

УДК 811.92

DOI <https://doi.org/10.31392/cult.alm.2023.1.14>

Григорак Юрій Дмитрович,
магістр філософії,
Київ, Україна
yuradovakin0@gmail.com

КОНЦЕПТ «РОБОЕТИКИ»: СТАТУС РОБОТА В КОНТЕКСТІ СУСПІЛЬНИХ ВІДНОСИН

У статті здійснюється експлікація та особливості розуміння концепту «робоетики» та місця робота в контексті суспільних відносин.

Розкрито, що в сучасній ситуації проблема взаємодії робота і людини є нагальною, оскільки розквіт робототехніки тільки починається. Однак передумови для інтенсивної інтеграції роботів до повсякденного життя людини вже існують. І завдання дослідників полягає не лише в тому, щоб спрогнозувати різке зростання розвитку даної тенденції, а й у тому, щоб окреслити коло питань і проблем, які безпосередньо впливатимуть на буття людини у майбутньому.

Доведено, що питання співвідношення етики та штучного інтелекту докорінно відрізняються від того, що розуміється, наприклад, під етичними проблемами генних технологій, інформатики, природознавства тощо. Ця відмінність визначається тим, що в штучному інтелекті етичні питання ближчі до розуміння етики у філософському чи соціо-гуманітарному сенсі, і ці етичні аспекти пов'язані передусім із тим, що вони стосуються питань поведінки та прийняття рішень.

Сформульовано визначення робоетики як галузі прикладної етики, яка має на меті розробку наукових, культурних та технічних рішень, що можуть бути загальними для різних соціальних груп та людей різних переконань. Головна проблема робоетики полягає в тому, що вчені-робототехніки, дослідники та користувачі оцінюють роботів з різних точок зору.

Обґрунтовано тезу про те, що етичні дилеми робототехніки є унікальними в тому плані, що є перенесеними зі сфери людських взаємин. Однак специфіка даної проблематики полягає в тому, що люди мають можливість закласти й формувати етичні та моральні принципи функціонування роботів. Тому вирішення якщо не всіх, але ключових етичних проблем робототехніки може бути досягнуто завдяки створенню міжнародних етичних принципів і кодексів, що буде досить важливим етапом розвитку людства, оскільки робототехніка не просто зачіпає життя суспільства, а фактично видозмінює його.

Ключові слова: робоетика, робот, етичні дилеми, соціальні відносини, прикладна етика.

Hryhorak Yuriy Dmytrovych,
Master of Philosophy,
Kyiv, Ukraine
yuradovakin0@gmail.com

THE CONCEPT OF "ROBOETHICS": STATUS ROBOT IN THE CONTEXT OF SOCIAL RELATIONS

The article explains and features the understanding of the concept of "roboethics" and the place of robots in the context of social relations.

It is revealed that in the modern situation, the problem of interaction between a robot and a person is urgent since the flowering of robotics is just beginning. However, the prerequisites for the intensive integration of robots into everyday human life already exist. And the task of the researchers is not only to predict a sharp increase in the development of this trend but also to outline the range of issues and problems that will directly affect human existence in the future.

It has been proven that the issue of the relationship between ethics and artificial intelligence is fundamentally different from what is meant, for example, by the ethical problems of gene technologies, informatics, natural science, etc. This difference is determined by the fact that in artificial intelligence, ethical issues are closer to the understanding

of ethics in a philosophical or socio-humanitarian sense, and these ethical aspects are primarily related to the fact that they relate to issues of behavior and decision-making.

The definition of roboethics is formulated as a field of applied ethics, which aims to develop scientific, cultural, and technical solutions that can be common to different social groups and people of different beliefs. A major problem in robotics is that robotics scientists, researchers, and users evaluate robots from different perspectives.

The thesis that the ethical dilemmas of robotics are unique in that they are transferred from the sphere of human relations is substantiated. However, the specificity of this issue is that people have the opportunity to lay down and form ethical and moral principles for the functioning of robots. Therefore, solving, if not all, the key ethical problems of robotics can be achieved thanks to the creation of international ethical principles and codes, which will be a rather important stage in the development of humanity, since robotics does not just affect the life of society, but actually changes it.

Key words: roboethics, robot, ethical dilemmas, social relations, applied ethics.

Актуальність дослідження пропонованої теми пов'язана з тим, що інтенсивний розвиток технологій, актуалізація міжнародних досліджень у галузі специфіки функціонування нервової системи та головного мозку людини, становлення кібернетичного спрямування у сучасній науці, сприяє активній розробці роботів.

Перш ніж розпочати безпосередній розгляд етичних проблем робототехніки, необхідно дати робоче визначення самому поняттю «робот». Наразі не існує єдиного підходу ані до того, що можна вважати роботом, ані до того, які властивості технічний засіб повинен мати, щоб бути зарахованим до цієї категорії. Опускаючи розгляд побутових, звичайних уявлень стосовно того, що є роботом, можна навести вузьке і широке визначення даного феномена.

У широкому, можна сказати, більш технологічно орієнтованому сенсі, робот – це універсальна автоматична конструкція, яка виконує механічні дії. При цьому зазначається, що дії, що виконуються такою конструкцією, подібні до людських, оскільки зразком для створення роботів служать фізичні можливості людей (Asaro, 2017).

Виходячи з цього положення, робот є лише інструментом, що опосередковує діяльність людини. Однак ми не можемо взяти за основу наведене визначення, оскільки в такому разі розгляд етичної сторони взаємодії робота та людини був би нераціональним. Вужче, або соціальне (оскільки роботу приписуються деякі людські якості) визначення робота включає такі критерії, як: здатність думати, відчувати і діяти автономно (Фурашев, 2017, с. 40).

Нині функціонує новітній науково-технічний напрямок – робототехніка, викликають порядок із проблемами технічного плану (специфіка

розробки та вдосконалення) проблеми також соціального, етичного порядку, яким незаслужено приділяється мало уваги.

Термін «робоетика» був запроваджений італійським професором Джанмарко Веруджо (Школа робототехніки Інституту електроніки, інформаційної техніки та телекомунікацій) у 2002 р. Під час Першого міжнародного симпозіуму з робототехніки в січні 2004 р. у Сан-Ремо робоетика підняла два питання, що залишаються актуальними до цього дня: чи можуть роботи виконувати хороші та погані дії; чи можуть вони бути небезпечними для людства?

Робоетика – прикладна етика, орієнтована на розробку наукових, культурних та технічних рішень, які можуть бути загальними для різних соціальних груп та людей різних переконань. Ці рішення мають бути спрямовані на стимулювання розвитку робототехніки для покращення життя як окремих індивідів, так і людського суспільства в цілому, а також на запобігання її неетичному та неправомірному використанню проти окремих людей та людства в цілому (Veruggio, 2010, с. 108).

У 2004 році провідні міжнародні інститути почали формулювати етичні норми робототехніки. У Фукуоці (Японія) під час Міжнародного ярмарку роботів у лютому 2004 р. учасники підписали Всесвітню декларацію роботів. Основні принципи даної декларації полягають у тому, що роботи наступного покоління будуть партнерами людей та допомагатимуть їм як фізично, так і психологічно. Роботи наступного покоління сприятимуть підтримці безпечного та мирного суспільства. Як бачимо, тут підкреслюються позитивні моменти співіснування людей із роботами.

У 2006 р. Дж. Веруджо була розроблена дорожня карта робоетики, в якій були визначені

етичні цінності людей (Veruggio, 2010, с. 107). Йшлося про повагу до цінностей, пов'язаних з гідністю та правами людини, сприяння рівності та справедливості у доступі до нових технологій. Цей документ має сприяти правильній оцінці користі та шкоди від використання роботів, захисту культурної різноманітності людей та плюралізму, запобіганню дискримінації та стигматизації. Тут акцентувалася увага на повазі до конфіденційності та необхідності поінформованої згоди, прийняття особистої відповідальності за біосферу.

У 2010 р. Дж. Веруджо уточнив поняття робоетики, пояснивши, що це не етика роботів і не штучна етика, а етика розробників, виробників та користувачів роботів. «Дослідження в галузі робототехніки та їх застосування все частіше призводять до етичних наслідків, що пов'язані з більш тісною взаємодією між роботами та людьми, а також із найбільш тісною взаємодією між самою робототехнікою та біологічною наукою» (Veruggio, 2010, с.105). Головна проблема робоетики полягає в тому, що вчені-робототехніки, дослідники та користувачі оцінюють роботів з різних точок зору. На думку інженера-механіка, професора Віденського технічного університету Петера Капачека, таких точок зору лише чотири:

- 1) роботи – це не що інше, як машини;
- 2) роботи мають етичні аспекти;
- 3) роботи визначаються як моральні агенти;
- 4) роботи – новий вид, що еволюціонує (Корасек, 2012, с. 70).

Тому у вивченні феномена робототехніки мають бути задіяні різні галузі знання, такі як інформатика, філософія, етика, богослов'я, біологія, фізіологія, психологія, когнітивні науки, нейронаука та юриспруденція.

Отже, робоетика – це міждисциплінарна дослідницька діяльність, що перебуває на стику етики та робототехніки та спрямована на вивчення процесу інтеграції роботів у людське суспільство. Основне завдання робоетики полягає в описі соціальних, моральних аспектів взаємодії робота та людини.

Питання співвідношення етики та штучного інтелекту докорінно відрізняються від того, що розуміється, наприклад, під етичними проблемами генних технологій, інформатики, природознавства тощо. Ця відмінність визначається тим, що в штучному інтелекті етичні питання

ближчі до розуміння етики у філософському чи соціо-гуманітарному сенсі, і ці етичні аспекти пов'язані передусім із тим, що вони стосуються питань поведінки та прийняття рішень.

Вважатимемо, що сам термін «штучний інтелект» тут і далі розуміється в метафоричному сенсі (як і термін «етика», під яким зазвичай розуміється філософська дисципліна, що досліджує питання моралі та моральності, що б під цим не малося на увазі). Зазначимо, що у цій роботі нам буде достатньо лише загального, схематичного уявлення про етику та пов'язані з нею аспекти, так само як і уявлення про те, що таке штучний інтелект.

Важливим аспектом є розгляд інтелектуальної системи не лише як когнітивної, але й як активної сутності. У цьому плані визначальною властивістю такої системи є можливість здійснення впливу на навколишній світ і насамперед соціум. Іншими словами, питання етики в штучному інтелекті зводяться до того, що ми маємо справу зі штучною системою, що реалізує процеси планування, цілепокладання, вибору та здійснення тієї чи іншої поведінки (Asaro, 2017).

При цьому вибір, здійснюваний такою системою, має визначатися певними етичними імперативами й нормами у найширшому сенсі. Наприклад, етичні норми можуть трактуватися як певні евристичні утворення, якими керується система штучного інтелекту під час здійснення вибору тієї чи іншої дії, формування системи оцінок, цільових функцій та ін.

З прагматичної точки зору дослідження в галузі етики систем штучного інтелекту призведуть в кінцевому підсумку до створення різноманітних стандартів та подальшої сертифікації систем штучного інтелекту. І тут виникають три найважливіші проблеми.

Перша стосується конструктивної формалізації етичних норм у формі, придатній для опису функціонування конкретних програмно-апаратних комплексів. Друга проблема – це здатність об'єктивного (інструментального, прямого або непрямого, заснованого на аналізі поведінки тощо) контролю відповідності компонентів системі штучного інтелекту етичним нормам. Третя – це те, який вплив справлятимуть надалі ці стандарти, й чи не виконуватимуть вони суто обмежувальну роль, що гальмує розвиток систем штучного інтелекту (Bickerstaff, 2017).

Безумовно, не можна говорити, що в сучасній ситуації проблема взаємодії робота і людини є нагальною, оскільки розквіт робототехніки тільки починається. Однак передумови для інтенсивної інтеграції роботів до повсякденного життя людини вже існують. І завдання дослідників полягає не лише в тому, щоб спрогнозувати різке зростання розвитку даної тенденції, а й у тому, щоб окреслити коло питань і проблем, які безпосередньо впливатимуть на буття людини у майбутньому.

Безумовно, наразі залишається відкритим питання щодо того, чи автоматична конструкція може мислити і відчувати подібно до людини. У даній роботі ми не будемо зосереджуватися на ньому, оскільки, можна припустити, що в найближчому майбутньому (з огляду на прогрес у галузі робототехніки) дана проблема буде вирішена більш-менш однозначно.

Однак слід звернути увагу на такий аспект, який виділяється у визначенні, як здатність робота діяти автономно, незалежно від людського чинника. Адже в такому випадку, робот із об'єкта перетворюється на суб'єкт, а враховуючи інтегративні тенденції (роботи як невід'ємні супутники людини), прийняті ним рішення, і, відповідно, чинні дії, безпосередньо зачіпатимуть фактично кожного індивіда, що неминуче призведе до актуалізації цілого ряду питань етичного плану.

Крім того, вже зараз існують проекти роботів, які самостійно приймають рішення, і, спираючись на них, виконують певні дії. До таких проектів можна віднести автономний автомобіль Google (self-driving car), який зараз перебуває на етапі активного тестування (Bickerstaff, 2017). Через це, при розгляді етичних питань таку властивість робота, як автономність, можна розглядати як ключову, що, втім, не зменшує ступінь значущості дилем, пов'язаних із використанням роботичних конструкцій як інструменту.

Зарубіжні дослідники окреслюють досить широке коло етичних проблем робототехніки. Це пов'язано, з одного боку, з тим, що на сучасному етапі науково-технічного розвитку в розробці роботів наявні проблеми інженерного (розробка штучного інтелекту, сенсорів тощо) та економічного плану (ціна та доступність).

Крім того, окремі етичні проблеми не є унікальними, а лише переносяться зі сфери

взаємодії людина-людина до сфери взаємодії людина-робот. Адже, згідно з принципами антропоморфізму, суб'єктом морального відношення може виступати не тільки людина, а й інші істоти, які мають здатність до усвідомленого вибору лінії поведінки (відповідно, суб'єктом морального відношення можуть виступати і роботи, якщо вони відповідають цій вимозі).

Пропонуємо класифікацію етичних проблем робототехніки з урахуванням виділення суб'єкта виконуваної дії та оцінки специфіки взаємодії робота та людини (Корасек, 2012):

1. Робот як об'єкт (інструмент) діяльності. На перший погляд, виділення даного критерію дещо суперечить ключовій якості робота – автономності. Однак не можна забувати про те, що в будь-якому випадку, автономність робота є відносною, оскільки він націлений на вирішення конкретних завдань (побутових, виробничих, соціальних), і в нього закладається певний функціонал (який може обмежувати чи спрямовувати його діяльність).

Проблеми етичного характеру в межах даного критерію (використання роботів людиною) можуть бути сформовані як питання: чи морально використовувати роботів у військових діях? У цьому слід зазначити, що в рамках військових операцій роботи можуть виконувати завдання різного характеру, починаючи з надання медичної допомоги (транспортування поранених), закінчуючи проведенням бойових операцій.

Відповідно, питання про етичність використання роботів у військових цілях розглядається в рамках виконання ними різних функцій. Так, з одного боку, роботи можуть опосередковано підтримувати етичні принципи, виступаючи або як об'єктивний спостерігач (фіксує події) або безпосередньо перешкоджаючи скоєнню військових злочинів (Фурашев, 2018, с. 30).

З іншого боку, беручи до уваги принцип справедливості війни (коли питання про етичність вбивства людей не порушується), надання можливості знищення людей на значній відстані (опосередковано за допомогою роботів) може призвести до інтенсифікації зловживань та порушень (емоційна дистанційність операторів). Крім того, актуальним залишається питання відповідальності (поза контекстом виконуваних роботом функцій) за ненавмисну

шкоду, завдану роботом, наприклад, внаслідок поломки або екстреного відключення живлення. Таким чином, є низка етичних проблем, що не мають наразі однозначного рішення, пов'язаних із залученням роботів у військові дії.

2. Робот як самостійно-функційний суб'єкт. Якщо відштовхуватися від позиції про те, що в майбутньому роботи володітимуть досить високим ступенем автономності, незалежністю від людського чинника, раціонально поставити етичне питання про те, чи слід надати роботам права і в якому обсязі.

Чи мають права, які надаються роботам, ґрунтуватися на правах людини? Безумовно, сама постановка подібного питання має на увазі з одного боку досягнення такого рівня технічного розвитку, коли роботи стануть невід'ємною частиною суспільства і будуть мати свідомість, а з іншого боку перестануть сприйматися людьми виключно як об'єкти.

Якщо вже на світовому рівні починають розроблятися принципи етичної поведінки роботів, слід очікувати, що з досягненням певного рівня розвитку робототехнічної сфери будуть розроблені закони про захист прав роботів (Moskalchuk & Turenko, 2021). Однак якщо роботам будуть надані права, постає питання, чи це не суперечитиме інтересам людей, оскільки основна мета створення роботів – служіння людині.

У сучасній ситуації однозначних відповідей на поставлені питання не існує. Крім того, якщо робот є самостійним суб'єктом, що приймає рішення, від яких залежать здоров'я та життя людини, виникає потреба проектувати його так, щоб він діяв етично.

Але яким конкретним етичним принципам має підпорядковуватися робот? Ця проблема актуалізувалася завдяки розробці та впровадженню автономних автомобілів, отримавши назву «проблема тунелю» (tunnel problem) (Харитонов & Харитонova, 2018, с. 43). Сенс її полягає в тому, що робот може опинитися в ситуації, коли необхідно буде зробити вибір – врятувати життя свого господаря, знищивши іншу людину, або врятувати життя іншої людини, знищивши господаря.

Коли говорять про етичні проблеми робототехніки, фактично завжди звертаються до трьох законів робототехніки, сформованих А. Азімовим. Адже, згідно з першим і головним законом,

робот не повинен завдавати шкоди людині або допускати, щоб людині було заподіяно шкоду. У наведеній ситуації дії робота в будь-якому випадку будуть неетичними, оскільки призведуть до смерті людини.

Ключовим питанням є визначення відповідності системи штучного інтелекту етичним нормам. Припустимо, що є, з одного боку, певна формальна система норм, шкал, принципів, з іншого – система штучного інтелекту, яка має розглядатися як «чорний» чи «сірий ящик». Наявність доступного програмного коду, алгоритмів, математичних моделей, що лежать в основі системи штучного інтелекту, жодним чином не зможе вирішити задачу визначення «ступеня етичності» цієї системи, тому що тут ми стикаємося з проблемами типу алгоритмічної верифікації.

При цьому важливими є такі аспекти.

Перший із них полягає в тому, що описані тестові ситуації, які вимагають вибору дії (ухвалення рішення), повинні бути не просто певними виділеними сценами, а мати зв'язок із контекстом ситуації у найширшому значенні цього терміну. Завдання контексту – визначення наслідків вибору дії, оцінка відповідності результатів вибору цієї дії тим чи іншим етичним імперативам. Чи має формуватися максимально повна картина чи модель світу, які механізми її реалізації, які протоколи подання ситуацій, системи оцінок тощо – це вже питання для окремого розгляду. Тут ми лише підкреслюємо, що система тестів повинна містити безліч поданих на вхід до штучного інтелекту ситуацій, що вимагають того чи іншого вибору, зумовленого евристико-етичними імперативами.

Другий аспект стосується питання, чи повинна система штучного інтелекту мати «пояснювальний компонент», тобто чи потрібна наявність механізму, що показує ланцюжок міркувань, що привела систему штучного інтелекту до того чи іншого рішення. З одного боку, здається, що таке пояснення дозволить краще визначити, наскільки система штучного інтелекту керувалась етичними засадами. Однак:

- по-перше, верифікація на етичність всього ланцюжка міркувань може виявитися дуже складним завданням.

- по-друге, послідовність дій, кожна з яких етично обумовлена, може призвести до «неетичного» результату. Це неминуче постає як

наслідок важливої неповноти та суперечливості етичних норм.

- по-третє, далеко не всі системи штучного інтелекту можуть мати пояснювальну компоненту через їхню природу. Прикладом цього можуть бути ті самі нейронні системи.

Таким чином, єдиною формою визначення ступеня етичності системи штучного інтелекту є тестування, яке за вхідними ситуаціями оцінює прийняте системою штучного інтелекту рішення. Основні етичні проблеми або навіть етичні дилеми робототехніки є унікальними

в тому плані, що є перенесеними зі сфери людських взаємин. Однак специфіка даної проблематики полягає в тому, що люди мають можливість закладати й формувати етичні та моральні принципи функціонування роботів. Тому вирішення якщо не всіх, але ключових етичних проблем робототехніки може бути досягнуто завдяки створенню міжнародних етичних принципів і кодексів, що буде досить важливим етапом розвитку людства, оскільки робототехніка не просто зачіпає життя суспільства, а фактично видозмінює його.

Список використаних джерел:

1. Фурашев, В.М. (2017). Інтернет речей і право, *Матеріали наук.-практ. конф. Інтернет речей: проблеми правового регулювання та впровадження*, м. Київ, 24 жовтня 2017 р., НТУУ “КПІ ім. Ігоря Сікорського”. Київ: Вид-во “Політехніка”, 39–43.
2. Фурашев, В.М. (2018). Право у світлі технології Інтернет-речей, *Матеріали другої наук.-практ. конф. Інтернет речей: проблеми правового регулювання та впровадження*. м. Київ, 29 листопада 2018 р., НТУУ “КПІ ім. Ігоря Сікорського”. Київ: Вид-во “Політехніка”, 29–32.
3. Харитонов, Є.О., Харитонova, О.І. (2018). До проблеми цивільної правосуб’єктності роботів, *Матеріали наук.-практ. конф. Інтернет речей: проблеми правового регулювання та впровадження*, м. Київ, 29 листопада 2018 р., НТУУ “КПІ ім. Ігоря Сікорського”. Київ: Вид-во “Політехніка”, 42–46.
4. Asaro, P.M. (2017). Robots and Responsibility from a Legal Perspective. <http://www.peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf>
5. Bickerstaff, R. (2017). Do we need robot law? <https://www.britac.ac.uk/events/do-we-need-robot-law>
6. Kopacek, P. (2012). Roboethics. *Proceedings of the IFAC Workshop on „Supplemental Ways for Improving International Stability”*, 67–72.
7. Moskalchuk, M., Turenko, V., & Yarmolitska, N. (2021). Tolerance in the Field of Education: Modern Challenges and New Paradigms, *Beytulhikme An International Journal of Philosophy*, 11, 1, 241–254.
8. Veruggio, G. (2010). Roboethics. *Robotics & Automation Magazine, IEEE*. Vol. 17, Issue 2, 105–109.

References:

1. Furashev, V.M. (2017). Internet rechey i pravo, *Materialy nauk.-prakt. konf. Internet rechey: problem pravovogo reguluyvannya ta vprovadzhennya*, m. Kyiv, 24 zhovtynya 2017 r., NTUU “KPI im. Igorya Sikorskogo”. Kyiv: Vyd-vo “Politehnika”, 39–43.
2. Furashev, V.M. (2018). Pravo u svitli tehnologiyi Internet-rechey, *Materialy drugoyi nauk.-prakt. konf. Internet rechey: problemy pravovogo reguluyvannya ta vprovadzhennya*. m. Kyiv, 29 lystopada 2018 r., NTUU “KPI im. Igorya Sikorskogo”. Kyiv: Vyd-vo “Politehnika”, 29–32.
3. Haritonov, E.O., Haritonova, O.I. (2018). Do problemy tsivilnoi pravosub’ektnosti robotiv, *Materialy nauk.-prakt. konf. internet rechey: problemi pravovogo reguluyvannya ta vprovadzhennya*, m. Kyiv, 29 lystopada 2018 r., NTUU “KPI im. Igorya Sikorskogo”. Kyiv: Vyd-vo “Politehnika”, 42–46.
4. Asaro, P.M. (2017). Robots and Responsibility from a Legal Perspective. <http://www.peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf>
5. Bickerstaff, R. (2017). Do we need robot law? <https://www.britac.ac.uk/events/do-we-need-robot-law>
6. Kopacek, P. (2012). Roboethics, *Proceedings of the IFAC Workshop on “Supplemental Ways for Improving International Stability”*, 67–72.
7. Moskalchuk, M., Turenko, V., & YARMOLITSKA, N. (2021). Tolerance in the Field of Education: Modern Challenges and New Paradigms, *Beytulhikme An International Journal of Philosophy*, 11, 1, 241–254.
8. Veruggio, G. (2010). Roboethics. *Robotics & Automation Magazine, IEEE*. Vol. 17, Issue 2, 105–109.