Continual Learning: On Machines that can Learn Continually

Official Open-Access Course @ University of Pisa, ContinualAI, AIDA

Lecture 1: Introduction and Motivation

Vincenzo Lomonaco
University of Pisa & ContinualAI
vincenzo.lomonaco@unipi.it

ContinualAI

A Non-profit Research Organization and Open Community on
Continual Learning for **AI**

Home | News & Events | Research | Lab | Forum

Supporters | About us | Join us!

ContinualAI.org

Focus on Deep Neural Networks

Vincenzo Lomonaco (Instructor)

Assistant Professor @ University of Pisa

Co-founding President and *Lab Director @* ContinualAI.org

Co-founder & *Board Member @* AIforPeople.org

Antonio Carta
(Teaching Assistant)

Post-Doc @ University of Pisa

Expert in Distributed Continual Learning

Andrea Cossu (Teaching Assistant)

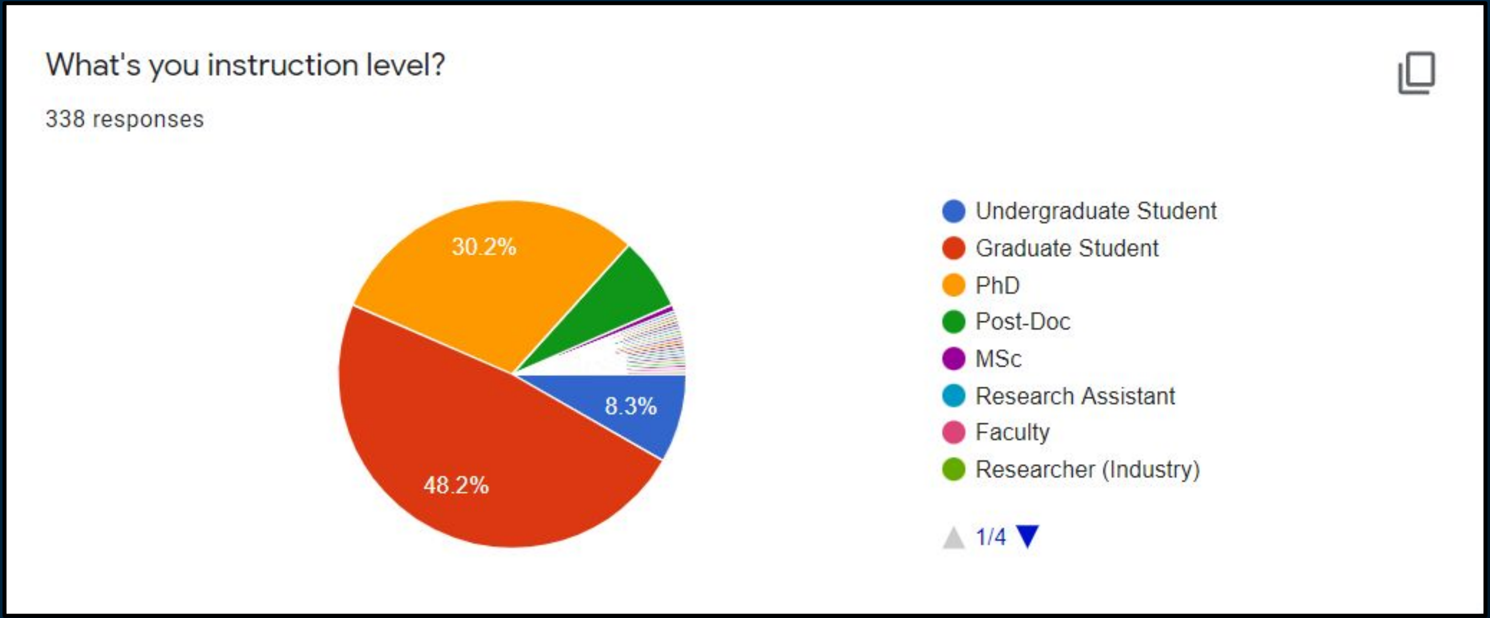
PhD Student @ Scuola Normale Superiore and University of Pisa

Expert in Continual Sequence Learning

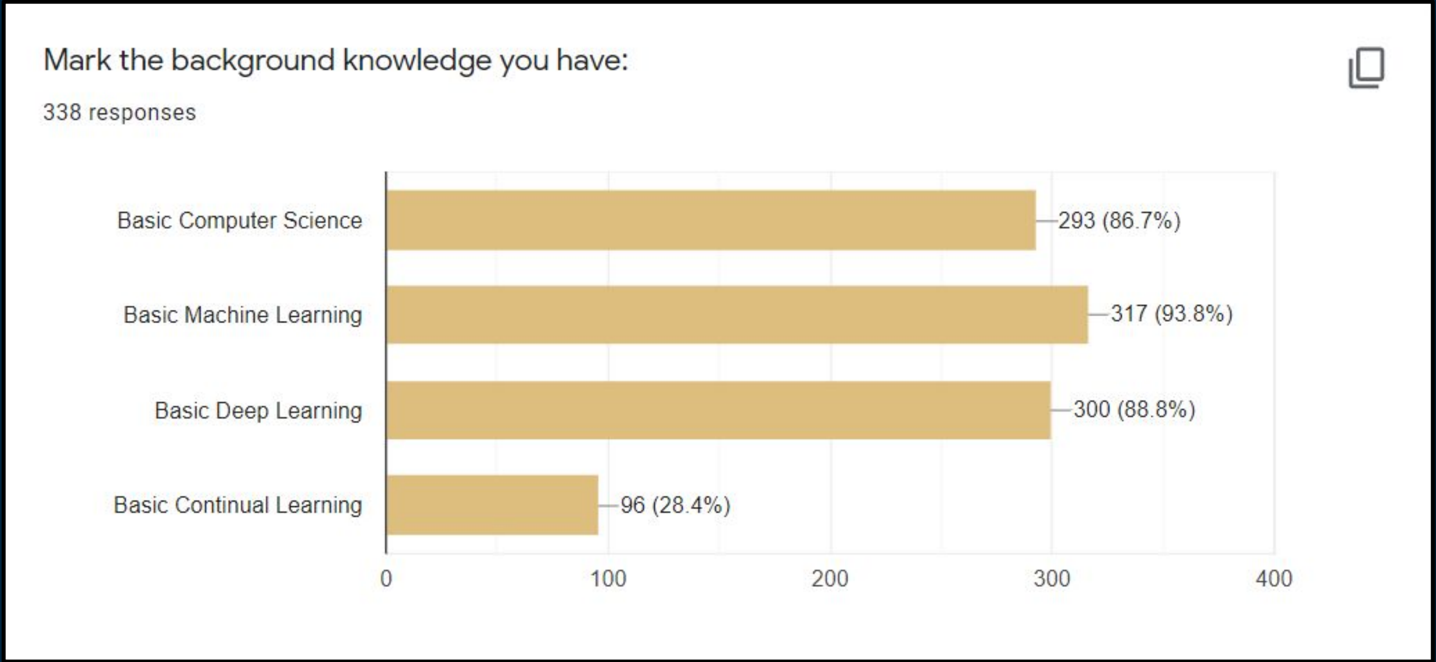
TABLE OF CONTENTS

About You

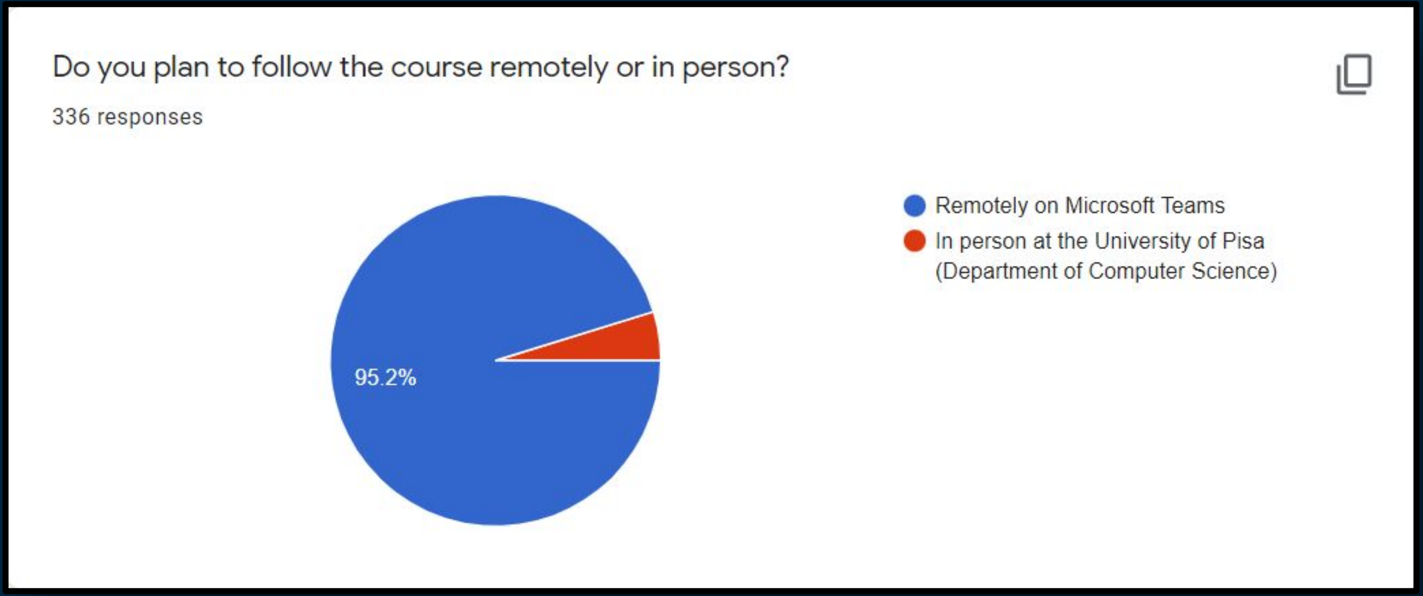
Education Level



Background Knowledge



Participation Modality



Course Modality & Structure

Course Details

Objectives

*In this course you'll learn the fundamentals of Continual Learning with Deep Architectures. At the end of this course you can expect to possess the **basic theoretical and practical knowledge** that will enable you to autonomously e**xplore more advanced topics** and frontiers in this exciting research area. You will also able to **apply such skills to your own research topics** and real-world applications.*

Mixed In-Person / Remote Modality

Where: University of Pisa, Department of Computer Science, "Sala Polifunzionale", Largo B. Pontecorvo, 356127, Pisa, Italy. The link to the **Microsoft Teams** will be sent via email to the registered participants.
Difficulty: Tailored for Graduate / PhD Students
Lectures plan: Every Monday and Wednesday 16-18 CET (Italian Time)
Period: 22/11/2021 - 20/12/2021
Language: English

Official Website: course.continualai.org

Certification and Exam

Please note that the certificate of attendance is only released after a project-based exam to be agreed with the course instructor!

The exam can be about:

1. A **google notebook showcasing an interesting CL exploration**, idea, algorithm, etc. to be added with a PR to the ContinualAI colab project.
2. **Implement a CL algorithm** and reproduce basic results creating a PR to the Reproducible-CL ContinualAI project.
3. **Add a feature to Avalanche** with a PR.
4. Develop a **custom continual learning project** based on your personal interests.

You can proceed autonomously but in case of doubt, please send an email to vincenzo.lomonaco@unipi.it with [CL-Course] in the Subject.

Prerequisites

This course has been designed for **Graduate** and **PhD Students** that have never been exposed to Continual Learning. However, it assumes basic knowledge in **Computer Science** (Bachelor level) and **Machine Learning**. In particular we assume basic knowledge in **Deep Learning**.

For students who do not have this background we suggest to follow at least an introductory *Machine Learning course*, such as the one offered by *Andrew Ng* at Coursera.

We also assume basic hands-on knowledge about:

- **Anaconda, Python and PyCharm**
- **Python Notebooks**
- **Google Colaboratory**
- **Git and GitHub**
- **PyTorch**

Make sure you learn the basics of these tools and languages as they will be used extensively across the course.

Setup & Tools

Before starting the course, please make sure you **mature some confidence** with the **following tools**:

- **Microsoft Teams**
- **Anaconda, Python and PyCharm** (or any other IDE of your choice for Python)
- **Google Colaboratory**
- **PyTorch**

Please make sure to setup your personal computer before the next lecture (it will come in handy).

Class Timetable

The course will be based on **8 main lectures** (2 hours each) and a final session composed of **2 invited talks**. Please note that these lectures will be recorded. If you don't want to be recorded you can follow the lecture on youtube.

1. **Introduction & Motivation (22-11)**
2. **Understanding Catastrophic Forgetting (24-11)**
3. **Scenarios & Benchmarks (29-11)**
4. **Evaluation & Metrics (1-12)**
5. **Methodologies [part 1] (6-12)**
6. **Methodologies [part 2] (8-12)**
7. **Methodologies [part 3] & Applications (13-12)**
8. **Frontiers in Continual Learning (15-12)**
9. **Guest Lectures (20-12)**

Introduction & Motivation

In this lecture we will address the following points:

1. ~~Course structure and modality~~
2. **What is continual learning?**
3. **Relationship with other learning paradigms**
4. **Brief history of continual Learning**

Understanding Catastrophic Forgetting

In this lecture we will address the following points:

1. What is **catastrophic forgetting**?
2. Understanding forgetting with **one neuron**
3. **A deep learning example**: Permuted and Split MNIST
4. **Avalanche**: an end-to-end library for continual learning research

Scenarios & Benchmarks

In this lecture we will address the following points:

1. Possible **continual learning scenarios**
2. **Existing and commonly used benchmarks**
3. **Avalanche Benchmarks**

Evaluation & Metrics

In this lecture we will address the following points:

1. Evaluation Protocols
2. Continual learning **metrics**
3. Avalanche Metrics & Loggers

Methodologies [Part 1]

In this lecture we will address the following points:

1. **Strategies Categorization and History**
2. **Replay Strategies**: Intro & Main Approaches
3. Avalanche Strategies & Plugins

Methodologies [Part 2]

In this lecture we will address the following points:

1. **Regularization Strategies**: Intro & Main Approaches
2. **Architectural Strategies**: Intro & Main Approaches
3. Avalanche Implementation

Methodologies [Part 3], Applications & Tools

In this lecture we will address the following points:

1. **Hybrid Strategies**: Intro & Main approaches
2. Avalanche Implementation
3. **Applications of Continual learning**: Past, Present and Future
4. The Continual Learner Toolbox

Frontiers in Continual Learning

In this lecture we will address the following points:

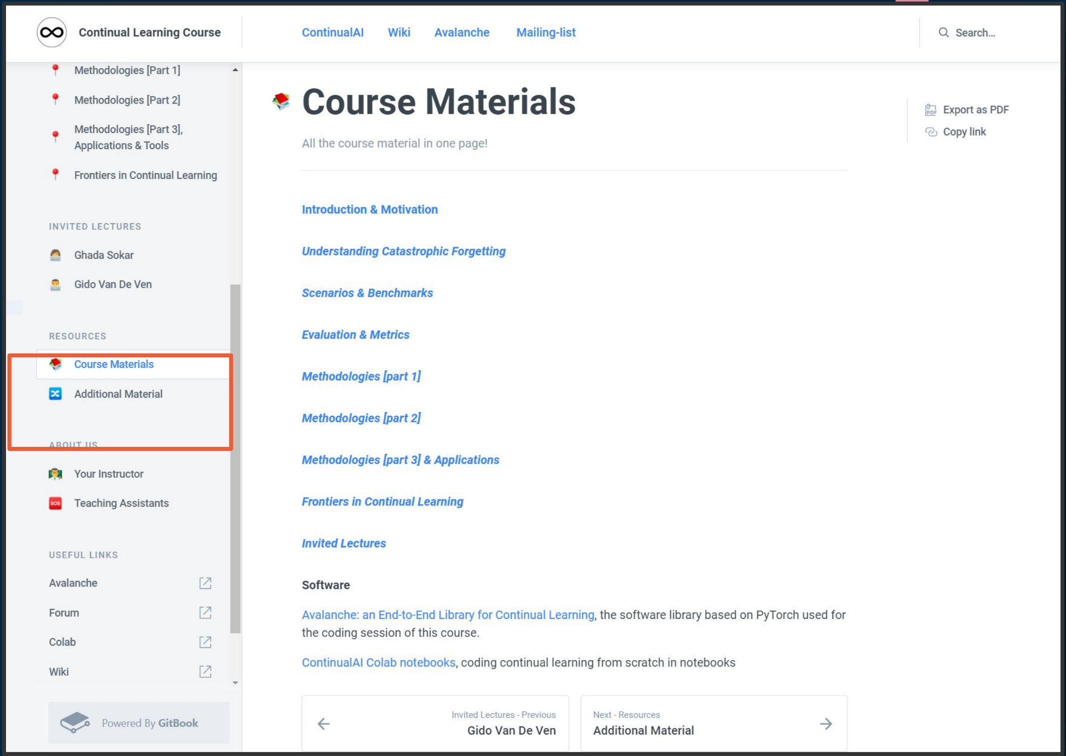
1. **Promising Future Directions**
2. **Distributed Continual Learning**
3. **Continual Sequence Learning**

Guest Lectures

In this last lecture we will host two invited talks:

1. **"Addressing the Stability-Plasticity Dilemma in Rehearsal-Free Continual Learning" - Ghada Sokar**
2. **"Using Generative Models for Continual Learning" - Gido Van De Ven**

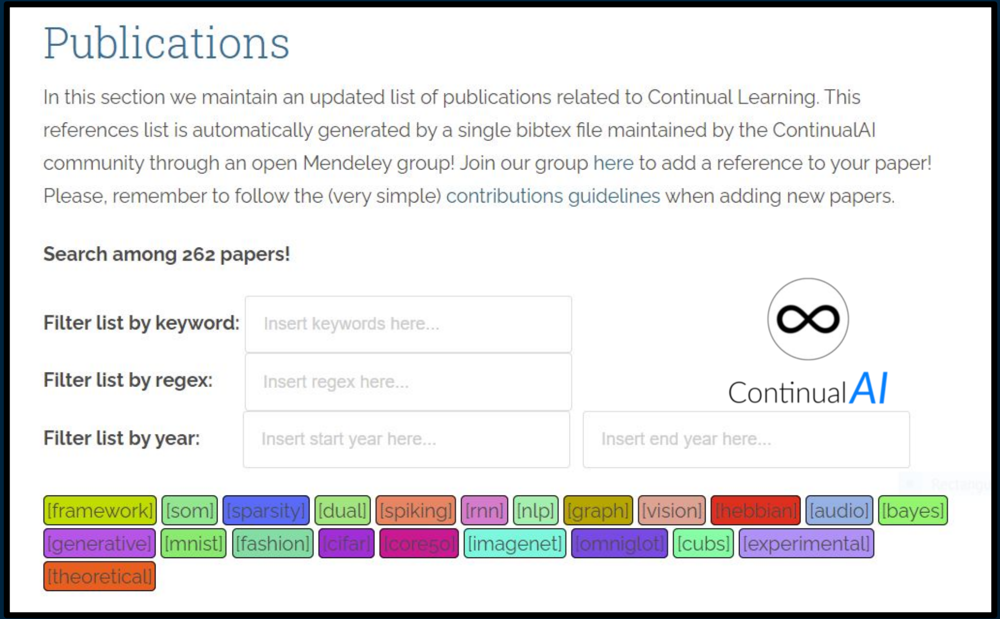
Courses Materials



You can find more at: *course.continualai.org*

Courses Materials

- **ContinualAI Wiki**: a shared and collaboratively maintained *knowledge base* for Continual Learning: tutorials, workshops, demos, tutorials, courses, etc.
- **Continual Learning Papers**: curated list of CL papers & books with meta-data by ContinualAI
- **ContinualAI Forum** + Slack: discussions / Q&As about Continual Learning
- **ContinualAI Research Consortium**: networks of Top CL Labs across the world.



Publications

In this section we maintain an updated list of publications related to Continual Learning. This references list is automatically generated by a single bibtex file maintained by the ContinualAI community through an open Mendeley group! Join our group here to add a reference to your paper! Please, remember to follow the (very simple) contributions guidelines when adding new papers.

Search among 262 papers!

Filter list by keyword: Insert keywords here...

Filter list by regex: Insert regex here...

Filter list by year: Insert start year here... Insert end year here...

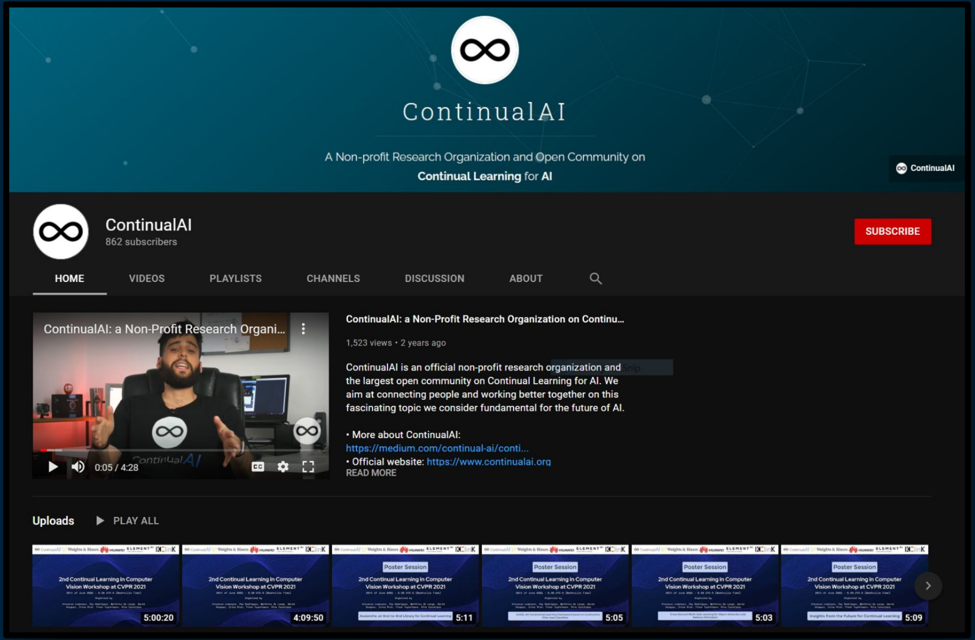
ContinualAI

[framework] [som] [sparsity] [dual] [spiking] [rnn] [nlp] [graph] [vision] [hebbian] [audio] [bayes] [generative] [mnist] [fashion] [cifar] [core50] [imagenet] [omniglot] [cubs] [experimental] [theoretical]

You can find more at: *www.continualai.org*

Courses Materials

- **ContinualAI Publication**: a curated list of original blog posts on CL.
- **Continual Learning & AI Mailing List**+: open mailing-list moderated by the ContinualAI community.
- **ContinualAI Newsletter**:news from the ContinualAI community and the CL World in one place.
- **ContinualAI Seminars**: weekly invited talks on CL.
- **ContinualAI YouTube**: collection of videos about CL.



You can find more at: *www.continualai.org*

Courses References

Here a **few reviews** and **books** you can refer to during the course:

1. **Lifelong Machine Learning**, Second Edition. by Zhiyuan Chen and Bing Liu. Synthesis Lectures on Artificial Intelligence and Machine Learning, 2018.
2. **A Continual Learning Survey: Defying Forgetting in Classification Tasks** by Matthias De Lange, Rahaf Aljundi, Marc Masana, Sarah Parisot, Xu Jia, Ales Leonardis, Gregory Slabaugh and Tinne Tuytelaars. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.
3. **Continual Learning for Robotics: Definition, Framework, Learning Strategies, Opportunities and Challenges** by Timothée Lesort, Vincenzo Lomonaco, Andrei Stoian, Davide Maltoni, David Filliat and Natalia Díaz-Rodr\ǵuez. Information Fusion, 52--68, 2020.
4. **Continual Lifelong Learning with Neural Networks**: **A Review** by German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan and Stefan Wermter. Neural Networks, 54--71, 2019.
5. **Continual Learning for Recurrent Neural Networks: An Empirical Evaluation** by Andrea Cossu, Antonio Carta, Vincenzo Lomonaco and Davide Bacciu. Neural Networks, 607--627, 2021.
6. **Replay in Deep Learning: Current Approaches and Missing Biological Elements** by Tyler L. Hayes, Giri P. Krishnan, Maxim Bazhenov, Hava T. Siegelmann, Terrence J. Sejnowski and Christopher Kanan. arXiv, 2021.
7. **Embracing Change: Continual Learning in Deep Neural Networks** by Raia Hadsell, Dushyant Rao, Andrei A Rusu and Razvan Pascanu. Trends in Cognitive Sciences, 2020.
8. **Towards Continual Reinforcement Learning: A Review and Perspectives** by Khimya Khetarpal, Matthew Riemer, Irina Rish, Doina Precup. arXiv:2012.13490, 2020.
9. A **Wholistic View of Continual Learning with Deep Neural Networks**: Forgotten Lessons and the Bridge to Active and Open World Learning by Martin Mundt, Yong Won Hong, Iuliia Pliushch and Visvanathan Ramesh. arXiv, 32, 2020.

.... More here: https://github.com/ContinualAI/continual-learning-papers#review-papers-and-books

Any Problems?

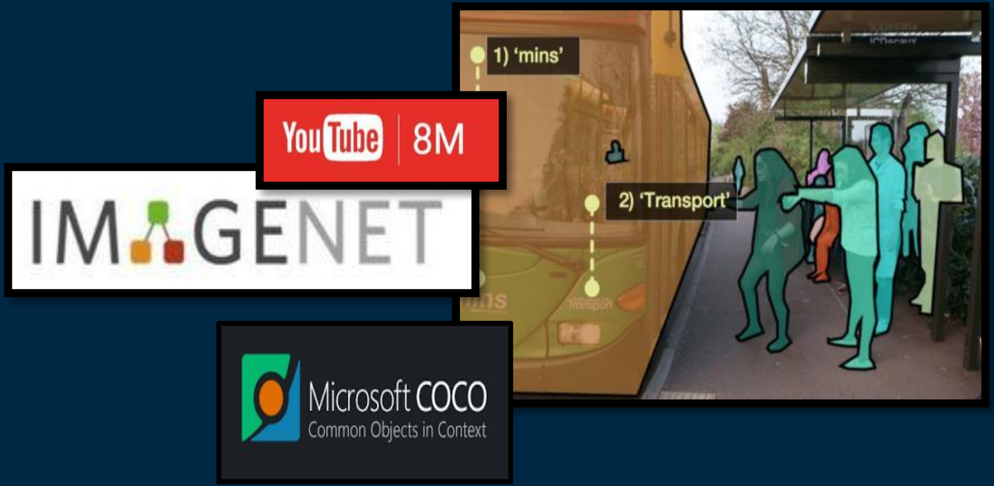
- If it's a problem related to the course material (slides, website, etc.) or modality **you can contact me**.
- If it's an Avalanche problem you can use the **Avalanche Issues** or **Discussion** pages.
- If it's a CL question you can post it on the **ContinualAI Forum**: **me**, **Antonio** and **Andrea** will try to answer there.

https://continualai.discourse.group/c/continual-learning-course-questions-discussions/15

Intro to Continual Learning

Machine Learning: State-Of-The-Art

- **Deep Learning** holds state-of-the-art performances in many tasks.
- Mainly supervised training with **huge** and **fixed** datasets.



The Curse of Dimensionality

$$256^{227 \cdot 227 \cdot 3}$$

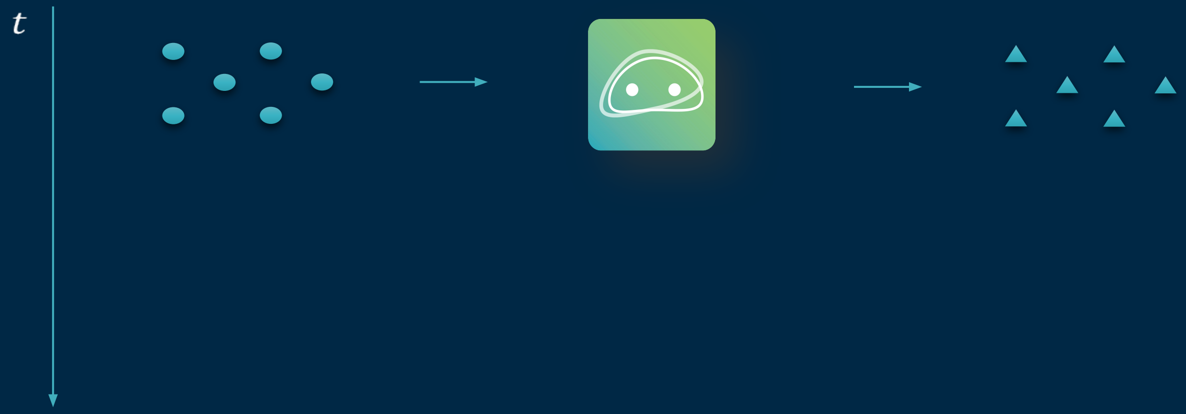
# of possible 227x227 RGB images

a group of people standing on top of a beach

How can we improve AI efficiency, scalability and adaptability?

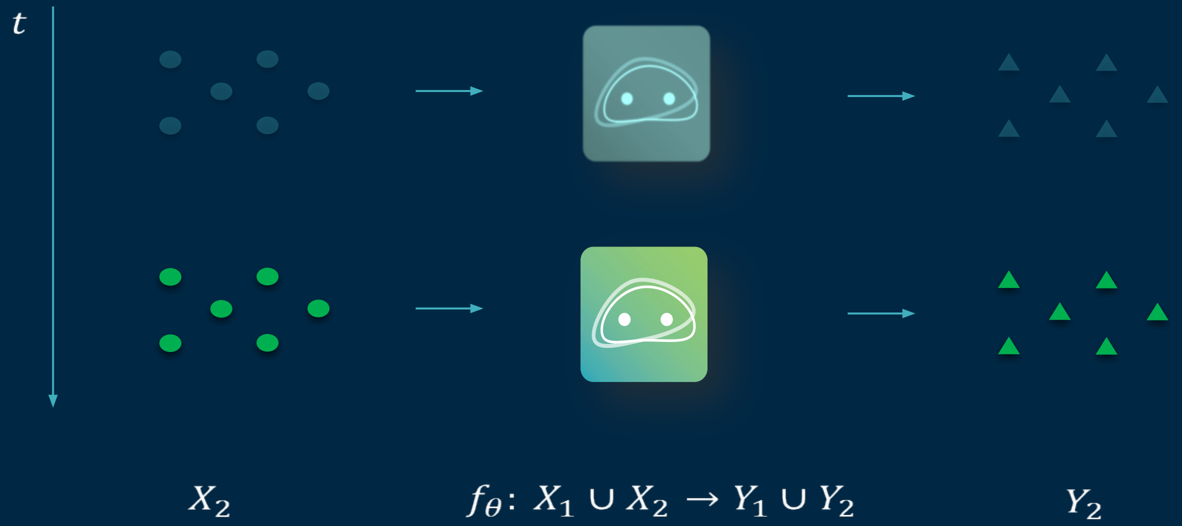
*(Hence making it **sustainable in the long term**)*

Continual Learning (CL)



X_1 $\quad f_\theta \colon X_1 \rightarrow Y_1$ $\quad Y_1$

Continual Learning (CL)



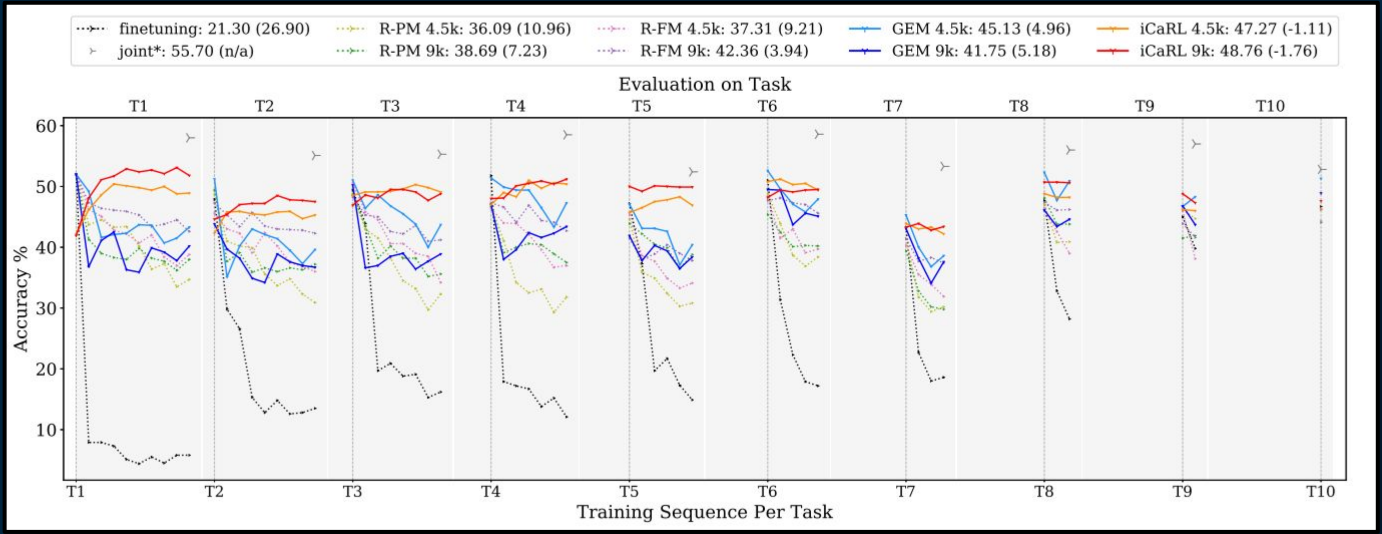
Continual Learning Desiderata

- Higher and **realistic time-scale** where data (and tasks) become available only during time.
- **No access** to previously encountered data.
- **Constant** computational and memory resources (efficiency)
- **Incremental development** of ever more complex knowledge and skills (scalability)
- **Efficiency** + **Scalability** = **Sustainability**

Sustainable Artificial Intelligence through Continual Learning. Cossu, A., Ziosi, M., Lomonaco, V., International Conference on AI for People (CAIP), 2021.

Catastrophic Forgetting (or Interference)

*Catastrophic interference, also known as **catastrophic forgetting**, is the tendency of an artificial neural networks to completely and abruptly forget previously learned information upon learning new information. ->* Mostly due to Gradient Descent.

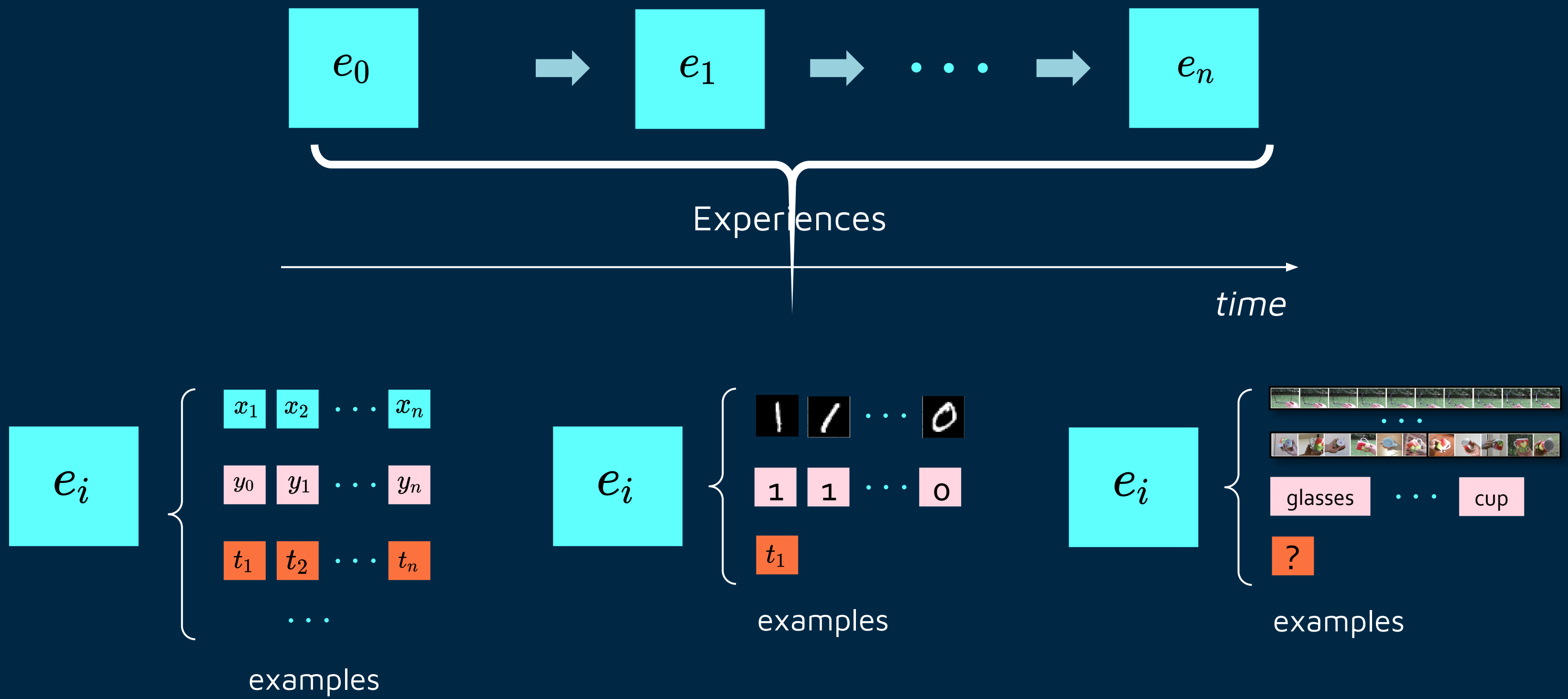


A Continual Learning Survey: Defying Forgetting in Classification Tasks by Matthias De Lange, Rahaf Aljundi, Marc Masana, Sarah Parisot, Xu Jia, Ales Leonardis, Gregory Slabaugh and Tinne Tuytelaars. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021.

“We are not looking for incremental improvements in state-of-the-art AI and neural networks, but rather paradigm-changing approaches to machine learning that will enable systems to continuously improve based on experience.”

– Hava Siegelmann, 2018

Continual Learning



Continual Learning (more formally)

In continual learning (CL) data arrives in a streaming fashion as a (possibly infinite) sequence of learning experiences $S = e_1, \ldots, e_n$. For a supervised classification problem, each experience e_i consists of a batch of samples $\mathcal{D}^i$, where each sample is a tuple $\langle x_k^i, y_k^i \rangle$ of input and target, respectively, and the labels y_k^i are from the set $\mathcal{Y}^i$, which is a subset of the entire universe of classes $\mathcal{Y}$. Usually $\mathcal{D}^i$ is split into a separate train set $\mathcal{D}^i_{train}$ and test set $\mathcal{D}^i_{test}$.

A continual learning algorithm $\mathcal{A}^{CL}$ is a function with the following signature:

$$\mathcal{A}^{CL} : \langle f_{i-1}^{CL}, \mathcal{D}^i_{train}, \mathcal{M}_{i-1}, t_i \rangle \rightarrow \langle f_i^{CL}, \mathcal{M}_i \rangle \quad (1)$$

where f_i^{CL} is the model learned after training on experience e_i, $\mathcal{M}_i$ a buffer of past knowledge, such as previous samples or activations, stored from the previous experiences and usually of fixed size. The term t_i is a task label which may be used to identify the correct data distribution.

The objective of a CL algorithm is to minimize the loss $\mathcal{L}_S$ over the entire stream of data S:

$$\mathcal{L}_S(f_n^{CL}, n) = \frac{1}{\sum_{i=1}^{n} |\mathcal{D}^i_{test}|} \sum_{i=1}^{n} \mathcal{L}_{exp}(f_n^{CL}, \mathcal{D}^i_{test}) \quad (2)$$

$$\mathcal{L}_{exp}(f_n^{CL}, \mathcal{D}^i_{test}) = \sum_{j=1}^{|\mathcal{D}^i_{test}|} \mathcal{L}(f_n^{CL}(\boldsymbol{x}_j^i), y_j^i), \quad (3)$$

where the loss $\mathcal{L}(f_n^{CL}(\boldsymbol{x}), y)$ is computed on a single sample $\langle \boldsymbol{x}, y \rangle$, such as cross-entropy in classification problems.

Desiderata

- Replay-Free Continual Learning
- Memory and Computationally Bounded
- Task-free Continual Learning
- Online Continual Learning

Continual Learning for Robotics: Definition, Framework, Learning Strategies, Opportunities and Challenges, Lesort et al. Information Fusion, 2020.

Brief History
Of CL Machines

Nomenclature & Related Paradigms

Unconsolidated Nomenclature

- **Continual Learning**
- **Incremental Learning**
- **Lifelong Learning**
- **Continuous Learning**

Related Paradigm

- **Multi-Task Learning**
- **Meta-Learning / Learning to Learn**
- **Transfer Learning & Domain Adaptation**
- **Online / Streaming Learning**

Continual Learning for Robotics: Definition, Framework, Learning Strategies, Opportunities and Challenges, Lesort et al. Information Fusion, 2020.

History Timeline

1. Incremental learning with rule-based systems (**`70s - `80s**)
2. Forgetting in Neural Networks (**French, 1989**)
3. Incremental learning with Kernel Machines (**`90s**)
4. Continual Learning (**Ring, 1998**)
5. Lifelong Learning (**Thrun, 1998**)
6. Never-Ending Learning (**Mitchell, 2009**)
7. Deep Continual Learning (**Kirkpatrick, 2016**)
8. Lifelong (Language) Learning (**Liu, 2018**)

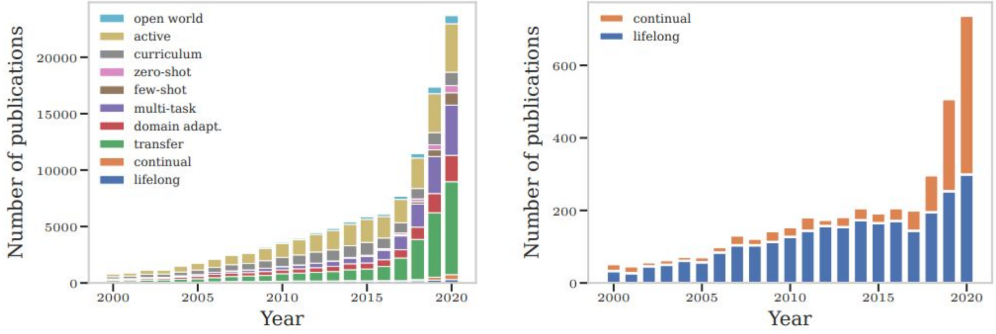


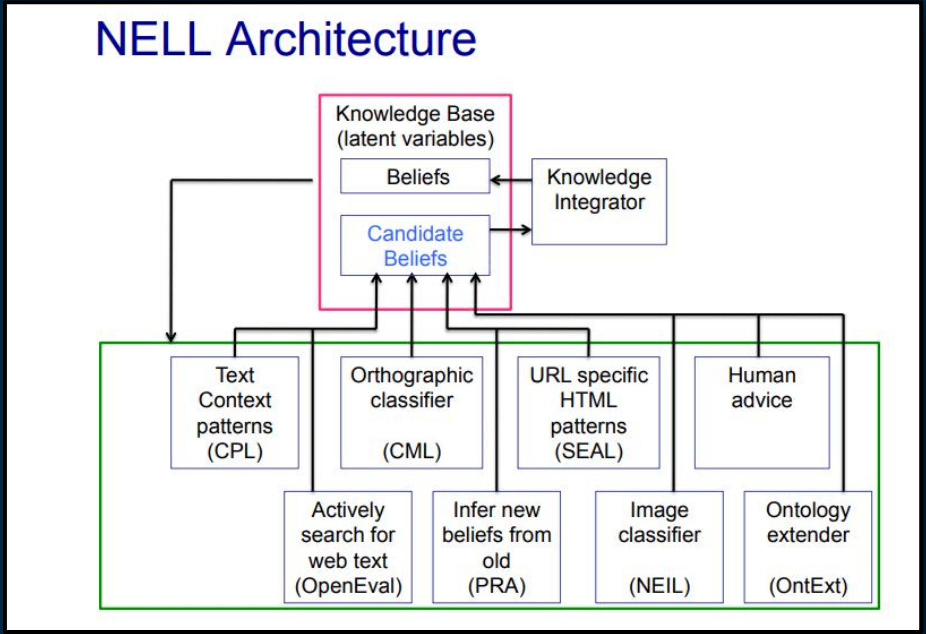
Figure 1: Per year machine learning publications. Left: cumulative amount of papers across keywords with continuous components that influence continual learning practice, see Section 2. Right: increasing use of "continual" machine learning, demonstrating a shift in use of terminology with respect to the preceding emphasis on the term "lifelong". Data queried using the Microsoft Academic Graph utilities (Sinha et al., 2015) based on keyword occurrence in the abstract.

CLEVA-Compass: A Continual Learning EValuation Assessment Compass to Promote Research Transparency and Comparability, Martin Mundt et al. preprint, 2021

NELL: Never-Ending Language Learning (a Key Milestone)

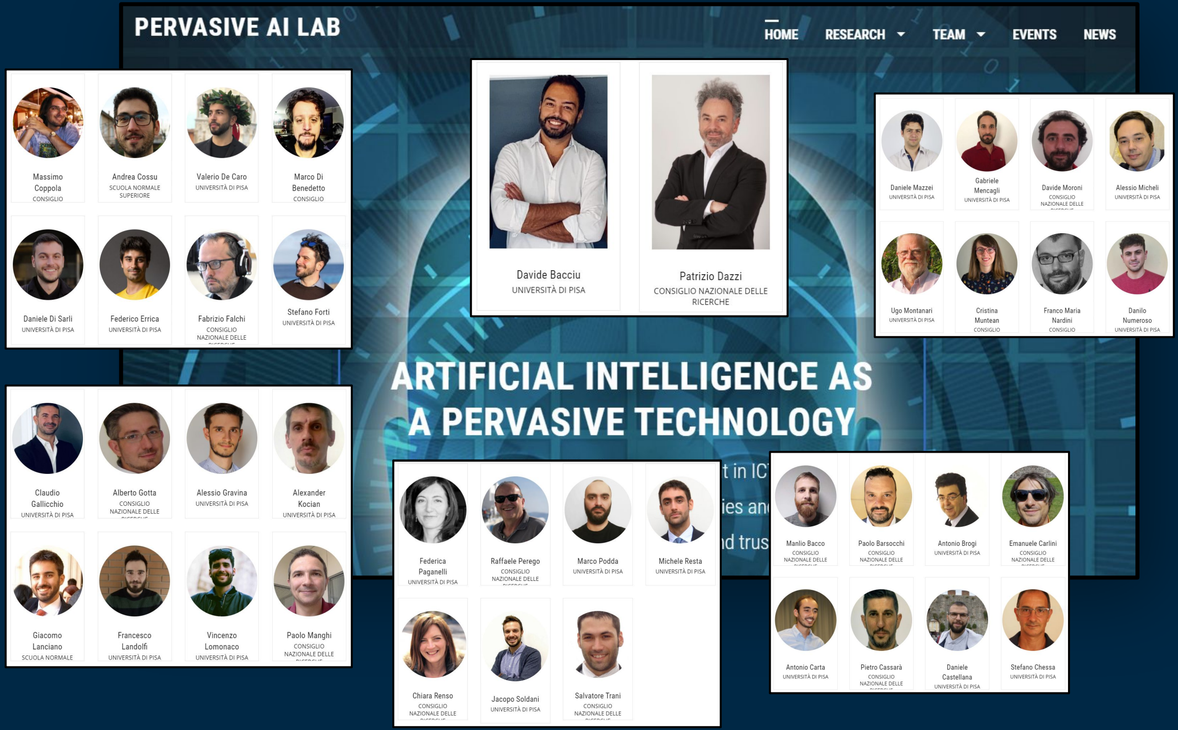
Key Ideas

- *Semi-Supervised Learning System*
- Ran 24x7, from January, 2010 to September 2018
- Combination of many learning algorithms (CPL, CML,SEAL, OpenEval, PRA, NEIL)
- Intended as a case-study for a **never-ending agent**



T. Mitchell et al. ***Never-Ending Learning***. Communications of the ACM, 2018.

Join our Pervasive AI Lab!



Teaching & Supervision

Research

Spin-off

Consultancy

The lab is in Pisa, Italy! Feel free to visit and get in touch with us anytime! Official website: Pervasive AI Lab (unipi.it)

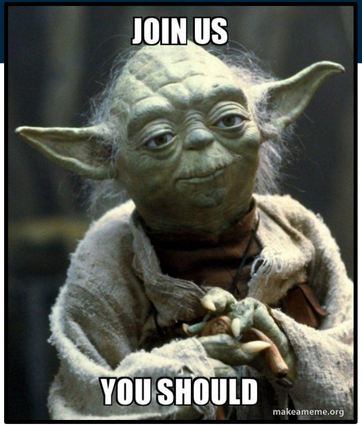
CL @ PAI Lab

At the **PAI Lab** we design and implement **deep continual learning algorithms** for enabling the next generation AI systems and study their applications to real-world problems. In particular, we are interested in:

- *Unsupervised / Self-Supervised/ Weekly/ Semi-Supervised Continual Learning*
- *Continual Sequence Learning*
- *Neuroscience-Inspired Continual Learning*
- *Continual Reinforcement Learning*
- *Continual Learning R&D Frameworks & Tools*
- *Continual Robot Learning*
- *Continual learning on the Edge*
- *Distributed Continual Learning*
- *Real-World Continual Learning Applications*
- *...and much more!*

Team

- **Davide Bacciu** – Associate Professor
- **Vincenzo Lomonaco** – Assistant Professor
- **Claudio Gallicchio** – Assistant Professor
- **Antonio Carta** – Post-Doc
- **Andrea Cossu** – PhD Student
- **Rudy Semola** – PhD Student
- **Michele Resta** – PhD Student
- **Valerio De Caro** – PhD Student
- **Hamed Hemati** – PhD Student (co-supervised with Damian Borth at University of St. Gallen)



The lab is in Pisa, Italy! Feel free to visit and get in touch with us anytime! Official website: [Pervasive AI Lab (unipi.it)]

Next:

Understanding Catastrophic Forgetting

Do you have any questions?

vincenzo.lomonaco@unipi.it
vincenzolomonaco.com
University of Pisa

THANKS

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon, and infographics & images by Freepik