

Технології штучного інтелекту та безпека даних

Освітня програма «Економіка і аналітика», 1 курс · 4 кредити ECTS

Лекція 11. Організація безпечного впровадження ШІ в компанії

Мета лекції: показати, як компанія переходить від окремих правильних дій співробітників до системного, підзвітного управління використанням ШІ — через політики, ролі, навчання й моніторинг. Це синтез усього блоку безпеки курсу (лекції 7–10) на рівні організаційної практики.

Зв'язок з попередніми лекціями: лекція 10 показала індивідуальні загрози й особисті практики захисту. Ця лекція показує, що окремі правильні дії співробітників не замінюють системного підходу — потрібне формальне управління (AI governance) на рівні всієї організації.

1. Місток від індивідуальних загроз до організаційного рівня

- Лекція 10 показала індивідуальні загрози (фішинг, дипфейки) і особисті практики захисту.
- Але окремі правильні дії співробітників не замінюють системного підходу організації — без єдиної політики кожен працівник самостійно вирішує, як (і чи) безпечно користуватися ШІ, що створює непослідовний і ризикований ландшафт.
- Ця лекція розглядає, як компанія системно впроваджує безпечне використання ШІ: від політик і ролей до навчання й моніторингу.

2. Що таке AI governance і навіщо воно потрібне

- **AI governance** (управління штучним інтелектом) — сукупність політик, процесів, ролей і механізмів контролю, які визначають, як організація розробляє, впроваджує й використовує системи ШІ.
- Мета — не заборонити інновації, а забезпечити, щоб використання ШІ відповідало юридичним вимогам (лекції 7–9), знижувало ризики (лекція 10) і водночас реалізовувало бізнес-цінність (лекція 6).

Без формального governance компанія покладається на випадкові, індивідуальні рішення співробітників — що суперечить принципу підзвітності з лекції 7.

3. Проблема «тіньового ШІ» (Shadow AI)

- Співробітники користуються ШІ-інструментами на роботі незалежно від того, чи є в компанії формальна політика щодо цього — просто тому, що інструменти корисні й легкодоступні.
- **«Тіньовий ШІ» (Shadow AI)** — використання неавторизованих ШІ-інструментів для робочих задач без відома чи контролю IT-відділу й керівництва, за аналогією з давнішою проблемою «тіньового IT» (Shadow IT).
- Ризик: співробітники можуть вводити чутливі корпоративні дані в публічні ШІ-сервіси (пор. лекцію 8), не усвідомлюючи юридичних і безпекових наслідків.

Практичний висновок: просто заборонити ШІ — неефективна стратегія (працівники однаково користуватимуться, тільки прихованіше); ефективніша стратегія — дати чіткі, реалістичні правила безпечного використання.

4. Класифікація дозволеного контенту: що можна вводити в ШІ

- Пряме організаційне застосування класифікації даних з лекції 7 і правила мінімізації з лекції 8.
- Компанії варто мати чітку, доступну всім співробітникам матрицю: які категорії інформації дозволено вводити в які категорії інструментів.

Категорія даних	Дозволений інструмент
Публічна інформація	Будь-який інструмент
Внутрішня операційна інформація	Лише схвалені корпоративні інструменти
Конфіденційна й персональні дані	Лише інструменти з підписаною угодою про обробку даних (DPA) чи заборонено

Ця матриця повинна бути конкретною й застосовною за лічені секунди, а не абстрактною декларацією — інакше співробітники її ігноруватимуть.

5. Рівні дозволеного доступу до ШІ-інструментів

- Компанії можуть впроваджувати рівневу модель доступу замість універсальної політики «все дозволено» чи «все заборонено».
- **Рівень 1:** публічні безкоштовні інструменти — лише для непрацюючих із конфіденційними даними задач (загальні дослідження, ідеї, публічний контент).
- **Рівень 2:** корпоративні (enterprise) підписки — для внутрішньої операційної роботи з даними середньої чутливості.
- **Рівень 3:** спеціалізовані, аудитовані рішення (власна розробка чи суворо контрольована інтеграція) — для задач з персональними даними, фінансовими рішеннями, юридично значущим контентом.

Ця модель прямо повторює структуру вибору «готове рішення / API / власна розробка» з лекції 6, тепер сформульовану як політика доступу, а не лише економічний вибір.

6. Ролі та відповідальність в AI governance

- **Комітет з управління ШІ (AI governance committee)** — міжфункціональна група (ІТ, юридичний відділ, безпека, бізнес-підрозділи), що визначає й переглядає політику використання ШІ.
- **Відповідальний за впровадження ШІ на рівні підрозділу** — людина, відповідальна за практичне дотримання політики в конкретному відділі.
- **Кінцеві користувачі** — усі співробітники, які використовують ШІ-інструменти в щоденній роботі й повинні дотримуватися затвердженої політики.

Це організаційне віддзеркалення розмежування володілець/розпорядник/суб'єкт даних з лекції 7 — тепер застосоване до управління самим ШІ, а не лише даними.

7. Політика прийнятного використання ШІ та перевірка постачальників

- **Політика прийнятного використання (AI Acceptable Use Policy)** визначає, які інструменти дозволені, для яких задач, з якими обмеженнями, і які наслідки порушення політики.
- Типові розділи: перелік схвалених інструментів; класифікація дозволеного контенту (пор. п. 4); вимоги до перевірки згенерованого результату перед використанням (пор. лекцію 4 — «чернетка, не фінал»); процедура запиту на додавання нового інструменту.
- **Перевірка нових постачальників (vendor vetting):** перш ніж дозволити новий ШІ-інструмент, варто перевірити умови обробки даних (пор. лекцію 8), відповідність AI Аст залежно від рівня ризику (пор. лекції 8–9), гарантії безпеки й договірну відповідальність за помилки (пор. чек-лист лекції 9).

***Дидактична примітка:** Політика повинна регулярно оновлюватися — швидкість розвитку технології (пор. криву Гартнера з лекції 3) означає, що статичний документ швидко застаріває.*

8. Навчання співробітників: від разового інструктажу до системи

- Лекція 10 вже наголошувала на важливості навчання щодо кіберзагроз; на організаційному рівні це вимагає системного підходу, а не разового семінару.
- **Компонент 1 — онбординг:** базове навчання з політики використання ШІ для нових співробітників.
- **Компонент 2 — регулярне оновлення:** періодичне навчання з актуальними прикладами нових загроз і можливостей.
- **Компонент 3 — рольове навчання:** поглиблені курси для тих, хто безпосередньо працює з чутливими даними чи ухвалює рішення на основі ШІ (пор. лекцію 6 — фінансовий скоринг).

9. Моніторинг та аудит використання ШІ

- Організація повинна мати механізми перевірки того, що політика використання ШІ реально дотримується, а не існує лише на папері.
- **Технічні механізми:** журналювання використання корпоративних ШІ-інструментів (без порушення приватності самих співробітників), періодичні перевірки на предмет тіньового ШІ (п. 3).
- **Організаційні механізми:** регулярний перегляд і оновлення реєстру схвалених інструментів; аудит відповідності політиці в різних підрозділах.

Це пряме застосування принципу підзвітності з лекції 7 (організація повинна вміти довести дотримання політики, а не просто декларувати).

10. Реагування на інциденти, пов'язані з ШІ

- Компанії потрібен чіткий план дій на випадок інциденту, пов'язаного з ШІ: випадкове розкриття конфіденційних даних через чат-бот, помилкове чи

упереджене автоматизоване рішення (пор. лекцію 4 — bias), успішна атака із застосуванням дипфейка чи клонованого голосу (пор. лекцію 10).

- **Ключові елементи плану:** чіткий канал повідомлення про підозрювані інциденти; визначені відповідальні за розслідування; процедура сповіщення постраждалих осіб і регуляторів, якщо цього вимагає закон (пор. лекцію 7 — юридичний обов'язок повідомити про витік).

Важливо: культура, що заохочує повідомляти про помилки без страху покарання, значно ефективніша за культуру приховування — раннє виявлення проблеми завжди дешевше за приховану, що розростається.

11. AI Act: обов'язки організацій-впроваджувачів (deployers)

- Лекція 9 розмежувала розробника моделі, компанію-впроваджувача (deployer) і кінцевого користувача; ця лекція деталізує саме обов'язки впроваджувача.
- Для систем високого ризику (пор. лекцію 9) впроваджувач зобов'язаний: використовувати систему відповідно до інструкцій розробника; забезпечити людський нагляд на практиці, а не лише на папері; моніторити роботу системи й повідомляти розробника чи регулятора про виявлені проблеми; зберігати автоматично згенеровані журнали роботи системи.

Це означає, що навіть компанія, яка сама не розробляє ШІ, а лише використовує готове рішення, несе власний, окремий набір юридичних обов'язків.

12. Культура відповідального використання ШІ

- Політики й процеси (пп. 4–11) працюють ефективно лише в поєднанні з відповідною організаційною культурою.
- **Психологічна безпека:** співробітники повинні відчувати, що можуть поставити питання чи повідомити про помилку без страху покарання (пор. п. 10).
- **Провідний приклад «згори»:** якщо керівництво компанії демонстративно ігнорує власну політику використання ШІ, годі очікувати її дотримання від рядових співробітників.
- **Баланс інновацій і обережності:** надмірно жорстка політика штовхає співробітників у «тіньовий ШІ» (п. 3); надмірно ліберальна — створює непередбачувані ризики.

13. Як вимірювати ефективність AI governance

- AI governance — це не одноразовий проєкт, а процес, що потребує вимірюваних показників для оцінки прогресу й обґрунтування подальших інвестицій.

Показник	Що він відображає
Частка виявленого тіньового ШІ	Наскільки повно компанія розуміє реальну картину використання ШІ (пор. п. 3)
Відсоток співробітників, що пройшли навчання	Охоплення систематичним навчанням (пор. п. 8)
Середній час схвалення нового	Ефективність процесу перевірки постачальників (пор. п.

Показник	Що він відображає
інструменту	7) — надто повільний процес підштовхує до тіньового ШІ
Кількість і серйозність зафіксованих інцидентів	Реальна дієвість політики й культури повідомлення про помилки (пор. п. 10, 12)

Практичний висновок: для майбутнього економіста чи аналітика оцінка ефективності AI governance — це той самий тип задачі, що й оцінка ефективності будь-якого іншого управлінського процесу: потрібні конкретні, регулярно відстежувані показники, а не лише формальне існування документа з політикою.

14. Практичний чек-лист запуску AI governance в організації

- Провести інвентаризацію: які ШІ-інструменти вже фактично використовуються в компанії (включно з тіньовим ШІ)?
- Класифікувати дані компанії за чутливістю (пор. лекцію 7) і визначити, що можна вводити в які інструменти.
- Сформувати міжфункціональний комітет з управління ШІ.
- Розробити й затвердити політику прийнятного використання.
- Запровадити процес схвалення нових інструментів.
- Організувати системне навчання співробітників.
- Впровадити механізми моніторингу й план реагування на інциденти.

Місток до наступної лекції: ми розглянули, як компанія переходить від окремих правильних дій співробітників до системного, підзвітного управління використанням ШІ. Це фактично синтезує весь блок безпеки курсу (лекції 7–10) на рівні організаційної практики. Остання лекція курсу зведе всі викладені принципи — і зі ШІ-блоку, і з блоку безпеки — у єдиний практичний чек-лист і підготує до підсумкового контролю. Це тема лекції 12.

Контрольні питання для самоперевірки студентів

- Поясніть, чому проста заборона ШІ на робочому місці зазвичай не працює і які ризики це створює.
- Розробіть спрощену матрицю «що можна вводити в ШІ» для уявної компанії з вашої майбутньої професійної сфери.
- Опишіть, чим роль «відповідального за впровадження ШІ» в підрозділі відрізняється від ролі кінцевого користувача.
- Назвіть три компоненти, які повинна містити політика прийнятного використання ШІ.
- Поясніть, чому компанія, що лише використовує готове ШІ-рішення (а не розробляє власне), усе одно має юридичні обов'язки за AI Act.
- Чому психологічна безпека важлива для ефективного реагування на інциденти, пов'язані з ШІ?