wile D, Tallis M, et al. 'Self-adaptive systems for information survivability: PMOP and AWDRAT'. MIT-CSAIL-TR-2007-023. MIT Cambridge: Computer Science and Artificial intelligent laboratory Technical Report, www.csail.mit.edu; April 10, 2007.

Shrobe H. 'Computational vulnerability analysis for information Survivability'. AI Magazine November 4, 2002;23.

ATAY, S., MASERA, M., "Challenges for the Security of Next Generation Networks", *Broadnets 2009, Madrid, Spain, 14-17 September 2009*, DOI:10.4108/ICST.BROADNETS2009.7470, *revised in 2010 and* ***published at Journal of Elsevier*** ***"*****Information Security Technical Report (2010)", doi:10.1016/j.istr.2010.10.010** (*This work was supported in part by the TUBITAK - Turkish Science and Technological Research Program- under Grant of Bideb2219 Post-Doctorate Research Fellowship.)* **Link: http://www.sciencedirect.com/science/article/pii/S136341271000035X**

DEEP LEARNING AND BIG DATA IN CYBERSECURITY

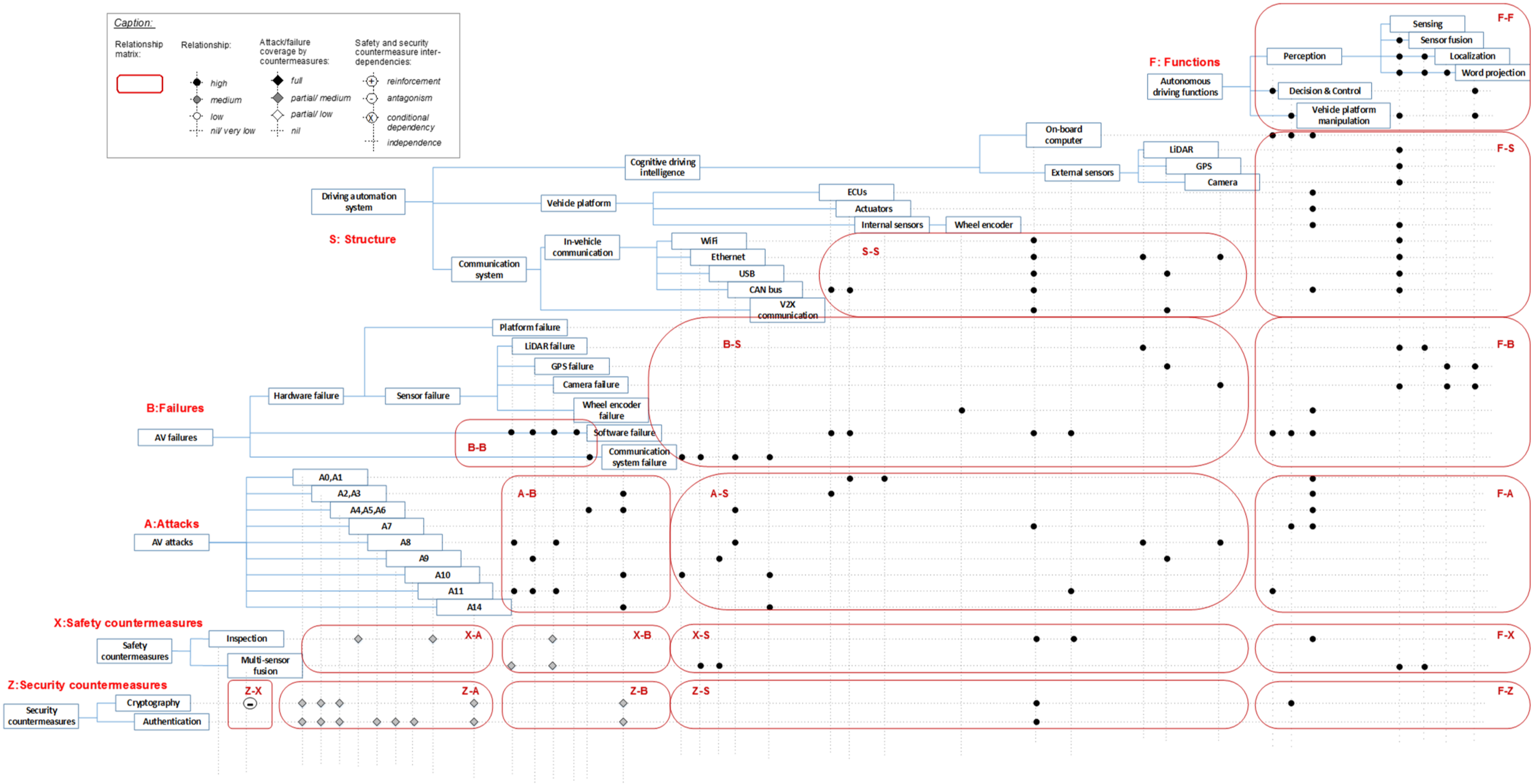
Jin Cui
Singapore University of
Technology and Design
Singapore

Data analysis for Autonomous vehicles

Jin CUI (SUTD)

Security for Autonomous vehicle

- Goal to guarantee the safety and cyber security of AVs.
- SSM model to analysis safety and security in parallel, thus, there will be different data from structure, functions, failure, attack, and countermeasures.
- How to effectively analyse this data?


Caption:
Relationship matrix:
Relationship:
high
medium
low
nil/ very low
Attack/failure coverage by countermeasures:
full
partial/ medium
partial/ low
nil
Safety and security countermeasure inter-dependencies:
reinforcement
antagonism
conditional dependency
independence
F: Functions
Autonomous driving functions
Perception
Sensing
Sensor fusion
Localization
Word projection
Decision & Control
Vehicle platform manipulation
F-F
F-S
F-B
F-A
F-X
F-Z
S: Structure
Driving automation system
Cognitive driving intelligence
On-board computer
External sensors
LiDAR
GPS
Camera
Vehicle platform
ECUs
Actuators
Internal sensors
Wheel encoder
Communication system
In-vehicle communication
WiFi
Ethernet
USB
CAN bus
V2X communication
S-S
B:Failures
AV failures
Hardware failure
Platform failure
Sensor failure
LiDAR failure
GPS failure
Camera failure
Wheel encoder failure
Software failure
Communication system failure
B-B
B-S
A:Attacks
AV attacks
A0,A1
A2,A3
A4,A5,A6
A7
A8
A9
A10
A11
A14
A-B
A-S
X:Safety countermeasures
Safety countermeasures
Inspection
Multi-sensor fusion
X-A
X-B
X-S
Z:Security countermeasures
Security countermeasures
Cryptography
Authentication
Z-X
Z-A
Z-B
Z-S

Challenges

- Relationship definition? High, median or low?
- Useful information?
- Implementation?

DEEP LEARNING AND BIG DATA IN CYBERSECURITY

Dimitris Kardaras
Athens University of
Economics and Business
Greece

Deep Learning in Business: Applications and Challenges

Dimitris K. Kardaras
Athens University of Economics and Business,
Athens, Greece

DL application areas in Business I

- **Recommender Systems** (high conversion rate approx. 60% ofAmazon's sales)

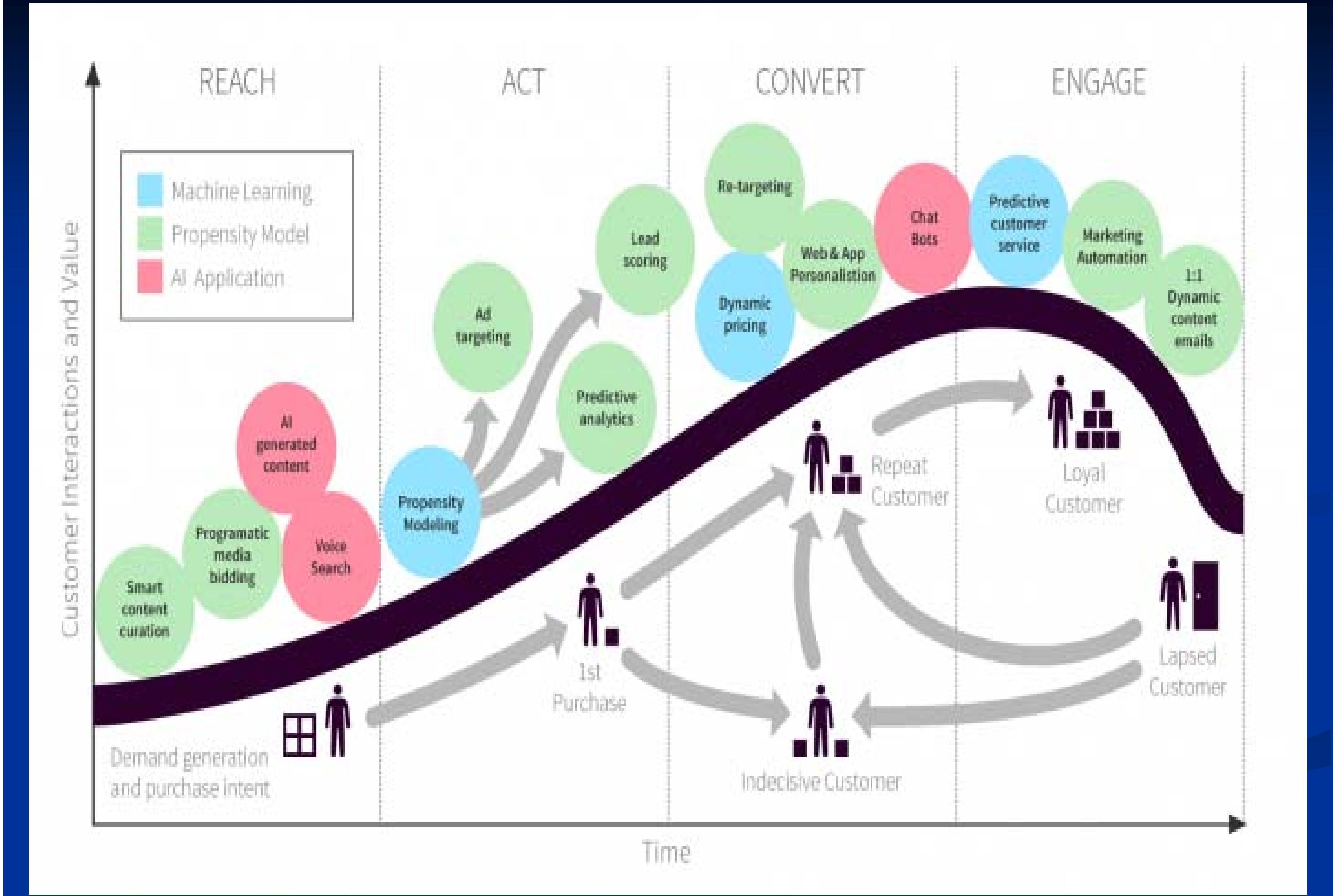
- **Semantic Analysis**; reviews analysis, building customer profiles and services/products models (customer before commit to purchase they read on average 4 reviews; important to know what they have read)

DL application areas in Business II

- **Propensity Modelling**
- knowing what, when, and why your customers are going to buy;
- predicting the likelihood of a customer to convert,
- predicting what price a customer is likely to convert at, or what customers are most likely to make repeat purchases;
- predictive customer service; what is the next step for the customer?

DL application areas in Business III

- **Personalised Communication**
- Chatbots, Personal Assistants (Mya recruiter);
- Ads Targeting (important both for agents and customers);
- Content Generation.


REACH
ACT
CONVERT
ENGAGE
🟦 Machine Learning
🟩 Propensity Model
🟥 AI Application
Smart content curation
Programatic media bidding
AI generated content
Voice Search
Propensity Modeling
Ad targeting
Predictive analytics
Lead scoring
Re-targeting
Dynamic pricing
Web & App Personalistion
Chat Bots
Predictive customer service
Marketing Automation
1:1 Dynamic content emails
Demand generation and purchase intent
1st Purchase
Repeat Customer
Indecisive Customer
Loyal Customer
Lapsed Customer
Customer Interactions and Value
Time

DL in Business: Challenges

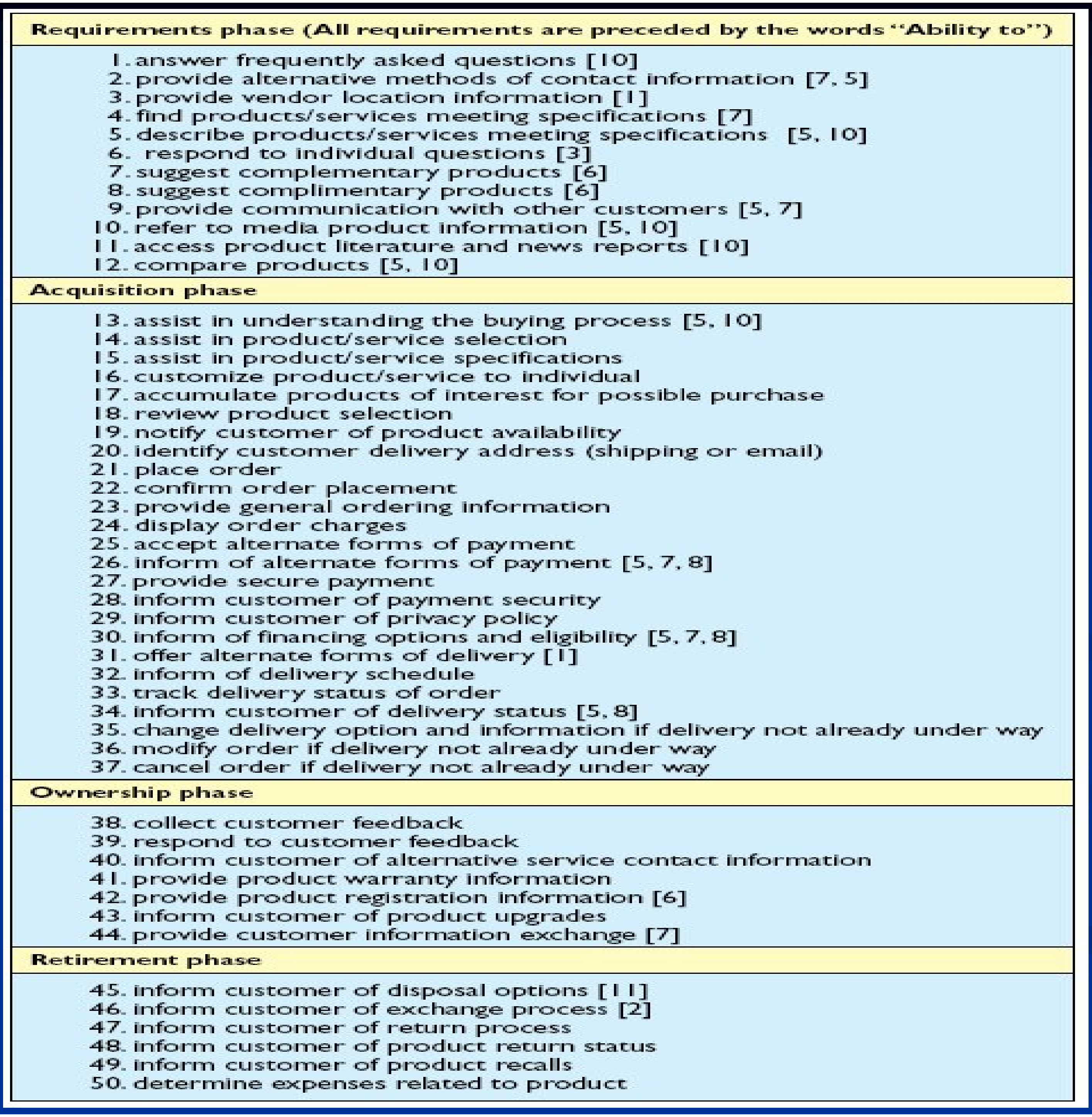
- Lack of data; Cold start;
- Customer are human beings…do not always follow rules, as they are constantly exposed to messages and information from competitors they may change their priorities for reasons…not always obvious… at least… to the algorithms!!!
- Legislation for data protection may restrict access and use of data

DL in Business: Suggestions?

- Use of Fuzzy Logic to accommodate large data sets, more human like segmentation and personalisation.

- Multi-disciplinary approaches, e.g. Use of frameworks such as the Customer Journey and Customer Service Life Cycle Model, to assist in modelling-predicting the content a customer may require

The Customer Service Life Cycle



Requirements phase (All requirements are preceded by the words "Ability to")

1. answer frequently asked questions [10]
2. provide alternative methods of contact information [7, 5]
3. provide vendor location information [1]
4. find products/services meeting specifications [7]
5. describe products/services meeting specifications [5, 10]
6. respond to individual questions [3]
7. suggest complementary products [6]
8. suggest complimentary products [6]
9. provide communication with other customers [5, 7]
10. refer to media product information [5, 10]
11. access product literature and news reports [10]
12. compare products [5, 10]

Acquisition phase

13. assist in understanding the buying process [5, 10]
14. assist in product/service selection
15. assist in product/service specifications
16. customize product/service to individual
17. accumulate products of interest for possible purchase
18. review product selection
19. notify customer of product availability
20. identify customer delivery address (shipping or email)
21. place order
22. confirm order placement
23. provide general ordering information
24. display order charges
25. accept alternate forms of payment
26. inform of alternate forms of payment [5, 7, 8]
27. provide secure payment
28. inform customer of payment security
29. inform customer of privacy policy
30. inform of financing options and eligibility [5, 7, 8]
31. offer alternate forms of delivery [1]
32. inform of delivery schedule
33. track delivery status of order
34. inform customer of delivery status [5, 8]
35. change delivery option and information if delivery not already under way
36. modify order if delivery not already under way
37. cancel order if delivery not already under way

Ownership phase

38. collect customer feedback
39. respond to customer feedback
40. inform customer of alternative service contact information
41. provide product warranty information
42. provide product registration information [6]
43. inform customer of product upgrades
44. provide customer information exchange [7]

Retirement phase

45. inform customer of disposal options [11]
46. inform customer of exchange process [2]
47. inform customer of return process
48. inform customer of product return status
49. inform customer of product recalls
50. determine expenses related to product

Thank you!!!