УДК 681.3:314.1:004.6

БРАЙЧЕВСЬКИЙ С.М., *кандидат фізико-математичних наук*

ПРОБЛЕМА ПЕРСОНАЛЬНИХ ДАНИХ В СИСТЕМАХ ІНТЕРНЕТУ РЕЧЕЙ З ЕЛЕМЕНТАМИ ШТУЧНОГО ІНТЕЛЕКТУ

Анотація. *В роботі розглядаються можливі механізми неконтрольованої генерації наборів персональних даних системами Інтернету речей з елементами штучного інтелекту*

Ключеві слова:*інформаційні технології, Інтернет речей, персональні дані.*

Summary. *Possible mechanisms of uncontrolled generation of sets of personal data by systems of the Internet of things with elements of artificial intelligence are considered in the work*

Keywords: *information technology, Internet of Things, personal data.*

Аннотация. *В работе рассматриваются возможные механизмы неконтролируемой генерации наборов персональных данных системами Интернета вещей с элементами искусственного интеллекта*

Ключевые слова: *информационные технологии, Интернет вещей, персональные данные.*

Постановка проблеми. Швидкий розвиток сучасних інформаційних технологій породжує нові виклики і нові ризики, що потребують ретельного вивчення. До їх числа відносяться і проблеми правового регулювання, пов'язані з використанням Інтернету речей (далі – ІР) [1 – 6]. Вони пов'язані з наявністю (принаймні, гіпотетичною) в поведінці систем ІР-елементів соціальної поведінки [2]. Питання про природу соціальних відносин між людиною та технологічною системою є, взагалі кажучи, досить нетривіальне. В пропонованій роботі ми не маємо наміру обговорювати цю проблему в повному обсязі.

Однією з проблем, які активно обговорюються у зв'язку з розвитком ІР, є захист персональних даних [7; 8]. Причина полягає перш за все в тому, що системи ІР за своєю природою призначені для збирання різноманітних даних, причому відповідно до певних алгоритмів, які не завжди відповідають загальноприйнятим нормам оперування конфіденційними відомостями. Важливо, що значна частина ризиків, що виникають, взагалі не пов'язані з штатними режимами експлуатації систем ІР. Дійсно, кібернетична система може оперувати даними, "не усвідомлюючи", що вони означають чи можуть означати в суб'єктивному сприйнятті людиною. Машина використовує дані з певною метою, тоді як хтось може використати ці ж самі дані зовсім з іншою метою.

В наявній літературі загалом обговорюються ситуації, пов'язані з безпосереднім отриманням даних за допомогою датчиків ІР та їх можливе несанкціоноване розповсюдження шляхом використання мережних технологій. Але мають бути розглянуті й складніші ситуації, що можуть виникати в процесі експлуатації систем, які містять елементи штучного інтелекту. В таких ситуаціях машина може оперувати даними, які вона самостійно збирає та опрацьовує в процесі вирішення задач, що виходять за межі лінійної обробки інформації, типової для "звичайних" кібернетичних систем.

В пропонованій роботі ми проаналізуємо принципову здатність систем з елементами штучного інтелекту в процесі експлуатації самостійно модифікувати алгоритми, закладені в них проектувальниками. А це означає, що вони можуть неконтрольовано генерувати непередбачені набори даних, які за своєю природою мають

бути віднесені до категорії персональних даних. Результатом може бути створення якісно нових комплексів персональних даних, які відсутні в інших наявних джерелах. Ми покажемо, що такі набори даних можуть формуватися за рахунок автоматичної побудови системи зв'язків між "стандартними" персональними даними.

Результати аналізу наукових публікацій. Правове регулювання в галузі використання технологічних (в тому числі інформаційних) систем саме по собі не є чимось новим. Мається на увазі правове регулювання відносин між людьми, які здійснюються за допомогою технологічних систем або у зв'язку з їх використанням.

При цьому виділяють дві основні категорії проблем:

- особливості функціонування технологічних систем як причина виникнення особливостей у додатковому правовому регулюванні;
- забезпечення захисту від наслідків нештатного функціонування технологічних систем.

Тобто суб'єктом права в будь-якому випадку є людина, а технологічна система виступає лише в ролі знаряддя в її руках. Отже, в ситуаціях, коли функціонування системи призводило до негативних наслідків, вважалося, що відповідальність за її дії несуть розробники, виробники та експлуатаційники, тобто люди.

Але сьогодні (принаймні, теоретично) розглядаються ситуації, в яких відповідальність може бути покладена саме на машину, незалежно від участі людини [2; 3]. Такий погляд на технологічні системи є принципово новим, оскільки передбачає можливість того, що їх функціонування може мати соціальні наслідки, а отже, вони самі можуть розглядатися як суб'єкти суспільних відносин. Фактично, сказане означає, що за певних умов технологічна система набуває елементів суб'єктності. На перший погляд, це суперечить загальноприйнятим уявленням про сутність технологічних систем. Адже вважається, що машина лише виконує програму, закладену в неї людиною. І разом з тим, розвиток сучасних інформаційних технологій, зокрема Інтернету речей, свідчить, що такі ситуації можливі. Якщо не вдаватися до наукової фантастики, то мова, очевидно, йде не про повноцінну суб'єктність машини, а про наявність в її функціонуванні окремих рис, характерних для справжнього суб'єкта – людини.

Вважаємо, що в рамках обраної нами теми ключовим чинником є здатність машини самостійно приймати рішення. Підкреслимо, що йдеться не про імітацію прийняття рішення, що, взагалі кажучи, на наш час не є чимось особливим (прикладом може служити комп'ютер, що грає в шахи). Ми маємо на увазі здатність машини приймати рішення, яке однозначно не визначається алгоритмом, обраними значеннями його параметрів та структурою вхідних даних. Саме така поведінка машини дає підстави говорити про її відповідальність за власні дії, що є предметом правового регулювання.

Загрози та ризики, що виникають в сфері використання IP, широко обговорюються в експертному середовищі. Стислий виклад поточного стану речей міститься, наприклад, в звітах групи Alliance for Internet of Things Innovation, (AIPI), створеної 2015 року у Європейській Комісії [9]:

- існуюча нормативно-правова база і регуляторні рамки, в основному, відповідають вимогам сучасного цифрового середовища;
- ключ до розвитку IP полягає у встановленні балансу між гарантуванням безпеки споживачів і стимулюванням інновацій;
- частина ризиків пов'язана з відповідальністю за якість продукції, якій надається особливе значення, хоча вона й застосовує IP, але це не є чимось унікальним для цієї продукції і платформ;

- виникають питання, пов’язані з відмінностями в поняттях “продукт” і “сервіс”, тому необхідні чіткі роз’яснення, щоб уникнути невизначеності;
- забезпечити такий розвиток регуляторної політики, щоб вона була досить гнучкою для можливості врахування схильності промисловості до постійного розвитку, що є для неї ключовим.

Окрему категорію становлять ризики, пов’язані з проблемою захисту персональних даних [7; 8; 10; 11]. ІР за своєю природою орієнтований на збирання великих обсягів даних. Серед них можуть бути і дані, які слід кваліфікувати як персональні.

Важливою є особливість систем ІР, яка полягає в тому, що активне використання великої кількості датчиків створює умови для формування комплексів даних, в тому числі і персональних [12].

Основні аспекти сучасної проблеми захисту персональних даних містяться, наприклад, в матеріалах звіту Федеральної торгової палати США [13]:

- переваги впровадження ІР зводяться до мінімуму наявністю негативних наслідків, наприклад, загрозами конфіденційності персональних даних;
- зайве регулювання в питаннях захисту персональних даних може призвести до уповільнення інвестицій в будь-який сектор;
- прийняття необхідного регулювання для гарантованого захисту персональних даних підвищить довіру споживачів до нових технологій;
- необхідно дочекатися проявів негативних наслідків і, тільки після цього, вживати заходів з регулювання;
- доцільно використовувати механізми саморегулювання замість регулювання законодавчими нормами.

Аналіз широкого кола джерел свідчить про те, що останнім часом проблема захисту персональних даних у використанні систем ІР активно переходить в сферу прийняття безпосередньо правових рішень [7; 8].

Метою статті є вивчення можливих механізмів генерації системами ІР з елементами штучного інтелекту наборів персональних даних, заснованих на використанні автоматично модифікованих алгоритмів..

Нижче ми проаналізуємо один із аспектів проблеми несанкціонованого поширення персональних даних системами ІР. А саме, принципову можливість системи ІР генерувати принципово нові набори персональних даних, засновану на використанні алгоритмів, що здійснюють агрегування вхідної інформації..

Виклад основного матеріалу. Перш за все зазначимо, що на наш час саме поняття персональних даних зазнало певного розширення в порівнянні з традиційним розумінням їх як “паспортні дані”. Відповідно до Загального регламенту про захист даних (GDPR), діючого в межах законодавства Європейського Союзу щодо захисту персональних даних, це поняття визначається як “…будь-яка інформація, що стосується фізичної особи, яку ідентифіковано чи можна ідентифікувати” [14]. Аналогічно це поняття визначається і Законом України “Про захист персональних даних”.

Для нас в цьому визначенні важливі два моменти:

- персональними даними може бути будь-яка інформація;
- визначальним чинником є ідентифікованість відповідної особи, або принципова можливість такої ідентифікації.

Прийнято вважати, що персональні дані належать до одного з таких видів даних:

- літери;
- числа;

- графічні зображення (малюнки або картини);
- фото;
- аудіо;
- відео.

Також останнім часом до персональних відносять такі специфічні дані:

- файли cookies;
- IP-адреси.

Таким чином, персональні дані в сучасному розумінні мають досить широкий спектр.

Головна особливість маніпуляції персональними даними в системах ІР полягає в тому, що її здійснює машина, яка, взагалі кажучи, "не знає" який сенс мають ті або інші дані з точки зору людини. Саме ця особливість породжує специфічні ризики, зумовлені тим, що людині надзвичайно важко контролювати такі аспекти функціонування кібернетичних пристроїв.

На рівні технологічної реалізації ІР є набором датчиків, що фіксують задані параметри навколишнього середовища, та пристроїв, що обробляють вхідні дані, отримані від датчиків. Для нас суттєво, що обмін даними здійснюється за допомогою мережі Інтернет. Метою створення такої системи є виключення безпосередньої участі людини принаймні в частині функціональних можливостей системи. Це, в свою чергу означає, що система ІР повинна на основі обробки отриманих вхідних даних приймати рішення, результатом яких буде отримання додаткових даних. Ці додаткові дані можуть мати різні джерела, які більш чи менш строго розподіляються на дві групи:

- дані датчиків, які входять до складу відповідної системи ІР;
- дані, що знаходяться в мережі Інтернет, до якого дана система ІР має доступ.

Саме доступність даних другої групи може створювати складні неконтрольовані ситуації. Адже проектувальник системи не може передбачити, запит на які дані сформує машина в певній ситуації, навіть, якщо сама ситуація прогнозована.

Отримання машиною додаткових даних гіпотетично є актуальним для систем з елементами штучного інтелекту [1].

Поняття "штучний інтелект" є надзвичайно популярним і, разом з тим, доволі погано визначеним. В літературі з ним пов'язано багато різноманітних спекуляцій від технічних непорозумінь до відвертих фантазій в дусі футуризму та наукової фантастики. В дійсності існує дві основні точки зору на поняття "штучний інтелект":

- технологія створення обчислювальних машин, здатних вирішувати завдання, що традиційно вважаються інтелектуальними [15];
- властивість обчислювальних машин вирішувати такі завдання [16].

Головна складність полягає в тому, щоб строго визначити, які задачі слід вважати інтелектуальними. Зазвичай кажуть, що це задачі, що вимагають виконання творчих функцій, але саме поняття творчості також потребує строгого визначення. Яскравим прикладом можуть служити шахи. Чи слід віднести комп'ютерні програми гри в шахи до реалізації штучного інтелекту? З одного боку, шахи вважаються інтелектуальною грою, а з іншого – такі програми принципово не відрізняються від програм, скажімо, аналітичного розв'язання рівнянь. Спрощено кажучи, вони просто перебирають всі можливі ходи із заданою глибиною (n-ходів вперед) і кожному з них за допомогою деякого алгоритму привласнюють ваговий множник. Перевага віддається ходу з максимальною вагою. Ефективність програми визначається досконалістю алгоритму обчислення ваг, який розробляє людина, а не власне роботою машини.

В рамках пропонованого дослідження обмежимося однією з можливих реалізацій системи, поведінка якої може в розумному наближенні вважатися інтелектуальною. В основі її (реалізації) лежить поділ систем на лінійні і нелінійні. Поняття лінійності ми використовуємо в досить широкому сенсі. Саме, під лінійною ми будемо розуміти таку систему, для якої нескінченно малі відхилення вхідного сигналу призводять до нескінченно малих відхилень вихідного сигналу. Відповідно, нескінченно мале відхилення вхідного сигналу нелінійної системи призводить до кінцевого відхилення вихідного сигналу. А це означає, що при нескінченно малих (і тому непомітних для нас) збуреннях даних, які машина так чи інакше отримує на вході, вона може виконувати дії, що не збігаються з тими, які повинні відбутися при незбурених значеннях цих даних. І тут мова не йде про імітацію інтелектуальної діяльності: машина дійсно веде себе самостійно з нашої точки зору, оскільки кінцевий результат не належить до наперед заданого набору можливих варіантів.

Таким чином, під штучним інтелектом ми розумітимемо технологію створення обчислювального комплексу, що представляє собою деяку нелінійну систему. Відзначимо, що мова йде про нелінійність всього комплексу, а не тільки програми, оскільки істотну роль можуть грати механізми отримання вхідних даних.

Безпосередньо нас цікавить ситуація, в якій машина використовує персональні дані, не передбачені при її створенні. В "звичайних" кібернетичних системах такі ситуації не виникають. Кожний конкретний програмно-апаратний комплекс від початку призначений для обробки певного набору даних, серед яких можуть бути і персональні. Процеси несанкціонованого збирання та поширення персональних даних є доволі простими і зрозумілими. Ми розуміємо їх причини і механізми. Проблема полягає лише в тому, щоб віднайти адекватні засоби відповідних дій.

Зовсім інший стан справ виникає тоді, коли машина здатна сама генерувати персональні дані, використовуючи інші дані, на перший погляд такі, що не мають відношення до персональних. При цьому, очевидно, так чи інакше машина повинна використовувати додаткові дані, які вона збирає самостійно, використовуючи алгоритми власного виробництва (наприклад, в рамках можливості самонавчання). Джерелами таких даних, звичайно, можуть бути і власні датчики системи ІР, і доступні для неї ресурси мережі Інтернет.

З точки зору проектувальників, відповідна система ІР може взагалі не оперувати персональними даними, або оперувати ними в обмежених рамках. Але ми вже казали, що до персональних даних можуть бути віднесені будь-які відомості, так чи інакше пов'язані з тією чи іншою особою. І вони можуть складатися з кількох компонентів. Частина з них передбачена штатним режимом експлуатації системи, а частина – збирається і обробляється машиною в рамках використання модифікованих алгоритмів.

Наведемо умовний (гіпотетичний) приклад, який ілюструє сказане вище. Нехай маємо систему ІР категорії "розумний будинок" (мається на увазі технологія керування різноманітними побутовими приладами, а також здійснення різноманітних видів віддаленого моніторингу, дані яких визначають прийняття рішень в конкретних ситуаціях). Шляхом вдосконалення алгоритмів, що керують засобами охорони будинку, система фіксує людину, яка наближається на критичну відстань, ідентифікує її за обличчям, а потім здійснює пошук в доступних базах на предмет реєстрації її як терориста, педофіла тощо. При цьому машина може завантажувати вміст баз даних, які містять дані на цю людину, створюючи її власний профіль. Оскільки такі дані машина використовує сама, ми можемо не мати про це жодного уявлення. Але за певних умов ці

дані можуть бути кимось використані або відповідно до характеру роботи машини, або внаслідок зламу системи сторонніми особами.

Такі ситуації породжують додаткові загрози в плані захисту персональних даних, важливість яких в першу чергу обумовлена принциповою неконтрольованістю наборів даних, якими фактично маніпулює машина.

Висновки.

Отже, ми бачимо, що за певних умов характер взаємодії ІР з оточуючим середовищем може призводити до нелінійних ефектів з елементами непередбачуваної поведінки системи. Один із можливих випадків пов'язаний з використанням машиною непередбачених наборів даних, серед яких можуть бути присутні і персональні дані.

Ми бачимо, що системи ІР з елементами штучного інтелекту в процесі експлуатації в принципі здатні розширювати штатний режим отримання та обробки даних, внаслідок чого машина стає здатна самостійно генерувати персональні дані, використовуючи інші дані, отримані в передбачений спосіб. При цьому машина так чи інакше повинна використовувати додаткові дані, які вона збирає самостійно, використовуючи модифіковані алгоритми (наприклад, в рамках можливості самонавчання). Джерелами таких даних, звичайно, можуть бути і власні датчики системи ІР, і доступні для неї ресурси мережі Інтернет.

В результаті виникають додаткові загрози в плані захисту персональних даних, важливість яких в першу чергу обумовлена принциповою неконтрольованістю наборів даних, якими фактично маніпулює машина. Отже, виникає необхідність врахування таких загроз при розробці норм законодавства щодо захисту персональних даних, а також адекватних механізмів реалізації цих норм на практиці.

Використана література

1. Баранов А.А. Интернет вещей и искусственный интеллект: истоки проблемы правового регулирования: збірник матеріалів ІІ-ї Міжнародної науково-практичної конференції *"ІТ-право: проблеми та перспективи розвитку в Україні"*, м. Львів, 17 листопада 2017 р. Львів: НУ "Львівська політехніка", 2017. 318 с. С. 18-42.

2. Рекомендации МСЭ-Т Y.2060 (06/2012). Серия Y: Глобальная информационная инфраструктура, аспекты протокола Интернет и сети последующих поколений. Сети последующих поколений. – Структура и функциональные модели архитектуры. Обзор Интернета вещей. URL: http://handle.itu.int/11.1002/1000/11559

3. Баранов О.А. "Інтернет речей" як правовий термін. *Юридична Україна*. 2016. № 5-6. С. 96-103. URL: http://www.irbis-nbuv.gov.ua/cgi-bin/irbis_nbuv/cgiirbis_64.exe?C21C OM=2&I21 DBN=UJRN&P21DBN=UJRN&IMAGE_FILE_DOWNLOAD=1&Image_file_name=PDF/urykr _2016_5-6_16.pdf

4. Леонид Черняк. Платформа Интернета вещей (рус.). *Открытые системы. СУБД*. 2012. № 7, URL: https://www.osp.ru/os/2012/07/13017643

5. Kevin Ashton. That 'Internet of Things'. In the real world, things matter more than ideas. (англ.). RFID Journal (22 June 2009) http://www.rfidjournal.com/articles/view?4986

6. 'Internet of Things' (англ.). Gartner IT glossary. Gartner (5 May 2012). – 'The Internet of Things is the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment'. URL: https://www.gartner.com/it-glossary/internet-of-things

7. Баранов О.А., Брижко В.М. Захист персональних даних в сфері Інтернет речей. *Інформація і право*. № 2(17)/2016. С. 85-91. URL: http://ippi.org.ua/sites/default/files/11_0.pdf

8. Брижка В.М. , Пилипчук В.Г. та ін. Становлення і розвиток правових основ та системи захисту персональних даних в Україні: монографія. Київ: ТОВ "Видавничий дім "АртЕк", 2017. 226 с.

9. Charlie Hawes. Hogan Lovells assists Internet of Things policy group in Brussels, 28 October 2015. URL: http://www.hlmediacomms.com/2015/10/28/hogan-lovells-assists-internet-of-things-policy-group-in-brussels

10. Интернет вещей: чем угрожает будущее. URL: http://igate.com.ua/news/3169-internet-veshhej-chem-ugrozhaet-budushhee

11. Как в 2015 году был взломан Интернет вещей. URL: http://igate.com.ua/ news/12342-kak-v-2015-godu-byl-vzloman-internet-veshhej

12. Recommendation CM/Rec(2010)13 of the Committee of Ministers to member states on the protectionof individuals with regard to automatic processing of personal data in the context of profiling (Adopted by theCommittee of Ministers on 23 November 2010 at the 1099th meeting of the Ministers' Deputies). URL: https://search.coe.int/cm/Pages/result_details.aspx? ObjectID=09000016805cdd00

13. Internet of Things: Privacy & Security in a Connected. World Federal Trade Commission (FTC) Staff Report. January 2015. URL: https://www.ftc.gov/system/files/documents /reports/ federal-trade-commission-staff-report-november-2013-workshop-entitled-internet-things-privacy/150127IPrpt.pdf

14. Регламент (ЄС) 2016/679 від 27.04.16 р. "Про захист фізичних осіб у зв'язку з обробкою персональних даних та про вільне переміщення таких даних, а також про скасування Директиви 95/46/ЄС (Загальний Регламент про захист даних)". URL: https://gdpr-text.com/?col =2&lang1 =ukr&lang2=en&lang3=rumain. – (Переклад Регламенту та ін. правових стандартів ЄС надано у кн.: *Сучасні правові стандарти Євросоюзу у сфері захисту персональних даних* / [І. Майстренко – пер. з англ.; В. Брижко – ред. тексту]. – (Науково-дослідний інститут інформатики і права Національної академії правових наук України). Київ: ТОВ "Видавничий дім "АртЕк", 2018. 177 с.).

15. What is Artificial Intelligence? FAQ от Джона Маккарти, 2007. URL: http://www-formal.stanford.edu/jmc/whatisai/whatisai.html

16. Аверкин А.Н., Гаазе-Рапопорт М.Г., Поспелов Д.А. Толковый словарь по искусственному интеллекту. Москва: Радио и связь, 1992. 256 с. URL:http://www.raai.org/library/tolk/aivoc.html#L208

~~~~~~~~~~~~~~ * * * ~~~~~~~~~~~~~~
~~~~~~~~~~~~~~