

Технології штучного інтелекту та безпека даних

Джерела для самостійного опрацювання студентами

Для кожної лекції подано 3 джерела: опрацюйте їх після лекції та перед практичною роботою. Орієнтовний час указано біля кожного джерела. Деякі ресурси оновлюються регулярно, тому звертайте увагу на найновішу версію. Якщо пряме посилання не відкривається, знайдіть матеріал за назвою на головній сторінці сайту.

Лекція 1. Основи ШІ: від ідеї до генеративних моделей

Джерело	Тип	Посилання	Що опрацювати
Elements of AI (Університет Гельсінкі)	Безкоштовний онлайн-курс	https://www.elementsofai.com	Розділи 1–2: що таке ШІ, розв’язання задач; ~3 год
Stanford HAI — AI Index Report	Щорічний аналітичний звіт	https://hai.stanford.edu/ai-index	Розділ «Top Takeaways» останнього звіту; ~30 хв
Gartner — Hype Cycle Methodology	Опис методології	https://www.gartner.com/en/methodologies/gartner-hype-cycle	Пояснення фаз кривої; ~15 хв

Лекція 2. Як працюють великі мовні моделі

Джерело	Тип	Посилання	Що опрацювати
3Blue1Brown — «Large Language Models explained briefly»	Відео (YouTube)	https://www.youtube.com/@3blue1brown	Знайдіть відео за назвою на каналі; ~8 хв
Andrej Karpathy — «Intro to Large Language Models»	Відеолекція (YouTube)	https://www.youtube.com/@AndrejKarpathy	Знайдіть відео за назвою на каналі; ~1 год
Jay Alammar — The Illustrated Transformer	Візуальна стаття	https://jalammar.github.io/illustrated-transformer/	Уся стаття; ~30 хв

Лекція 3. Обмеження ШІ: галюцинації, упередження, колапс моделей

Джерело	Тип	Посилання	Що опрацювати
Ji et al. — Survey of Hallucination in Natural Language Generation	Оглядова стаття (arXiv:2202.03629)	https://arxiv.org/abs/2202.03629	Вступ і класифікація галюцинацій; ~30 хв
Shumailov et al. — AI models collapse when trained on recursively generated data (Nature, 2024)	Наукова стаття	https://www.nature.com/articles/s41586-024-07566-y	Вступ і висновки; ~30 хв

Джерело	Тип	Посилання	Що опрацювати
Mehrabi et al. — A Survey on Bias and Fairness in Machine Learning	Оглядова стаття (arXiv:1908.09635)	https://arxiv.org/abs/1908.09635	Розділи про джерела упередженості; ~40 хв

Лекція 4. Основи промпт-інжинірингу

Джерело	Тип	Посилання	Що опрацювати
OpenAI — Prompt engineering guide	Офіційна документація	https://platform.openai.com/docs/guides/prompt-engineering	Шість стратегій і приклади; ~30 хв
Anthropic — Prompt engineering overview	Офіційна документація	https://docs.claude.com/en/docs/build-with-claude/prompt-engineering/overview	Розділи про чіткість, ролі та приклади; ~30 хв
Google — Gemini API: Prompting strategies	Офіційна документація	https://ai.google.dev/gemini-api/docs/prompting-strategies	Весь розділ; ~20 хв

Лекція 5. Шаблони промптів, Chain-of-Thought та ітерації

Джерело	Тип	Посилання	Що опрацювати
Wei et al. — Chain-of-Thought Prompting Elicits Reasoning in LLMs	Наукова стаття (arXiv:2201.11903)	https://arxiv.org/abs/2201.11903	Вступ, приклади та рис. 1; ~30 хв
Prompt Engineering Guide (DAIR.AI)	Відкритий довідник	https://www.promptingguide.ai	Розділи Techniques: Zero-shot, Few-shot, CoT; ~40 хв
Learn Prompting	Безкоштовний курс	https://learnprompting.org	Модулі Basics та Intermediate; ~2 год

Лекція 6. Коли і як делегувати завдання ШІ: матриця рішень

Джерело	Тип	Посилання	Що опрацювати
Dell'Acqua et al. — Navigating the Jagged Technological Frontier (Harvard Business School, 2023)	Дослідження	https://www.hbs.edu/faculty/Pages/item.aspx?num=64700	Резюме та висновки; ~30 хв
NIST — AI Risk Management Framework 1.0	Офіційний стандарт-рамка	https://www.nist.gov/itl/ai-risk-management-framework	Основні функції Govern, Map, Measure, Manage; ~40 хв
Amershi et al. — Guidelines for Human-AI Interaction (Microsoft Research, 2019)	Дослідження	https://www.microsoft.com/en-us/research/publication/guidelines-for-human-ai-interaction/	18 настанов; ~20 хв

Лекція 7. AI-агенти та автоматизація

Джерело	Тип	Посилання	Що опрацювати
Anthropic — Building effective agents	Технічна стаття	https://www.anthropic.com/research/building-effective-agents	Уся стаття; ~30 хв
OWASP — Top 10 for LLM Applications	Перелік ризиків	https://genai.owasp.org	Пункти про prompt injection та надмірні повноваження; ~30 хв
Simon Willison — Prompt injection series	Блог-серія	https://simonwillison.net/series/prompt-injection/	Перші 3 дописи; ~40 хв

Лекція 8. Кібербезпека та ШІ: тріада CIA, фішинг, дипфейки

Джерело	Тип	Посилання	Що опрацювати
CERT-UA	Державна команда реагування на кіберінциденти	https://cert.gov.ua	Рекомендації та оголошення про актуальні загрози; ~30 хв
ENISA — Threat Landscape	Щорічний звіт агентства ЄС	https://www.enisa.europa.eu/topics/cyber-threats/threats-and-trends	Розділи про соціальну інженерію й фішинг; ~30 хв
OWASP — Top 10 for LLM Applications	Перелік ризиків	https://genai.owasp.org	Розділ про витік чутливої інформації; ~20 хв

Лекція 9. Захист персональних даних при роботі з ШІ

Джерело	Тип	Посилання	Що опрацювати
Закон України «Про захист персональних даних»	Нормативний акт	https://zakon.rada.gov.ua/laws/show/2297-17	Статті 1–6 та 11–12; ~30 хв
Regulation (EU) 2016/679 (GDPR)	Нормативний акт	https://eur-lex.europa.eu/eli/reg/2016/679/oj	Стаття 4 (визначення) та стаття 5 (принципи); ~30 хв
EDPB — Opinion 28/2024 on AI models and personal data	Офіційний висновок	https://www.edpb.europa.eu	Знайдіть за назвою в розділі Publications; ~40 хв

Лекція 10. Регулювання ШІ: AI Act та підходи в Україні

Джерело	Тип	Посилання	Що опрацювати
Regulation (EU) 2024/1689 (AI Act)	Нормативний акт	https://eur-lex.europa.eu/eli/reg/2024/1689/oj	Статті 5 (заборонені практики) та 6 (високий ризик); ~40 хв
EU AI Act Explorer	Інтерактивний довідник	https://artificialintelligenceact.eu	Розділи Explorer та Compliance Checker; ~30 хв
Мінцифри України — Біла книга з регулювання ШІ в Україні	Політичний документ	https://thedigital.gov.ua	Знайдіть за назвою в розділі «Штучний інтелект»; ~40 хв

Лекція 11. Авторське право та відповідальність при використанні ШІ

Джерело	Тип	Посилання	Що опрацювати
U.S. Copyright Office — Copyright and Artificial Intelligence	Офіційний звіт	https://www.copyright.gov/ai/	Частина про охороноздатність (copyrightability); ~30 хв
Закон України «Про авторське право і суміжні права»	Нормативний акт	https://zakon.rada.gov.ua/laws/show/2811-20	Положення про об'єкти, створені із застосуванням ШІ; ~20 хв
UNESCO — Recommendation on the Ethics of AI	Міжнародний документ	https://www.unesco.org/en/artificial-intelligence/recommendation-ethics	Розділ «Values and principles»; ~30 хв

Лекція 12. Управління ШІ в організації: governance та shadow AI

Джерело	Тип	Посилання	Що опрацювати
ISO/IEC 42001:2023 — AI management system	Міжнародний стандарт	https://www.iso.org/standard/81230.html	Опис і структура стандарту (сторінка ISO); ~15 хв
Microsoft & LinkedIn — Work Trend Index	Щорічний звіт	https://www.microsoft.com/en-us/worklab/work-trend-index	Розділ про використання ШІ працівниками (BYOAI); ~30 хв
NIST — AI RMF Playbook	Практичний посібник	https://airc.nist.gov/AI-RMF-Knowledge-Base/Playbook	Функція Govern; ~40 хв