

# **DATASCI 185 - Introduction to AI Applications**

Danilo Freire

Fall 2026

## **Course description**

Welcome to [DATASCI 185!](#) This course offers a non-technical introduction to artificial intelligence and its effects on institutions, work and everyday life. The course is practical: students will learn how modern systems are designed and how to ask the right questions about data, reliability and harms. The course combines short demonstrations (no programming experience required), case studies, and project work. It is suitable for undergraduate students from any faculty who want a grounded understanding of what AI can and cannot do.

## **Learning objectives**

By the end of the course, students will be able to:

- Explain the main ideas behind contemporary AI systems in plain language.
- Identify common failure modes of AI systems and the data issues that cause them.
- Read and assess claims about AI in news articles, product pages and policy documents.
- Design a small, realistic plan for an AI application, including data needs, evaluation, and a basic harm-mitigation strategy.
- Reflect critically on ethical, legal and social questions raised by AI deployment.

# Course logistics

- **Lectures:** Mondays and Wednesdays, 4-4:50pm.
- **Location:** Psychology Building, Room 290.
- **Instructor:** Danilo Freire ([daniolo.freire@emory.edu](mailto:daniolo.freire@emory.edu)). Office hours by appointment; please email me to arrange a time.
- **Materials:** Lecture notes, tutorials, and assignment templates are on [the course GitHub repository](#). Students must check the repository often for updates.

## Prerequisites and software

No formal prerequisites. Familiarity with spreadsheets, basic statistics or introductory programming is useful but not required. We will discuss some simple Python code snippets in class, but no programming experience is needed.

## Readings and resources

Readings are short papers, essays and tutorials rather than a single textbook. Links and PDFs will be provided on the course website. Please note that the syllabus may be updated during the semester. For those interested in digging deeper into some of the topics we cover, here are some recommended books and resources:

### Books

- [Artificial Intelligence: A Guide for Thinking Humans](#) by Melanie Mitchell (2020). A clear, non-technical overview of AI concepts and history.
- [The Master Algorithm](#) by Pedro Domingos (2015). A readable introduction to machine learning concepts and applications.
- [The Coming Wave: AI, Power, and Our Future](#) by Mustafa Suleyman and Michael Bhaskar (2025). An award-winning book on the transformative potential of AI and its societal implications.

- [The Alignment Problem: Machine Learning and Human Values](#) by Brian Christian (2020). An exploration of what can go wrong when AI systems are deployed in the real world.
- [AI & I: An intellectual history of artificial intelligence](#). by Eugene Charniak (2024). A new book by a leading AI researcher, covering the history of the field from the 1950s to the present day.
- [Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy](#) by Cathy O’Neil (2016). A book on the dangers of unregulated mathematical models.
- [Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence](#) by Kate Crawford (2021). A critical look at the social and environmental impacts of AI technologies.

## Online courses

- [AI For Everyone](#) by Andrew Ng (Coursera). A non-technical introduction to AI concepts and applications. Free to audit, with a paid certificate option.
- [Elements of AI](#) by the University of Helsinki. A free, friendly introduction to AI and some of its methods.
- [AI Essentials](#) by Google (Coursera). This programme is designed for people who want to gain practical AI skills for the workplace with no experience required.
- [Google Prompting Essentials](#) by Google (Coursera). A short course on how to effectively use and design prompts for large language models.
- [Radical Ideas in AI Ethics](#) by Pragmatic AI Labs (edX). A course that discusses AI through the lens of human rights and digital autonomy.

## Other resources

- [In Machines We Trust](#). A podcast series by MIT Technology Review about the promises and perils of AI. Very accessible and engaging.
- [AI for the Rest of Us](#). 25-30 minute episodes focused on explaining AI concepts to non-technical listeners.
- [The Gradient](#). A publication that features accessible articles on AI research.
- [AI Now Institute](#). Research institute focused on the social implications of AI.

- [AI for Good Lab](#). A Microsoft research group that highlights how AI can change society for the better.
- [The AI Incident Database](#). A collection of real-world cases where AI systems have caused harm or failed.
- [Teachable Machine](#). A web-based tool by Google that allows users to create simple machine learning models without coding.
- [The Algorithm](#). A newsletter by MIT Technology Review that covers the latest developments in AI.

## Assessments

- **Problem sets (10) – 50%**. Short conceptual and practical tasks. Submit Jupyter notebooks (.ipynb), Word documents, or PDFs. Late submissions incur a 10% penalty per day unless authorised in advance. Collaboration for discussion is allowed but answers must be written independently; list collaborators on submission.
- **In-class quizzes (5) – 30%**. Each quiz occupies a full lecture. Quizzes are open-book/open-notes and individual assessments. Discussion during quizzes is not permitted.
- **Final group project – 20%**. Groups of 3–4 produce a report and a short demo video. The project involves conceptual elements and a harms assessment. More details will be provided in class.

## Grading scale

Grade	A	A-	B+	B	B-	C	D	F
Range	91–100	86–90	81–85	76–80	71–75	66–70	60–65	<60

Final grades are rounded to the nearest point.

## **Academic integrity and AI use**

Generative AI tools (for example ChatGPT, Claude, GitHub Copilot) may be used as aids for brainstorming or drafting, but all AI use must be declared on submissions. Students must take responsibility for correctness and must attribute any direct text or code produced by AI. The Emory Honour Code applies.

## **Emory Honour Code**

The Emory Undergraduate Academic Honour Code is in effect throughout the semester. The Honour Code applies to any action or inaction that fails to meet the communal expectations of academic integrity. Students should strive to excel in their academic pursuits in a just way with honesty and fairness in mind and avoid all instances of cheating, lying, plagiarizing, or engaging in other acts that violate the Honor Code. Such violations undermine both the individual pursuit of knowledge and the collective trust of the Emory community. Students who violate the Honour Code may be subject to failure of the course, a reportable record, suspension, permanent expulsion, or a combination of these and other sanctions. The Honor Code may be reviewed at: <http://catalog.college.emory.edu/academic/policies-regulations/honor-code.html>.

## **Accessibility**

Students who require adjustments should contact the Department of Accessibility Services early in the semester and provide documentation. Reasonable accommodations will be made for assessments, including quizzes, according to DAS guidance. Please contact me privately to discuss any specific needs.

## **Course schedule**

### **Module 0 – Orientation**

#### **1. Lecture 1 – Course introduction; what AI means today**

Syllabus, assessment, expectations. A factual tour of AI in products and media and the questions

we will address.

Readings:

- Stanford University's Human-Centered Artificial Intelligence (2025). [AI Index Annual Report](#). An excellent report on the state of AI, with many useful charts and references.
- Our World in Data (2025). [Artificial intelligence](#). A concise overview of AI trends and applications, with links to further resources.
- Menlo Ventures (2025). [2025: the state of consumer AI](#). How AI is being used in consumer products, with examples and market data.
- Morgan Stanley (2025). [AI's next leap: 5 trends shaping innovation and ROI](#). A short text on the current state of AI applications in industry.
- McKinsey & Company (2025). [The state of AI: how organizations are rewiring to capture value](#). A business-focused overview of AI adoption and impact.

## 2. Lecture 2 – A brief history of AI and the recent shift

From symbolic approaches to data-driven learning; the transformer architecture explained in plain language.

Readings:

- Dick, S. (2019). [Artificial intelligence](#). *Harvard Data Science Review*, 1(1). A brief history of AI research and applications. Written for a general audience.
- Grzybowski, A., Pawlikowska-Łagód, K., & Lambert, W. C. (2024). [A history of artificial intelligence](#). *Clinics in Dermatology*, 42(3), 221-229. A “pre-history” overview, focusing on early ideas and milestones.
- Cao, Y., Liu, L., Li, M., Huang, Y., & Li, S. (2023). [A comprehensive survey of AI-generated content \(AIGC\): A history of generative AI from GAN to ChatGPT](#). *arXiv preprint*. A somewhat technical but comprehensive overview of generative AI methods.
- Halevy, A., Norvig, P., & Pereira, F. (2009). [The unreasonable effectiveness of data](#). *IEEE Intelligent Systems*, 24(2), 8-12. A classic paper arguing that large datasets are often more important than sophisticated algorithms.

- Pradhan, M. (2023). [A non-technical introduction to Transformers](#). A clear explanation of the transformer architecture that underpins many modern AI systems.
- Doshi, K. (2025). [Transformers explained visually, part 1: overview of functionality](#). Another explanation of transformers, with diagrams.
- Horstman, M. (2024). [Transformers for dummies: a peek inside AI models](#). *Yet another* explanation of transformers. This one is a little more technical.
- Georgia Tech's Polo Club of Data Science (2025). [Transformer explainer](#). An interactive visualisation of how transformers work.

## Module 1 – How AI systems are designed

### 3. Lecture 3 – Dataset design, labels, and tasks

What data are, how tasks are framed, and common pitfalls in dataset design.

Readings:

- Athalye, A., Northcutt, C., & Mueller, J. (2024). [Dataset creation and curation](#) in *Data-Centric AI*. A chapter from a forthcoming book on data-centric AI, covering the main ideas in dataset design.
- Sapien Labs. (2024). [Labeling data for machine learning: best practices and quality control](#). A practical guide to data labeling, with tips and common pitfalls.
- Liang, W., Tadesse, G. A., Ho, D., Fei-Fei, L., Zaharia, M., Zhang, C., & Zou, J. (2022). [Advances, challenges and opportunities in creating data for trustworthy AI](#). *Nature Machine Intelligence*, 4(8), 669-677. A good discussion on the challenges of creating high-quality datasets.
- Orr, W., & Crawford, K. (2024). [The social construction of datasets: on the practices, processes, and challenges of dataset creation for machine learning](#). *New Media & Society*, 26(9), 4955-4972. A sociological perspective on dataset creation, with interviews with practitioners.
- Nazer, L. H., Zatarah, R., Waldrip, S., Ke, J. X. C., Moukheiber, M., Khanna, A. K., ... & Mathur, P. (2023). [Bias in artificial intelligence algorithms and recommendations for mitigation](#). *PLOS Digital Health*, 2(6), e0000278. A review of different types of bias that can arise in datasets and what to do about them. Focused on healthcare but broadly applicable.

#### **4. Lecture 4 – Supervised, unsupervised, and hybrid approaches to learning**

What learning means and why generalisation is difficult.

Readings:

- Amazon Web Services (2025). [What is the difference between supervised and unsupervised learning?](#) A practical overview of supervised and unsupervised learning, with examples.
- Lee, F. (2025). [What is bias-variance tradeoff?](#) An explanation of a key concept in machine learning. It has some formulas but the main ideas are easy to grasp.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). [Chapter 5: Machine learning basics](#) in *Deep Learning*. A technical but clear introduction to the main ideas behind machine learning.
- Elsevier (2025). [Science direct: supervised learning.](#) A short introduction to supervised learning, with links to further readings.
- Elsevier (2025). [Science direct: unsupervised learning.](#) Same as above, but for unsupervised learning.

#### **5. Lecture 5 – Metrics, validation and overfitting**

Confusion matrices, precision and recall, cross-validation and the limits of single-number metrics.

Readings:

- Akinkugbe, A. (2025). [The essential guide to model evaluation metrics for classification.](#) A practical guide to common evaluation metrics, with examples and case studies.
- Thomas, R. L., & Umansky, D. (2022). [Reliance on metrics is a fundamental challenge for AI.](#) *Patterns*, 3(5).100440. A discussion of the limitations of metrics and how they can be gamed or misinterpreted.
- Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). [Overfitting, model tuning, and evaluation of prediction performance.](#) In *Multivariate statistical machine learning methods for genomic prediction* (pp. 109-139). Cham: Springer International Publishing. Good chapter, worth reading even if you don't understand all the maths.
- Coursera (2025). [Precision vs. recall in machine learning: what's the difference?](#) A short article about those two commonly confused metrics.

- Palumbo, G., Carneiro, D., & Alves, V. (2024). [Objective metrics for ethical AI: a systematic literature review](#). *International Journal of Data Science and Analytics*, 1-21. New ideas about how to evaluate AI systems beyond accuracy.

## 6. **Lecture 6 – Quiz 1**

# **Module 2 – Language and perception**

## 7. **Lecture 7 – How machines handle language: tokens, embeddings and context**

An intuitive view of tokenisation, vector representations and contextual meaning.

Readings:

- Neptune.ai (2025). [The ultimate guide to word embeddings](#). The title is a bit grandiose, but this is a clear and practical introduction to word embeddings. Includes Python code snippets.
- Neptune.ai (2025). [Tokenization in NLP: types, challenges, examples, tools](#). Another article by the same group, this time on tokenisation. Also includes code snippets.
- StackOverflow (2023). [An intuitive introduction to text embeddings](#). A practical introduction to embeddings, with simple examples.
- DeepLearning.ai (2023). [A complete guide to natural language processing](#). Despite another grandiose title, this is a good overview of NLP concepts and techniques.
- Gastaldi, J. L., Terilla, J., Malagutti, L., DuSell, B., Vieira, T., & Cotterell, R. (2024). [The foundations of tokenization: statistical and computational concerns](#). *arXiv preprint arXiv:2407.11606*. Very technical, for those interested in the mathematical details of tokenisation.

## 8. **Lecture 8 – How machines analyse images and audio**

Abstraction, convolution and the role of large datasets.

Readings:

- O’Shea, K., & Nash, R. (2015). [An introduction to convolutional neural networks](#). *arXiv preprint arXiv:1511.08458*. A clear and concise tutorial.

- DataCamp (2025). [What is computer vision? A beginner guide to image analysis](#). Simple and illustrated introduction to computer vision concepts.
- AltexSoft (2025). [How AI sound, music, and voice generation works](#). Another practical introduction. A nice reading!
- Singal, A. (2023). [Hearing is believing: revolutionizing AI with audio classification via computer vision](#). How machines classify audio using image-based techniques. Includes Python code for practitioners.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). [Deep learning](#). MIT Press. Chapters 9 and 10 provide a solid foundation in convolutional networks. A bit technical, quite dense, but this is one of the best books on deep learning.

## 9. Lecture 9 – Behavioural testing: examples and demonstrations

Live demos: prompt responses, image classifiers, and stress-tests.

No required readings for this lecture. We will do live demonstrations in class.

## 10. Lecture 10 – Quiz 2

Covers Lectures 7–9.

# Module 3 – Retrieval, generation and pipelines

## 11. Lecture 11 – Retrieval-augmented workflows and semantic search

Augmenting generation with retrieval; vector search and chunking explained for practitioners.

Readings:

- Amazon Web Services (2025). [What is retrieval-augmented generation \(RAG\)?](#) Beginners' guide to RAG by people who know what they're talking about.
- Erikson, J. (2024). [What is vector search? The ultimate guide](#). Yes, the grandiose titles are a trend. The article is pretty good, though.
- Mastering LLM (2024). [11 chunking strategies for RAG – simplified & visualized](#) Very nice explanation of chunking strategies, with diagrams.

- Slack (2025). [Semantic search explained: how it works, and why it matters](#). Good explainer from a top tech company.
- Wang, X., Wang, Z., Gao, X., Zhang, F., Wu, Y., Xu, Z., ... & Huang, X. (2024). [Searching for best practices in retrieval-augmented generation](#). *arXiv preprint arXiv:2407.01219*. Technical in nature, for those interested in the research behind RAG.

## 12. Lecture 12 – Creativity and hallucination

Why models invent facts and how to spot and reduce hallucination.

Readings:

- Jiang, X., Tian, Y., Hua, F., Xu, C., Wang, Y., & Guo, J. (2024). [A survey on large language model hallucination via a creativity perspective](#). *arXiv preprint arXiv:2402.06647*. Very nice article showing how hallucinations can be used creatively.
- Tonmoy, S. M. T. I., Zaman, S. M., Jain, V., Rani, A., Rawte, V., Chadha, A., & Das, A. (2024). [A comprehensive survey of hallucination mitigation techniques in large language models](#). *arXiv preprint arXiv:2401.01313*, 6. The article is quite detailed and lists a series of techniques to reduce hallucinations. No need to know them all, just to know they exist and how they work in principle.
- Wikipedia (2025). [Hallucination \(artificial intelligence\)](#). Wikipedia actually has some good articles on AI topics. This one is a well-written overview of hallucinations.
- Chakraborty, T., & Masud, S. (2024). [The promethean dilemma of AI at the intersection of hallucination and creativity](#). *Communications of the ACM*, 67(10), 26-28. Full of interesting ideas and perspectives.
- OpenAI (2024). [Does ChatGPT tell the truth?](#) I'm sure you know who these people are. Practical advice.
- Kalai, A. T., Nachum, O., Vempala, S. S., & Zhang, E. (2025). [Why language models hallucinate](#). *arXiv preprint arXiv:2509.04664*. Written by members of OpenAI. Not for the faint of heart by any means, but if you're feeling adventurous...

### **13. Lecture 13 – Building reliable pipelines: monitoring and testing**

Overview of simple strategies for AI evaluation; what operational teams track.

Readings:

- Google (2025). [Production ML systems: monitoring pipelines](#). If there's one company that knows how to run ML systems at scale, it's Google. This is a practical guide to monitoring ML pipelines. If you are unfamiliar with some terms, don't worry too much; the main ideas are easy to grasp. Use AI to help you if needed!
- Allen, J. (2024). [Data pipeline monitoring: best practices for full observability](#). Informal and full of practical tips. Some technical terms too, but nothing too difficult.
- Steidl, M., Felderer, M., & Ramler, R. (2023). [The pipeline for the continuous development of artificial intelligence models–current state of research and practice](#). *arXiv e-prints*, arXiv-2301. *Super long article* (sorry!), but it covers everything we discuss in class and more. Just read what you need.

### **14. Lecture 14 – Quiz 3**

Covers Lectures 11–13.

## **Module 4 – Data ethics and bias**

### **15. Lecture 15 – Types of bias and how they arise**

Selection bias, label bias and measurement error.

Readings:

- Barocas, S., Hardt, M., & Narayanan, A. (2019). [Fairness and machine learning](#). Chapter 2: When is automated decision making legitimate? A very good book on fairness in machine learning. This chapter covers the main types of bias that can arise in datasets.
- Mehrabi, N., Morstatter, F., Saxena, N., Lettich, A., & Galstyan, A. (2021). [A survey on bias and fairness in machine learning](#). *ACM Computing Surveys (CSUR)*, 54(6), 1-35. A comprehensive survey of bias and fairness in machine learning. Quite technical, but the introduction and conclusion are accessible.

- Suresh, H., & Guttag, J. V. (2019). [A framework for understanding unintended consequences of machine learning](#). In *Conference on Fairness, Accountability, and Transparency* (pp. 1-10). A good framework for thinking about biases.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). [On the dangers of stochastic parrots: Can language models be too big?](#) In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). A critical look at large language models and the risks they pose.
- Silberg, J., & Manyika, J. (2019). [Notes from the AI frontier: tackling bias in AI \(and in humans\)](#). McKinsey Global Institute, 1(6), 1-31. An easy-to-read report on bias in AI, with practical recommendations for business and policy.

## 16. Lecture 16 – Documentation and dataset governance

Datasheets, model cards and provenance.

Readings:

- Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J. W., Wallach, H., Daumé III, H., & Crawford, K. (2018). [Datasheets for datasets](#). *arXiv preprint arXiv:1803.09010*. The original paper proposing datasheets for datasets.
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., ... & Gebru, T. (2019). [Model cards for model reporting](#). In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (pp. 220-229). The original paper proposing model cards.
- Bender, E. M., & Friedman, B. (2018). [Data statements for natural language processing: Toward mitigating system bias and enabling better science](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing* (pp. 1-14). A proposal for data statements to accompany NLP datasets.
- IBM (2025). [What is data lineage?](#) A short guide to data lineage, with use cases.
- IBM (2025). [What is data provenance?](#) Another practical guide, this time on data provenance.

## 17. Lecture 17 – Case studies: biased outcomes in health, hiring and policing

Examples that illustrate impact and possible mitigations.

Readings:

- Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature machine intelligence*, 1(5), 206-215. A strong argument for using interpretable models in important aspects of life.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453. A landmark study showing how a widely used health algorithm exhibited racial bias.
- Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1-15. A study revealing bias in facial recognition systems.
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica*. A widely cited investigation into bias in criminal risk assessment tools.
- Wilson, K., & Caliskan, A. (2024, October). Gender, race, and intersectional bias in resume screening via language model retrieval. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (Vol. 7, pp. 1578-1590). A recent paper on how AI may be biased in hiring.
- AI Now Institute (2025). Artificial Power: 2025 Landscape Report. A strongly critical report on the power dynamics in the AI ecosystem. Worth reading even if you disagree with the arguments.

#### 18. **Lecture 18 – Quiz 4**

Covers Lectures 15–17.

### **Module 5 – Policy, governance and social impact**

#### 19. **Lecture 19 – AI regulation and standards around the world**

American AI Initiative, OECD principles, EU AI Act summary and how rules influence design.

Readings:

- The White House. (2025). [Winning the race: America's AI action plan](#). What the US government wants to do about AI.
- OECD. (2019). [OECD AI principles](#). The original text of the OECD AI principles. Just skim the main points.
- OECD. (2024). [Recommendation of the council on artificial intelligence](#). Again, just skim the main points.
- European Parliament. (2023). [EU AI act](#). The original text.
- Cole, M. D. (2024). [AI regulation and governance on a global scale: An overview of international, regional and national instruments](#). *Journal of AI Law and Regulation*, 1(1), 126-142. How different countries are approaching AI regulation.
- Chun, J., de Witt, C. S., & Elkins, K. (2024). [Comparative global AI regulation: Policy perspectives from the EU, China, and the US](#). *arXiv preprint*. Focused on the big three players in AI regulation.
- Feldstein, S. (2024). [Evaluating Europe's push to enact AI regulations: How will this influence global norms?](#) *Democratization*, 31(5), 1049-1066. Will the EU's approach to AI regulation become a global standard?

## 20. Lecture 20 – Privacy and data protection

Practical implications of GDPR-style rules for datasets and deployed systems.

Readings:

- European Union. (2016). [General Data Protection Regulation \(GDPR\)](#). The original text of the GDPR. Don't worry about reading the whole thing; just skim the main points.
- Hoofnagle, C. J., Van Der Sloot, B., & Borgesius, F. Z. (2019). [The European Union general data protection regulation: what it is and what it means](#). *Information & Communications Technology Law*, 28(1), 65-98. Much easier to read than the original GDPR text. Read this one instead if you want a summary.
- Wachter, S., Mittelstadt, B., & Floridi, L. (2017). [Why a right to explanation of automated decision-making does not exist in the general data protection regulation](#). *International data*

*privacy law*, 7(2), 76-99. A critical look at the “right to explanation” in GDPR.

- Ryngaert, C., & Taylor, M. (2020). [The GDPR as Global Data Protection Regulation? A \*JL Unbound\*, 114](#), 5-9. A discussion of the GDPR’s (possible) influence beyond the EU.
- Taddeo, M., & Floridi, L. (2018). [How AI can be a force for good](#). *Science*, 361(6404), 751-752. A positive perspective on how AI can be used ethically and responsibly.

## 21. Lecture 21 – Labour markets and economic effects

How automation affects jobs; augmentation strategies and policy responses.

Readings:

- Korinek, A., & Suh, D. (2024). [Scenarios for the Transition to AGI \(No. w32255\)](#). National Bureau of Economic Research. How output and wages behave under different scenarios for technological progress.
- Acemoglu, D., & Restrepo, P. (2019). [Automation and new tasks: How technology displaces and reinstates labor](#). *Journal of Economic Perspectives*, 33(2), 3-30. A comprehensive review of how automation affects jobs.
- Restrepo, P. (2025). [We won't be missed: work and growth in the era of AGI](#). NBER Chapters. Some scholars suggest we will live in an age of abundance, others that we will all be unemployed. What if the truth is somewhere in between?
- Bessen, J. E. (2019). [AI and jobs: The role of demand](#). *NBER Working Paper* No. 24235. A nuanced look at how AI affects labour markets, focusing on demand-side effects.
- World Economic Forum. (2025). [The future of jobs report 2025](#). A comprehensive report on how jobs are expected to change in the last years.
- Manyika, J., Chui, M., Miremadi, M., Bughin, J., George, K., Willmott, P., & Dewhurst, M. (2017). [A future that works: Automation, employment, and productivity](#). McKinsey Global Institute. A widely cited report on automation and jobs.

## 22. Lecture 22 – Quiz 5

Covers Lectures 19–21.

## Module 6 – Applications, limits and projects

### 23. Lecture 23 – AI in health, education and public services

Risks and benefits for public-sector applications.

Readings:

- Davenport, T., & Kalakota, R. (2019). The potential for artificial intelligence in healthcare. *Future healthcare journal*, 6(2), 94-98.
- Aung, Y. Y., Wong, D. C., & Ting, D. S. (2021). The promise of artificial intelligence: a review of the opportunities and challenges of artificial intelligence in healthcare. *British medical bulletin*, 139(1), 4-15. This article complements the previous one.
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A review of artificial intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021(1), 8812542. A comprehensive review on that topic.
- Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in Education and Research: Opportunities, concerns, and solutions. *Journal of Chemical Education*, 100 (8), 2965-2971.
- Osborne, S. P., Cucciniello, M., Nasi, G., & Zhu, E. (2022). Digital transformation, artificial intelligence and effective public services: challenges and opportunities. *Global Public Policy and Governance*, 2(4), 377-380. This is the introduction of a special issue on AI in public services. The whole issue is worth reading.

### 24. Lecture 24 – Misinformation, deepfakes and trust online

Technical features and societal responses to synthetic media.

Readings:

- Helmus, T. C. (2022). Artificial intelligence, deepfakes, and disinformation: A primer. *RAND Corporation*. Very accessible introduction to the topic.
- Mustak, M., Salminen, J., Mäntymäki, M., Rahman, A., & Dwivedi, Y. K. (2023). Deepfakes: Deceptions, mitigations, and opportunities. *Journal of Business Research*, 154, 113368. Focused on business applications.

- Chesney, R., & Citron, D. K. (2019). Deepfakes and the new disinformation war: The coming age of post-truth geopolitics. *Foreign Affairs*. A legal perspective on deepfakes and misinformation.
- Vaccari, C., & Chadwick, A. (2020). Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1), 2056305120903408. An empirical study that finds that deepfakes may contribute toward generalised indeterminacy and cynicism.
- Kietzmann, J., Lee, L. W., McCarthy, I. P., & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons*, 63(2), 135-146. The authors provide a framework to manage deepfake risks.

## 25. Lecture 25 – Course revision

Short course revision and Q&A.

## 26. Lecture 26 – Project work

In-class time for project work and consultations.