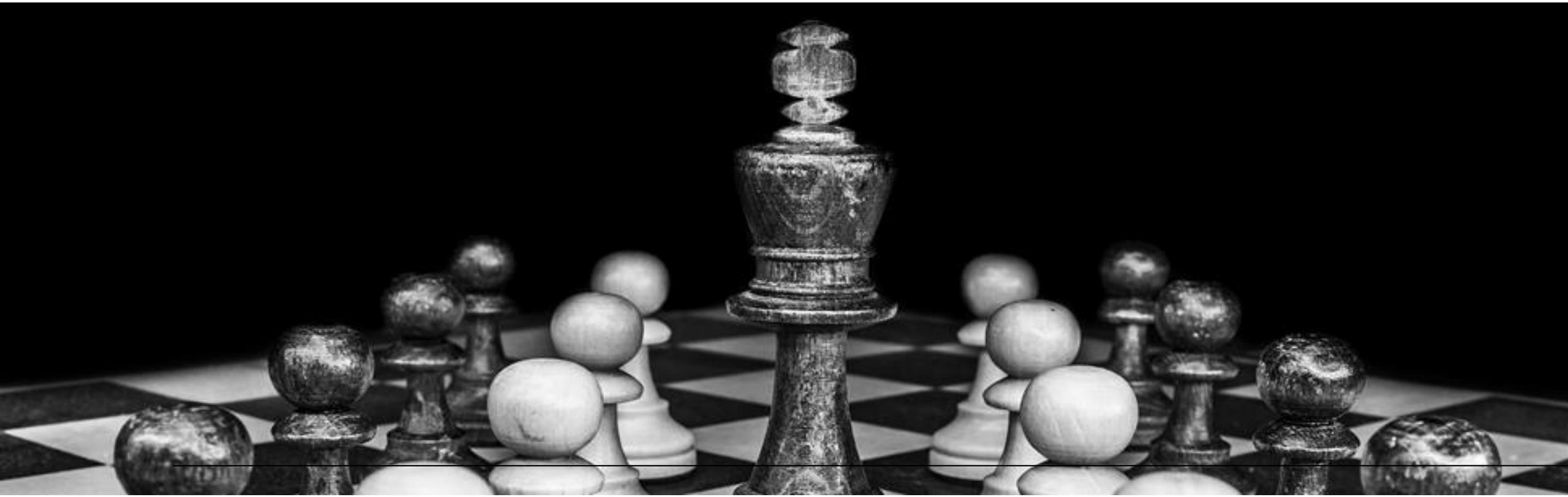
Grow with AI

Ulrich Bodenhausen AI Coaching



Deep Learning and Functional Safety

ESE Kongress 2018, Dr. Ulrich Bodenhausen, AI Coach and Vector Consulting Services

V1.0 | 2018-06-12

About me

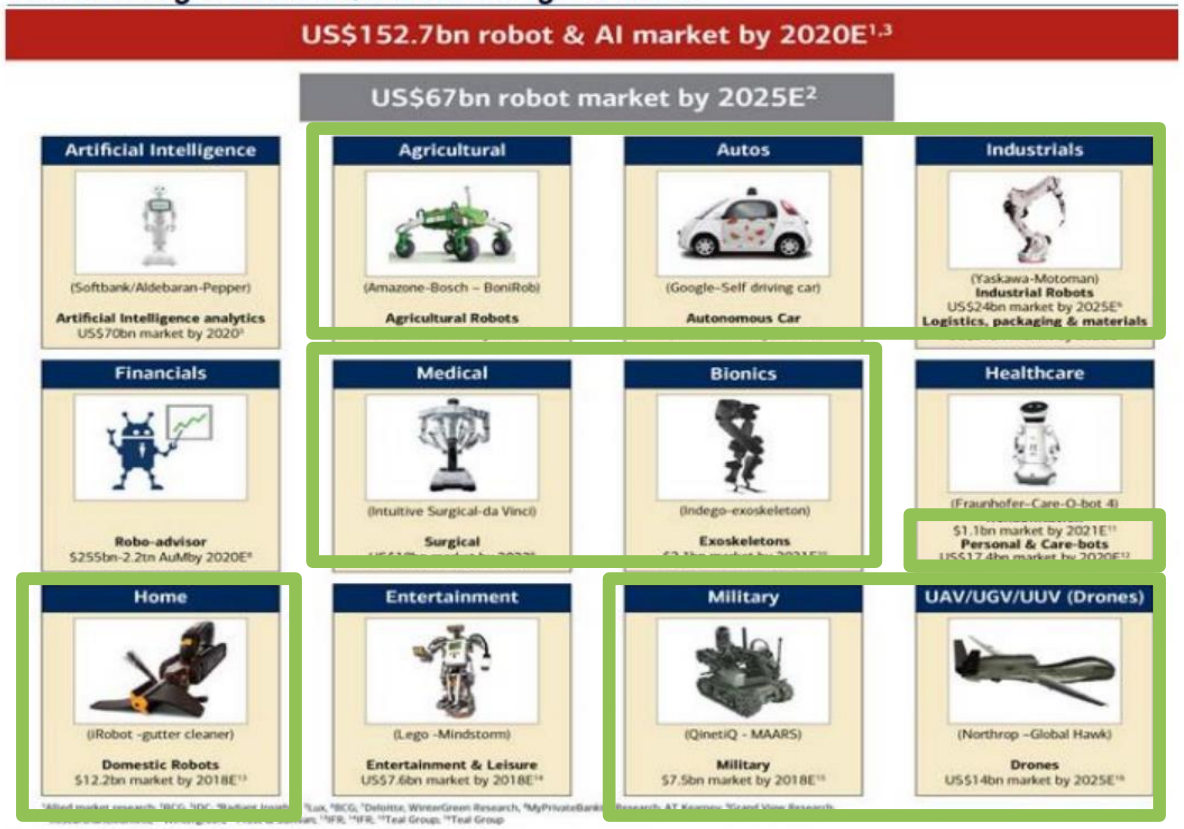
- More than 25 years ago I was excited to see that some formulas could result in an learning algorithm that was able to recognize **speech or gestures**.
- **PhD** in Machine Learning from KIT.
- Application of neural networks to speech and gesture recognition at **Carnegie Mellon University and KIT**.
- Consulting using applied **predictive analytics**.
- Strategic **business planning** at a 1st tier automotive supplier.
- Responsibility for **innovation management**, development processes and knowledge management at 1st tier automotive supplier.
- Managing a team of consultants **for transition programs** and optimization of product development (e.g. agile methods, development methods, functional safety) in the automotive industry at Vector Consulting Services.

AI Market and Functional Safety

Market analysis reports predict significant **growth of revenue with AI**. Selected quotes:

- Tractica Research: Growth rate (CAGR) of **56.8%.**
- BCC Research expects autonomous robot category to account for the largest share - **22.8%** - of annual market growth until 2024 and thus to dominate the smart machine market.
- GII Research (based in Japan): Artificial intelligence market is estimated to have cumulated growth rate of **53.65%** from 2015 to 2020.
- Bank of America Merrill Lynch: **35%** share of robots and artificial intelligence market for safety critical products.

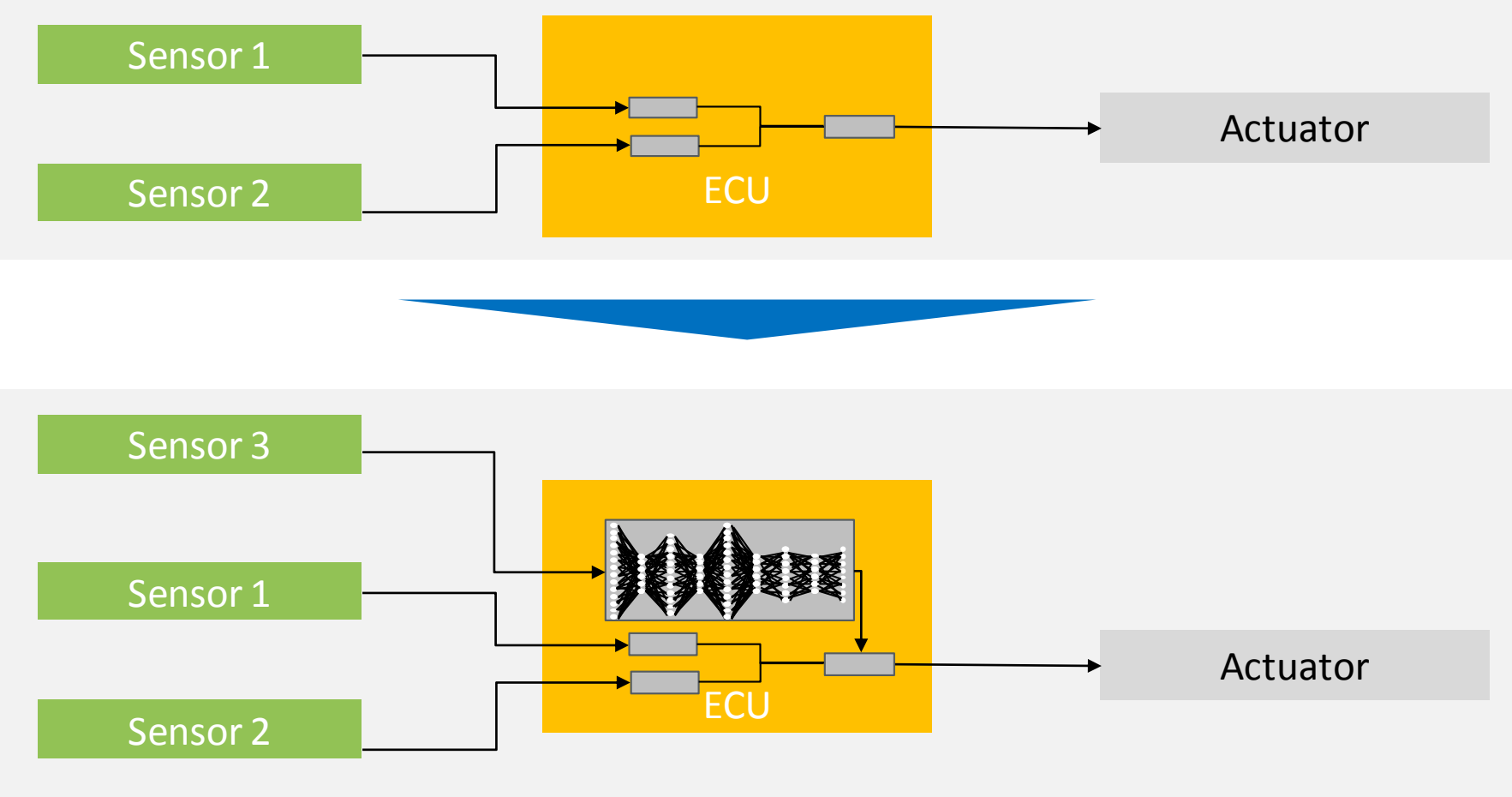
Exhibit 1: The global robots & artificial intelligence market



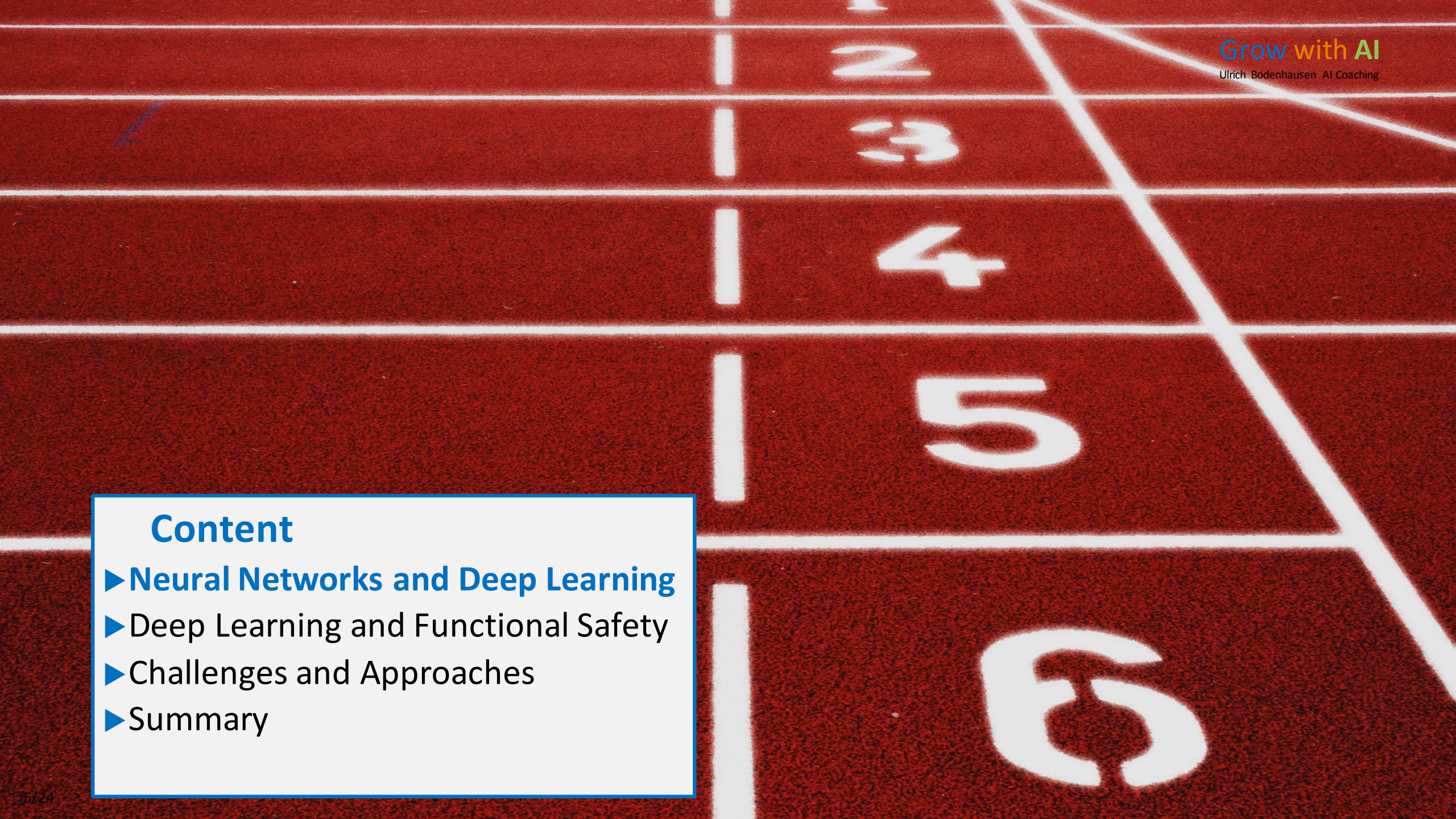
Source: BofA Merrill Lynch Global Research

1. AI is likely to become very important driver for safety critical products
2. Functional safety is an important requirement for large portion of AI based products

Potential of AI in Safety Critical Systems: Augmentation of ECUs with Additional Sensors and AI



Autonomous cars are large and important AI market share.
There are other attractive safety critical AI applications.



Content

- **Neural Networks and Deep Learning**
- Deep Learning and Functional Safety
- Challenges and Approaches
- Summary

Machine Learning and Goals for Real-World Applications

Machine Learning (ML):

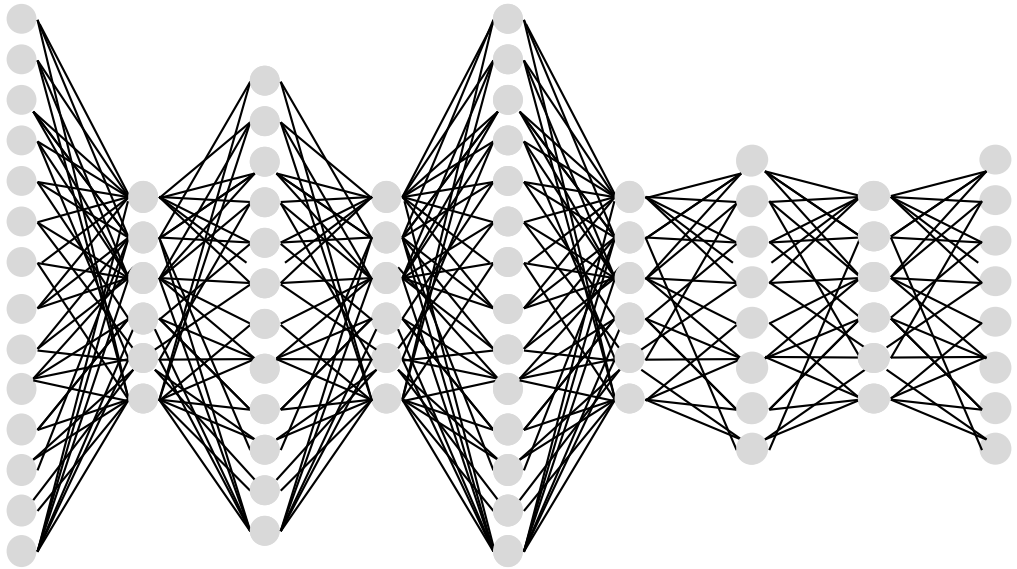
- Approach using inductive learning techniques for system design
- The run-time system uses the results of a learning process to perform algorithmic operations (e.g. running a neural network having precomputed weights).
- Assumption: Weights are fixed before validation by test data (e.g. no dynamic adaptation during runtime).

Goals for ML in Real-World Applications:

- The run-time systems performs with low deviation from desired system functionality in the defined application
- The functionality is robust against variances (i.e. rotation, size, shape, color, …)
- The performance is uniform across classes of inputs (i.e. critical classes do not yield unacceptable performance)
- The system is robust against differences between its training and testing sets

Neural Networks and Deep Learning

Generic (Deep Learning) Neural Network (NN)



Input

Output

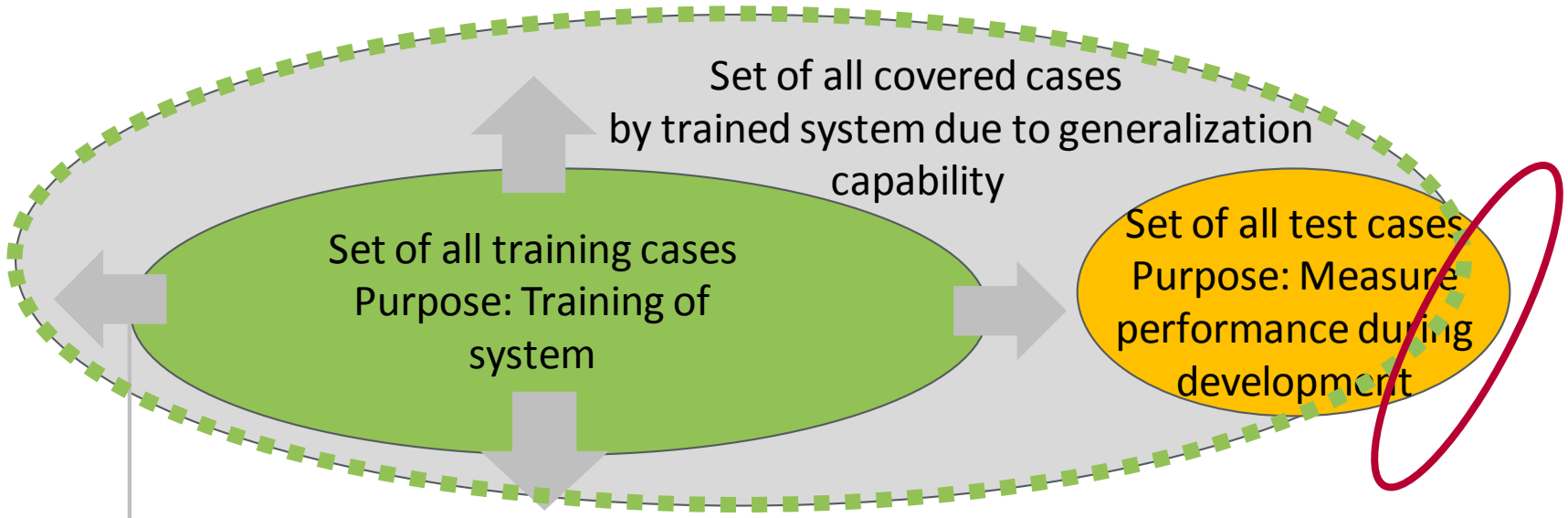
- Each node connected to each node of next layer
- Interconnection “weights” are computed by iterative method:
  - Using training examples
  - Gradient descent method on an energy function measuring the deviation to the expected output.
- Deep Learning = Learning the interconnections for several hidden layers

Deep Learning NN have proven to be very powerful approaches to Machine Learning in real-world applications. Why?

Generalization Capability

Training and testing set must be completely separate sets of data to assure evaluation on real-world data

Generalization Capability Enables Coverage of Testing Data



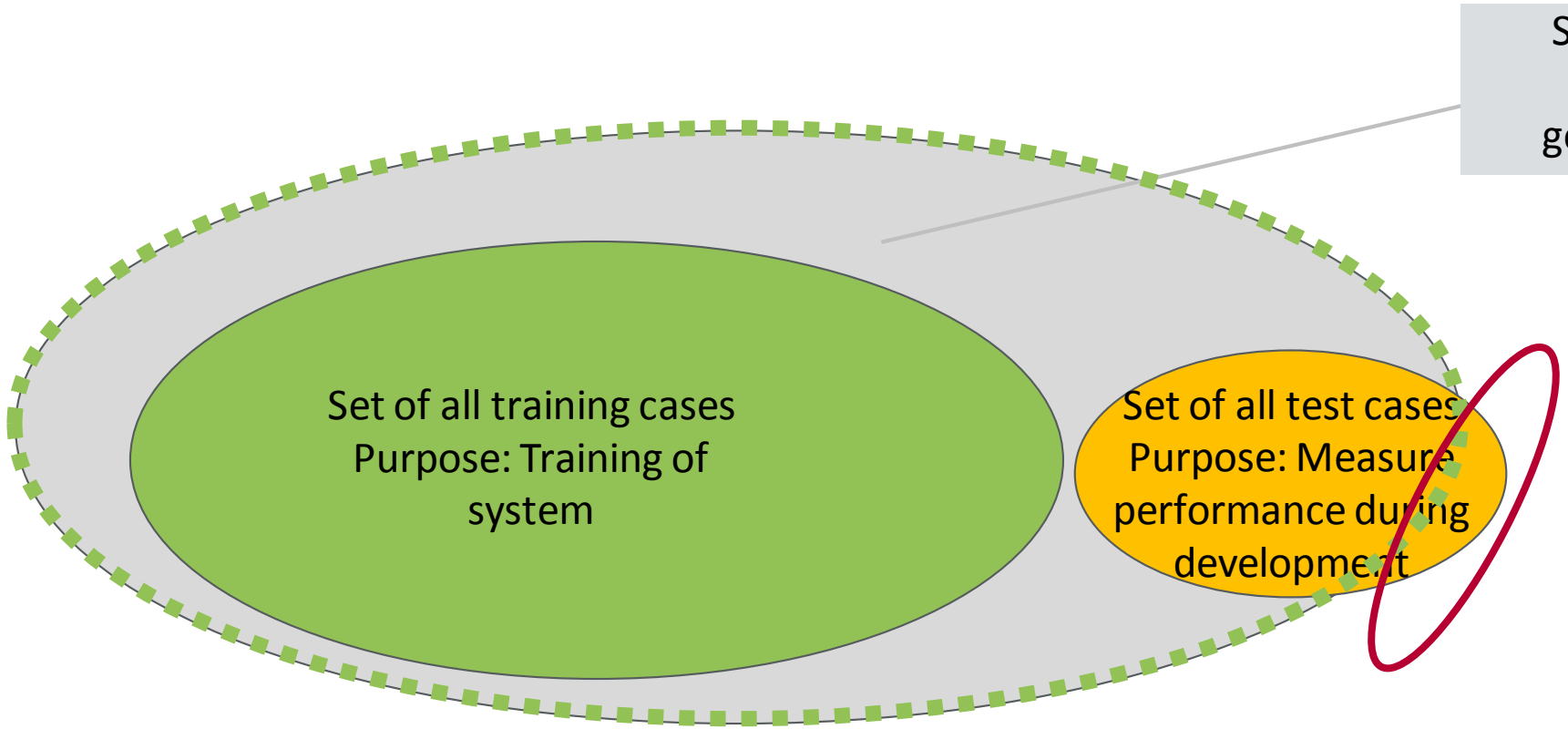
Approaches to enhance generalization capability of trained system beyond set of training cases

Not all test cases may be covered

Generalization capability: The trained system covers more cases than contained in the training data set due to the capability to extract features instead of memorization

THERE IS NO DATA LIKE MORE DATA!

Increasing Training Data Set without Enforcement of Generalization Capability



Still not all test cases may be covered

Experience: Increasing training data set improves generalization capability but still may not result in optimum performance on test data

Recent Results of Joint Research by Google and MIT

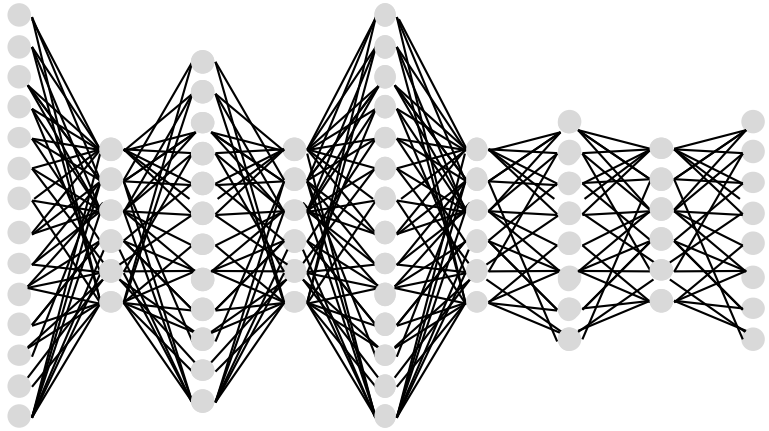
- Many NN are able to learn training data set, but generalization capability is very different.
- Under certain conditions Deep Learning NN have the proven ability to develop strong generalization capability.
- Conditions have recently been analyzed in detail* with the result that NN can generalize surprisingly well (better than expected).
- Conditions need to be analyzed further, but what ca be stated already is:
  - Changes in architecture resulted in bigger improvements in generalization capability than other approaches, i.e. regularization techniques.

Changes of architecture can be done manually (with very high effort from experts) or by automatic optimization algorithms.

* Source: Zhang, C., Bengio, S., Hardt, M., Recht, B., Vinyals, O., „Understanding Deep Learning Requires Rethinking Generalization“, ICLR 2017, arXiv:1611.03530 [cs.LG]

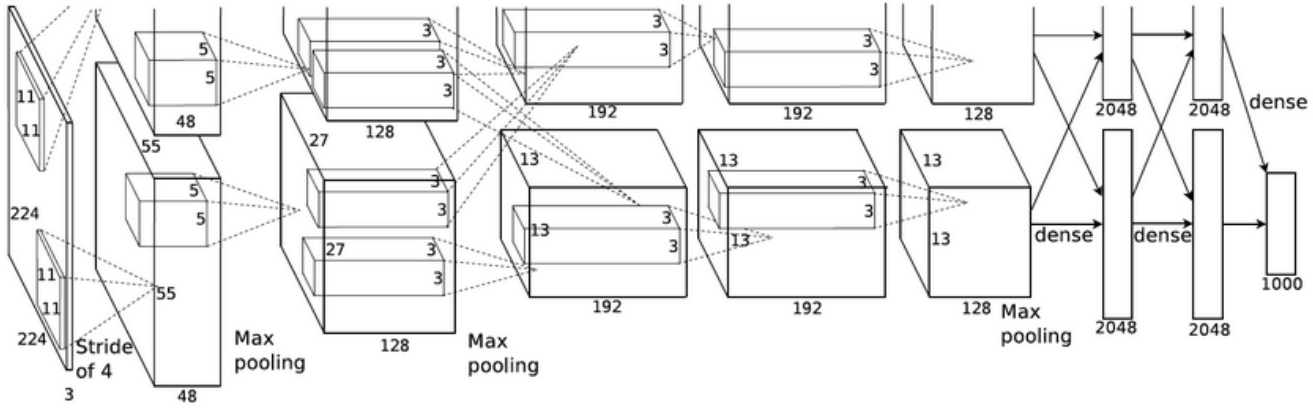
Architecture is Important: Tailored Architecture Massively Improves Performance

Generic Deep Learning Neural Network



- Each node connected to each node of next layer
- Generic Neural Network with 3 hidden layer achieves approx**.**
  50% test accuracy
  on difficult image classification task (CIFAR-10)

Example AlexNet*: Taylored to enforce feature extraction



- 5 convolutional layers and 3 fully connected layers
- **> 80% test accuracy**
  on CIFAR-10 (2012 results)

* Example; Best in class applications of rely on even more tuning of architectures

* Source: Krizhevsky et al: ImageNet Classification with Deep Convolutional Neural Networks, 2012

Content

- Neural Networks and Deep Learning
- **Deep Learning and Functional Safety**
- Challenges and Approaches
- Summary

Machine Learning/Deep Learning and Functional Safety Standards

Published Standards

- IEC61508 explicitly mentions artificial intelligence as a not to be proposed technic/measure (IEC61508 table A2-5).
- ISO26262-1: Very detailed methodology, but not well adapted to Machine Learning/Deep Learning. Examples:
  - Development and verification methods within part 6 do not address:
    - Functionality is embedded in highly dimensional data matrices
    - Verification methods such as coding guideline, white-box code coverage provide no relevant insights
    - Functionality is opaque to humans
    - Insufficiency of the system due to inherent restrictions in sensors, actuators

Standards in preparation and current practices

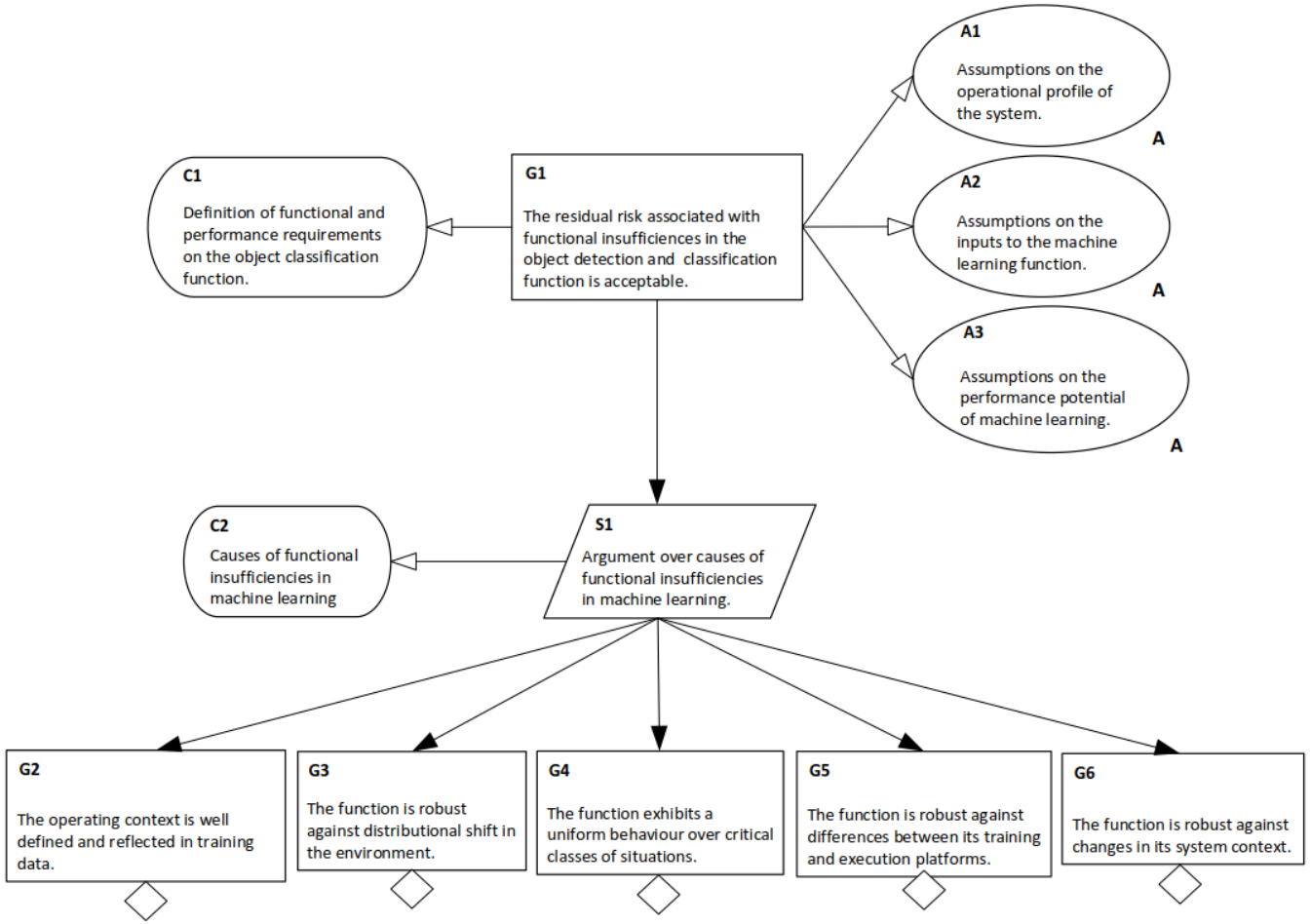
- ISO26262-2: under preparation for its final publication.
- ISO/WD PAS 21448 Safety of the Intended Functionality (SOTIF): under publication, but not published yet.
- Practice for Highly Automated Vehicles (HAV) as spearhead of ML technology:
  - Extensive road test – many Mio. Miles
  - Proprietary approaches, considered to be sophisticated, but not open for public evaluation.

Publishes standards do not yet provide guidance on how to proof safety for AI based products.

Concepts for Highly Automated Vehicles (HAV): Assurance Case for HAV Containing ML

One important Goal of the assurance case to be met:

- Acceptable residual risk of functional insufficiencies, proven by sub-goals:
  - Operational context is well defined and reflected in training data
  - Function is robust against distributional shift
  - Function exhibits uniform behavior over critical classes
  - Function is robust against differences between training and execution platform
  - Function is robust against changes in system context



Sub-Goals are in line with design goals and strengths of Deep Learning NN.
Now the what and how need to be addressed.

Source: Making the Case for Safety of machine Learning in Highly Automated Driving, S. Burton, L. Gauerhof, C. Heinzemann, SAFECOMP 2017

Concepts for HAV: Koopman

Koopman and colleagues propose safety validation approach going beyond the brute force on-road testing campaigns. Elements:

- Phased simulation and testing approach, emphasize testing to mitigate residual risk from previous phase
- Observability points to produce human- interpretable data; demonstrate the system is doing the right thing for the right reason
- Differentiation of various roles of testing form checking of requirements gaps to checking of design faults
- Run-time monitoring approach to manage identified risks

Need to go beyond "Deep Learning NN is a black box" safety case:
Put in observability points and validate safety for these points under given assumptions

Source: Toward a Framework for Highly Automated Vehicle Safety Validation, P. Koopman, M. Wagner, SAE World Congress 2018

Content

- Neural Networks and Deep Learning
- Deep Learning and Functional Safety
- **Challenges and Approaches**
- Summary

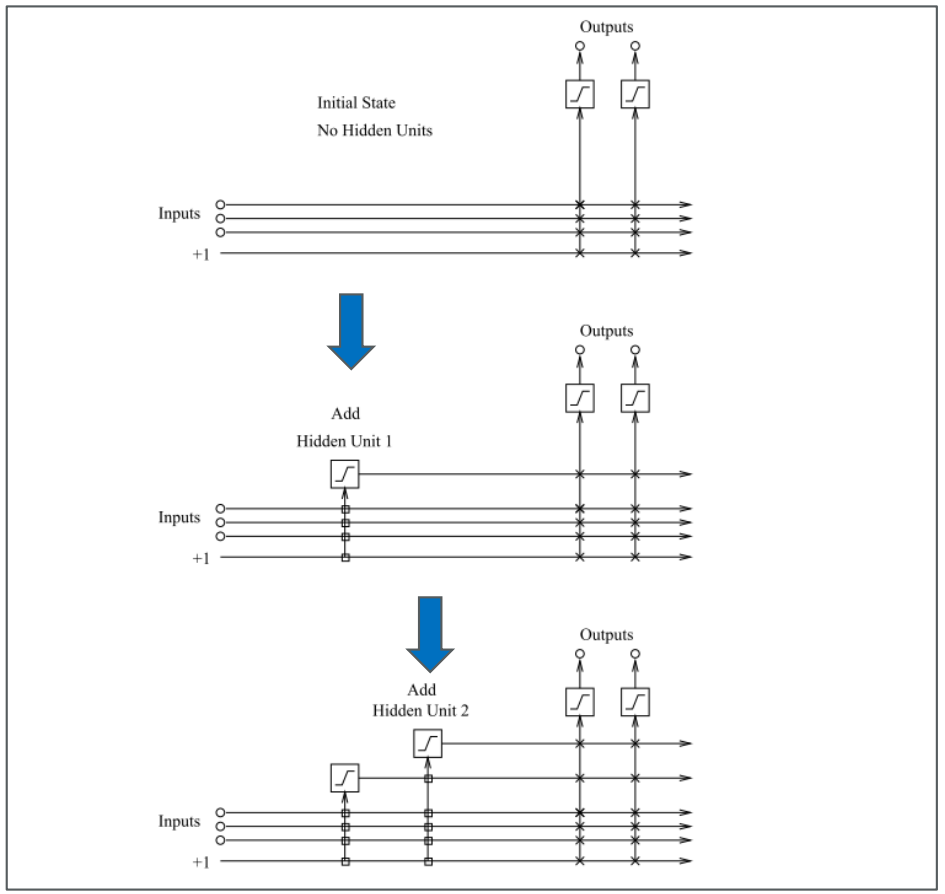
Challenges and Approaches

Challenges:	Common Approaches:	Advanced Approaches:
▪ Acceptable residual risk of functional insufficiencies	▪ Increase amount of training data ▪ Increase amount of testing data ▪ Enforce generalization capabilities: Optimize architecture and learning parameters to optimize generalization capability	▪ Automatic architecture optimization ▪ Observability points and test levels for verification and validation ▪ Use decomposition scheme from ISO 26262: Provide the functionality of one system with high ASIL level by two subsystems with lower ASIL level

▪ Internal structure is inherently hard to analyze, especially in Deep Learning NN layers with full connectivity	▪ -	▪ Enforce structure in architecture that can be analyzed. i.e.
		  ▪ Enforce tailored connectivity
		  ▪ Insert observability points
		  ▪ Explicitly test the system is doing the right thing for the right reason

Approaches: Automatic Optimization of Architectures

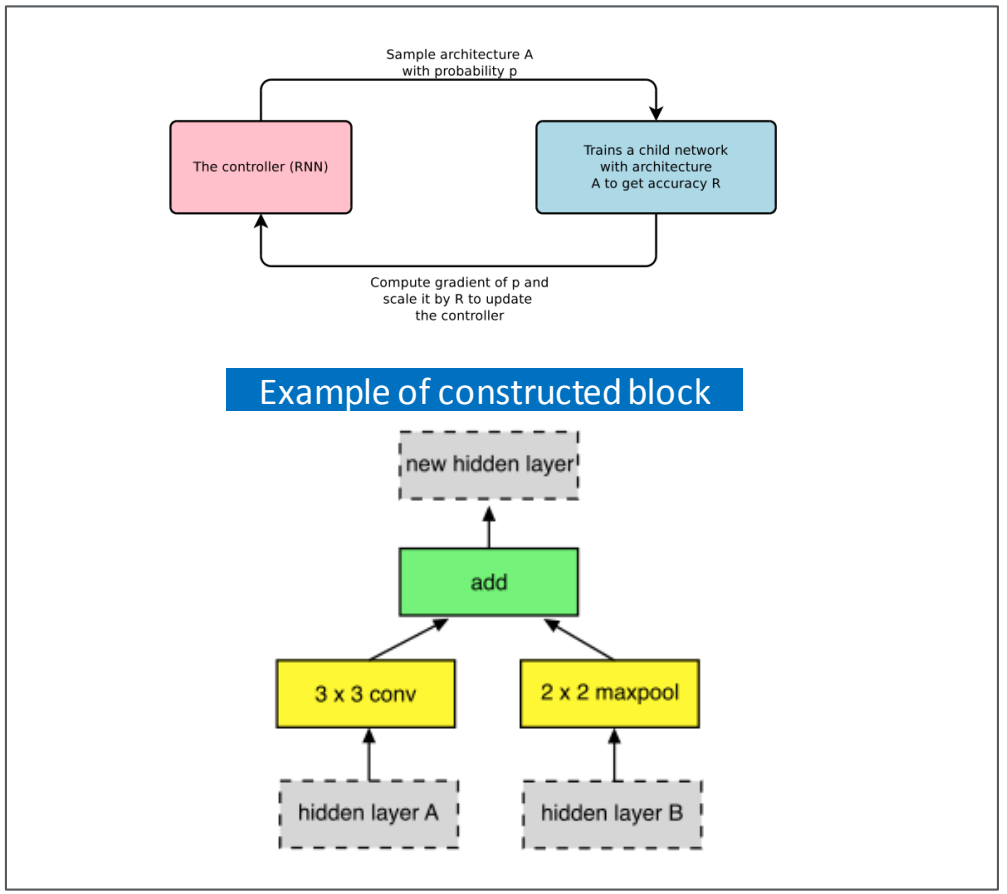
Cascade Correlation Algorithm, 1991*



Approach: Enforce units and connectivity to solve dedicated classification problems

More than 25 years of research

Neural Architecture Search NASNet, 2017**



Approach: Enforce building blocks to solve dedicated classification problems

* Fahlman, S., The Cascade Correlation Algorithm, 1991
** Zoph, B., "Learning Transferable Architectures for Scalable Image Recognition", arXiv:1707.07012v4 [cs.CV] 11 Apr 2018

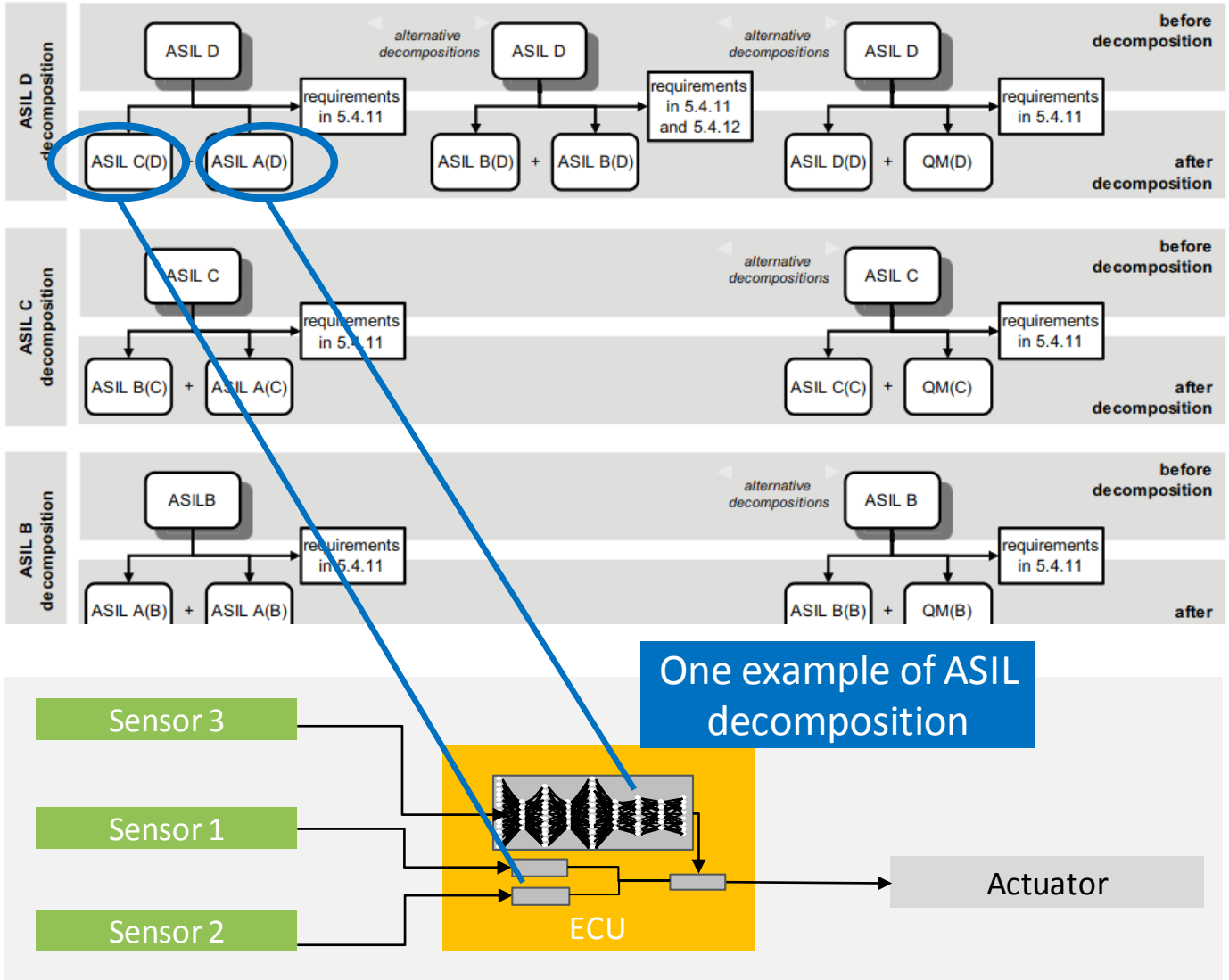
Approaches: Confusion Matrix for Automatic Optimization of Architecture

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	74	1			2					6															1	
B	1	65		2							2						1				1	2				
C			74																3							3
D		3		80		1	2													2			2			1
E	1				73					1								1			4					
F					1	77	1						1	1		1			2	1		3	1	1		
G			1				78	19									1									
H	3	1			3			63		1							1			1		1				
I	1								73					2	3			3								
J							1			76	1				1											1
K		1									79		2			1									1	1
L			1		1				1			79			2											
M		2			1				1				74	37	2			3							1	1
N				1	1								4	44	1											
O	2								6			3	1		74			1							2	
P		2		1		1		1			1					81	2	1					1	1		
Q			2				2	2			2						78			1				2		
R									3				2	1	1			74				1			2	
S			3			1	1					1							78				1			
T			1														1			80				1	
U					2																79	1				
V	3	9			1	4							1				1	1			1	75	3	1	4	
W		1		1		1										2						1	76			
X																						1		77		1
Y									1	1					1			1							74	
Z			3																2				1			79

Source: Bodenhausen, Automatic Structuring of Neural Networks …, PhD Thesis 1994

- Algorithms from Google for automatic optimization of architectures use confusion matrix to optimize uniform behavior over critical classes.
- Example from 1994 alphabet spelling task: “N” and “M” confused.

Approaches: Decomposition of High ASIL to Two Subsystems with Lower ASIL



Source: ISO 26262-1

- ISO 26262 considers ASIL decomposition as a concept to provide the functionality of one system with high ASIL level by two subsystems with lower ASIL level.
- Condition: Evidence of sufficient independence of the elements.

Content

- Neural Networks and Deep Learning
- Deep Learning and Functional Safety
- Challenges and Approaches
- **Summary**

Summary: 4 Conclusions and Success Factors for Deep Learning and Safety Critical Applications

AI is coming and it will be a very important driver for safety critical products:

- Deep Learning NN have proven to be very powerful approaches in real-world ML applications.
- The current state of not yet published standards will not stop this

Arguing safety will rely on:

- Acceptable residual risk of functional insufficiencies
- Transparency on what is going on in NN

Approaches to complete the argument need:

- Advanced approaches from optimization of Deep Learning architectures
- Intelligent bundling with proven approaches from Safety, like ASIL decomposition.

Machine learning and safety experts need to collaborate closely to make it successful and safe.