COMPUTER SYSTEMS AND NETWORKS

Vol. 6, No. 1, 2024

DOI: https://doi.org/10.23939/csn2024.01.036

УДК 004.056.5

МЕТОДОЛОГІЯ ЗБОРУ, ОБРОБКИ, ЗБЕРІГАННЯ ТА КЛАСИФІКАЦІЇ ДАНИХ ВІДПОВІДНО ДО ВИМОГ SOC 2 TYPE 2

О. Р. Дейнека, Л. Л. Бортнік

Національний університет "Львівська політехніка",
кафедра захисту інформації
E-mail: oleh.r.deineka@lpnu.ua, leonid.l.bortnik@lpnu.ua,

У сучасному світі швидкого зростання обсягів інформації виникають виклики з її класифікації, зберігання, передання та захисту. Фахівці з кібербезпеки розробляють стандарти, як-от ISO 27001 та SOC 2, для контролю доступу та забезпечення конфіденційності даних. Звіт SOC 2 Type 2, як стандарт оцінки контролів безпеки, підвищує довіру до організацій, надаючи підтвердження ефективності їхніх заходів безпеки від незалежних авдиторів. Відповідність цьому стандарту дає можливість компаніям не лише уникати штрафів за невідповідність законодавчим вимогам, а й виявляти та усувати вразливості в системах безпеки. Отримання звіту SOC 2 Type 2 сприяє вдосконаленню внутрішніх процедур та підвищенню розуміння ризиків і контролів серед співробітників.

Початковим та критично важливим кроком у формуванні надійної стратегії безпеки даних є класифікація даних, яка допомагає організаціям розпізнавати дані та призначати рівень чутливості, що керує відповідними заходами безпеки.

Запропонована в роботі методика збору, обробки, класифікації та зберігання даних зорієнтована на відповідність SOC 2 Type 2, містить розгляд різних типів даних та метаданих, а також застосування інструментів інтеграції для ефективного управління даними. Процес описує створення моделі даних, класифікацію даних та посилання їх на метадані, що забезпечує безпеку та відповідність нормативним вимогам.

Запропонований підхід надає гнучкість у виборі технологій та складу команди, адаптуючи процес під потреби проєкту для досягнення високої ефективності управління даними. Таким способом компанії можуть забезпечити безпеку, доступність та відповідність даних до сучасних стандартів, підвищуючи свою конкурентоспроможність та довіру клієнтів.

Ключові слова: SOC 2 Type 2, стандарти зберігання, класифікація даних, зберігання даних, безпека даних.

Вступ

У сучасному світі, де дані стають все більш цінним ресурсом, питання їх ефективного збору, обробки, класифікації та зберігання набувають особливої актуальності. Від того, наскільки правильно і безпечно ці процеси організовані, залежить не лише продуктивність роботи компанії, але й захист її комерційної таємниці, конфіденційної інформації та персональних даних клієнтів.

Одним з основних інструментів забезпечення безпеки даних є стандарт SOC 2 Type 2, розроблений Американським інститутом сертифікованих бухгалтерів. Цей стандарт встановлює вимоги до процесів обробки інформації, зокрема, з питань її зберігання, доступу до неї та контролю за її використанням.

Сподіваємося, що результати цього дослідження будуть корисними для фахівців у галузі інформаційної безпеки, а також для керівників організацій, які прагнуть поліпшити процеси збору, обробки, класифікації та зберігання даних у своїх компаніях.

1. Огляд літературних джерел

Сучасний світ характеризується швидким зростанням обсягів інформації, яка містить велику кількість критичної інформації. Великі обсяги такої інформації насамперед потребують класифікації за різними параметрами та особливостями, їх надійного зберігання та передання, а також захисту від несанкціонованого доступу. Останнім часом кількість можливих атак на інформаційні ресурси постійно зростає [1–3]. Спеціалісти з кібербезпеки постійно розробляють нові стандарти, підходи та методи протидії таким зловмисним діям, а також відбувається розвиток інфраструктури в цьому напрямку [4–9]. Важливим напрямом є розробка стандартів безпечного зберігання даних. Стандарти безпеки дають можливість краще зрозуміти, як саме установа контролює доступ до даних і забезпечує їх безпеку та конфіденційність [10–11].

Стандарти та вимоги до зберігання даних для організацій можуть варіюватися залежно від країни, специфіки сфери, діяльності, ступеню конфіденційності та додаткових обставин. Організація може мати специфічні стандарти та вимоги, які диктують її потреби та законодавчі вимоги.

Більша частина організацій або установ формують свою політику безпеки на основі міжнародних стандартів, які переважно проводяться за участю зовнішніх авдиторських компаній, які сертифікують на відповідність до відповідного стандарту.

Проте фахівці, що спеціалізуються на захисті та збереженні масштабних датасетів, досі натрапляють на численні виклики у своїй роботі. Наприклад, вони змушені боротися з проблемами цілісності, конфіденційності та доступності даних. Забезпечення того, щоб інформація залишалася незмінною від створення в процесі подальшого зберігання та відтворення, може бути складним завданням. До того ж професіонали мають гарантувати конфіденційність так, щоб тільки уповноважені особи мали доступ до даних. Вони також мають забезпечити, щоб дані були легко доступні за потреб, що може бути складним в епоху швидкого зростання обсягів даних.

Хоча є багато ефективних підходів, методів та способів організації зберігання великих даних, все ще є певні проблеми в цій сфері. Зокрема, як дуже актуальну проблему можна виділити пошук потрібної інформації в неструктурованих даних.

ISO 27001 – це стандарт, розроблений для забезпечення належного управління цифровими активами компанії, включно з фінансовою інформацією, інтелектуальною власністю, даними працівників та інформацією від довірених третіх сторін.

SOC 2 більш визнаний і такий, якому зазвичай віддають перевагу американські та канадські компанії. SOC 2 Type 2 – це звіт, який підтверджує здатність організації дотримуватися безпеки, доступності, цілісності обробки, конфіденційності та приватності.

Інша важлива деталь: SOC поділяється на SOC 1, SOC 2 та SOC 3. У першому йдеться лише про фінансовий контроль, а третій переважно використовується для маркетингових цілей, тому постачальники SaaS можуть зосередитися лише на SOC 2 [12–14].

2. Постановка завдання

Основною метою цієї статті є розроблення та впровадження методики збору, обробки, класифікації та зберігання даних відповідно до стандарту SOC 2 Type 2 для забезпечення безпеки, доступності та підвищення конкурентоздатності і довіри клієнтів. Основний акцент робиться на поліпшенні внутрішніх процедур та підвищенні розуміння ризиків і контролів серед співробітників. Методика має бути спрямована на створення моделі даних, їх класифікацію та посилання на метадані для забезпечення безпеки та відповідності нормативним вимогам. Водночас потрібно забезпечити баланс між безпекою та зручністю використання, автоматизацію, інтеграцію з іншими політиками і контрольними заходами, а також юридичну та регулятивну відповідність. Методика

має надавати гнучкість у виборі технологій та складу команди, що сприяє досягненню високої ефективності управління даними та підвищенню конкурентоспроможності компаній, бути простою в організації та відповідати бюджетним обмеженням організації.

3. Роль класифікації інформації у безпеці даних

Класифікація даних – це процес організації даних у категорії, що полегшують їх управління та захист на основі рівня чутливості і впливу на організацію у разі несанкціонованого доступу, зміни або знищення. Це критичний перший крок у встановленні міцної стратегії безпеки даних, оскільки допомагає організаціям зрозуміти, які дані вони мають, і призначити рівень чутливості для цих даних, що визначає безпечність, яку потрібно застосовувати. Визнання цього важливе, оскільки великі дані відіграють важливу роль в аналізі даних. Саме аналітика дає можливість нам правильно розуміти й інтерпретувати ці дані, щоб їх можна було використовувати для ухвалення правильних і обґрунтованих рішень, передбачення тенденцій і т. ін. Треба усвідомлювати, що сховища даних значно відрізняються від простої концепції "об'ємної бази даних". Основна відмінність в тому, що бази даних зазвичай зберігають структуровані дані та мають фіксовану схему, тоді як сховища неструктурованих даних також можуть зберігати неструктуровані дані та обробляти великі обсяги інформації. Основні цілі класифікації даних полягають у тому, щоб організувати та керувати ними таким способом, щоб поліпшити їх захист та відповідати загальній стратегії безпеки даних організації. Процес передбачає надання категорій даним на основі їх рівня чутливості та потенційного впливу на організацію у разі неправомірного доступу, зміни або знищення цих даних.

Для забезпечення цього процесу можна виділити такі цілі:

1. *Визначення чутливих даних*.

Класифікація даних дає можливість організаціям визначити, які дані чутливі та потребують надійніших заходів захисту. До цих даних входять персональні дані (PII), фінансові дані, медичні записи та інтелектуальна власність.

2. *Управління ризиками*.

За допомогою класифікації даних організації можуть краще розуміти ризики, пов'язані з кожним типом даних. Вищі рівні класифікації зазвичай вказують на більшу потребу в захисті через збільшений ризик.

3. *Дотримання регуляторних вимог*.

Багато галузей підпорядковані регуляціям, які потребують захисту певних типів даних або встановлюють конкретні правила для їх зберігання та доступу. Класифікація даних є критичною для дотримання таких регуляцій, як вимога про захист персональних даних (GDPR), закон медичного страхування та підзвітності (HIPAA) та інших.

4. *Забезпечення заходів безпеки*.

За допомогою класифікації даних організації можуть застосовувати відповідні засоби безпеки там, де вони найбільше потрібні. Цей цілеспрямований підхід забезпечує те, що найчутливіші дані одержують найвищий рівень захисту, оптимізуючи використання ресурсів безпеки.

5. *Підтримка контролю доступу*.

Правильна класифікація даних допомагає впровадженню ефективних засобів контролю доступу. Вона забезпечує те, що доступ до чутливих даних обмежений авторизованим особам залежно від їх ролей та принципу необхідності знання.

6. *Визначення управління життєвим циклом даних*.

Класифікація допомагає визначити, як треба поводитися з даними впродовж їх життєвого циклу, включно із зберіганням, архівуванням, організацією доступу до них та політикою безпечного знищення.

7. *Пріоритезація*.

У разі виникнення події безпеки розуміння класифікації постраждалих даних може допомогти пріоритезувати зусилля з реагування та відновлення, таким способом мінімізуючи потенційний вплив на організацію.

8. *Підвищення освіти та відповідальності*.

Сприяє підвищенню свідомості серед працівників про типи даних, з якими вони працюють, та їхню відповідальність у забезпеченні їх безпеки, таким способом слугуючи культурі безпеки та відповідальності в організації [15–16].

4. Цінність SOC 2 Type 2

Користь SOC 2 Type 2 полягає в тому, що цей звіт став стандартом для підприємств, які бажають забезпечити клієнтів, партнерів та зацікавлених сторін щодо безпеки їхніх даних та систем. Цей звіт, який буде виданий незалежним авдитором, дає глибокий огляд та підтвердження ефективності контролів інформаційної безпеки компанії впродовж певного періоду часу.

Основна причина, чому звіт SOC 2 Type 2 цінний для компанії, полягає в тому, що він надає чіткі докази того, що у компанії встановлені надійні та ефективні контролі для захисту даних клієнтів. У цифрову епоху безпека даних є одним із головних пріоритетів для підприємств та клієнтів. Порушення безпеки даних призводить не лише до економічних збитків, але й може завдати ударів по діловій репутації компанії.

Звіт SOC 2 Type 2 допомагає будувати довіру у клієнтів, демонструючи, що компанія прийняла неодмінні заходи для захисту їхніх даних. Це чіткий сигнал клієнтам, що їхні дані захищені, безпечні та обробляються відповідно до стандартів галузі або навіть перевищують їх. Ще однією перевагою відповідності SOC 2 Type 2 є можливість отримання конкурентної переваги. Компанії, які досягли відповідності SOC 2 Type 2, можуть відрізнятися від конкурентів, які цього не зробили. Це може бути вирішальним фактором для потенційних клієнтів під час вибору між різними постачальниками послуг. Крім того, звіт допомагає компаніям уникнути штрафів, пов’язаних з невідповідністю. Різні закони та нормативні акти вимагають від підприємств вжити певних заходів для захисту даних клієнтів. Завдяки відповідності SOC 2 Type 2 компанії можуть продемонструвати, що вони виконують ці вимоги, уникнувши потенційних штрафів та правових ускладнень.

Звіт SOC 2 Type 2 також допомагає компаніям виявити та усунути вразливості у контролі інформаційної безпеки. Процес досягнення відповідності потребує комплексного огляду політик та процедур інформаційної безпеки компанії. Це дає можливість виявити будь-які слабкі місця чи прогалини, які потребують уваги, таким способом зміцнюючи загальну позицію компанії щодо безпеки.

Варто також зазначити, що такий звіт може допомогти поліпшити внутрішні процеси компанії. Процес досягнення відповідності потребує від компанії документування і формалізації політик та процедур інформаційної безпеки. Це може призвести до ефективніших процесів та глибшого розуміння ризиків і контролів інформаційної безпеки компанії серед співробітників.

5. Результати дослідження

Пропонуємо таку схему збору, обробки, класифікації даних відповідно до стандарту SOC 2 Type 2 (див. рисунок).

Для того, щоб реалізувати процес класифікації даних відповідно до SOC 2 Type 2, треба виконати низку кроків:

Крок 1: *Розуміння типів даних, що належать вашій компанії.*

Дані можна широко класифікувати на три категорії: структуровані, напівструктуровані та неструктуровані.

Структуровані належать до даних, організованих за попередньо визначеною схемою, таких як дані, збережені в реляційній базі даних. Структуровані дані легко шукати, аналізувати та маніпулювати, оскільки вони мають послідовний формат.

Напівструктуровані належать до даних, які мають певний рівень організації, але не відповідають жорсткій схемі. Приклади напівструктурованих даних містять файли XML та JSON з даними в ієрархічному форматі, але не мають фіксованої схеми.

Схема збору, обробки, класифікації даних

Неструктуровані дані належать до таких, які не мають вбудованої структури або організації. Приклади неструктурованих даних містять текстові документи, зображення та відео. Неструктуровані дані можуть бути складними для пошуку, аналізу та маніпулювання, оскільки вони не мають послідовного формату [17–18].

Крок 2: *Розуміння метаданих, пов'язаних із вашими даними.*

Після визначення типів даних, що належать вашій компанії, наступним кроком є розуміння метаданих, пов'язаних з цими даними. Метадані належать до даних, які надають інформацію про інші дані. Наприклад, метадані, пов'язані з текстовим документом, можуть вміщати автора, дату створення та розмір файлу. Розуміння метаданих, пов'язаних з вашими даними, може допомогти краще їх організувати, керувати ними та їх аналізувати [19].

Крок 3: *Використання інструментів інтеграції для керування та зберігання даних.*

Після визначення типів даних, що належать вашій компанії, та метаданих, пов'язаних з цими даними, наступним кроком є використання інструментів інтеграції для керування та зберігання даних. Інструменти інтеграції дають можливість витягувати інформації з різноманітних ресурсів, перетворювати їх на уніфікований формат та завантажувати у сховище даних. Цей процес, відомий як Extract, Transform, Load (ETL), дає можливість об'єднувати дані в одному місці, що полегшує їх керування та аналіз [20–23].

Крок 4: *Створення моделі даних.*

Після того як дані були видобуті, перетворені та завантажені у сховище даних, наступним кроком є створення моделі даних. Модель даних – це візуальне відображення взаємозв'язків між різними елементами даних. Вона надає межі для організації та структуризації даних і може допомогти виявити закономірності та тенденції у них [24].

Крок 5: *Класифікація та посилання ваших даних на метадані.*

Після створення моделі даних наступним кроком є класифікація ваших даних та посилання їх на пов'язані з ними метадані. Це присвоєння рівня чутливості вашим даним на основі їх важливості та потенційного впливу, якщо вони були загублені або вкрадені. Після класифікації ваших даних ви можете посилати їх на пов'язані з ними метадані, надаючи додатковий контекст та інформацію про дані [25].

Крок 6: *Візуалізація та керування вашими даними.*

Останнім кроком є створення додатку, який дає можливість візуалізувати та керувати вашими даними. Цей додаток має надавати користувачам зручний інтерфейс для доступу, аналізу та маніпулювання даними. Він також має вміщати логіку для керування доступом, запитами та

інцидентами та бути інтегрований з вашою системою управління IT-сервісами (ITSM), щоб забезпечити обробку даних відповідно до політик і процедур вашої компанії [26–27].

Це рішення має безліч переваг порівняно з традиційними продуктовими пропозиціями різних компаній. Однією з основних переваг є можливість вибору середовища розміщення, яке найкраще відповідає вашим потребам: локального або хмарного. Це дає можливість будувати рішення, що буде ґрунтуватися на операційних вимогах та можливостях інфраструктури.

Крім того, у вас є можливість вибрати набір технологій, який найкраще підходить для вашого проєкту. Це означає, що ви не обмежені заздалегідь визначеним набором технологій, але можете налаштувати рішення для використання найактуальніших та найефективніших інструментів для конкретних потреб.

Щодо складу команди, то можна створити команду, яка якнайкраще підходить для конкретного проєкту. Ця гнучкість забезпечує застосування правильних знань та навичок для досягнення найкращих результатів.

Ще однією перевагою є гнучкість у плануванні бюджету. На відміну від рішень, що пропонують конкретні постачальники, які можуть мати фіксовані вартості ліцензування, бюджет на це рішення може бути адаптований відповідно до вашої фінансової можливості та вимог проєкту. Це може призвести до значних економічних переваг без втрати якості або продуктивності.

Підсумовуючи, варто зазначити, що запропоноване нами рішення подає надійні можливості управління змінами та функціоналом. Це означає, що воно легко адаптується до змінних бізнес-потреб з можливістю долучення нових функцій та внесення потрібних змін вчасно та ефективно. Рішення здешевлює вартість від 50 до 80 % запропонованих уже на ринку продуктів, з другого боку, ми отримуємо 100-відсоткову гнучкість у реалізації будь-яких нових потреб.

Висновки

Розробка політики класифікації даних для відповідності SOC 2 Type 2 є складним, але критично важливим завданням для організацій. SOC 2 Type 2 – це сертифікація, яка свідчить про здатність сервісної організації відповідати критеріям служб, що охоплюють безпеку, доступність, цілісність обробки, конфіденційність та приватність.

Класифікація даних є критично важливим першим кроком у створенні стратегії безпеки даних, оскільки вона допомагає організаціям зрозуміти, які дані вони мають, і призначає рівень чутливості цим даним, який визначає застосування контрольних заходів безпеки.

Основними цілями класифікації даних є організація та управління даними таким способом, щоб підвищити їх захист і відповідність загальній стратегії безпеки даних організації. Розробка політики класифікації даних для відповідності SOC 2 Type 2 містить кілька викликів, які організації мають враховувати, щоб ефективно захистити конфіденційну інформацію та зберегти цілісність своєї доставки послуг. Ці виклики становлять розуміння обсягу даних, відповідність критеріям, баланс між безпекою та зручністю використання, навчання та освіту, регулярні оновлення та перегляди, визначення рівнів класифікації, забезпечення послідовності, автоматизацію класифікації, інтеграцію з іншими політиками та контрольними заходами, співпрацю з третіми сторонами, моніторинг та підтримку, а також юридичну і регулятивну відповідність.

Розв’язання цих викликів потребує стратегічного підходу та постійного зобов’язання зберегти політику класифікації даних. Організації можуть звернутися по консультацію до експертів у цій галузі, юридичних радників та професіоналів з авдиту SOC 2, щоб розробити та впровадити політику, яка не тільки відповідає вимогам SOC 2 Type 2, але й підтримує загальну стратегію управління даними організації.

Запропоноване нами рішення має на меті продемонструвати простоту та гнучкість процесу, який можна розробити за допомогою технологій та ресурсів, що прийнятні для компанії в межах доступного бюджету.

Список літератури

1. Matturdi B., Zhou X., Li S. and Lin F. Big Data security and privacy: A review, in China Communications, vol. 11, No. 14, pp. 135–145, 2014. DOI: 10.1109/CC.2014.7085614

2. Susukailo V., Opirskyy I., Vasylyshyn S. Analysis of the attack vectors used by threat actors during the pandemic // 2020 IEEE 15th International Scientific and Technical Conference on Computer Sciences and Information Technologies, CSIT 2020 – Proceedings, 2020, 2, pp. 261–264, 9321897. DOI: 10.1109/CSIT49958.2020.9321897

3. Islam M. N., Zaki T., Uddin M. S., Hasan M. M. Security threats for big data: An empirical study. Int J Inf Commun Technol Human Dev (IJICTHD). 2018; 10(4): pp. 1–18. DOI:10.4018/IJICTHD.2018100101

4. Singh A., Kumar A., Namasudra S. DNACDS: Cloud IoE big data security and accessing scheme based on DNA cryptography. Frontiers Comput. Sci. 18(1): 181801 (2024). DOI: 10.1007/s11704-022-2193-3

5. Harasymchuk O. I., Kostiv Yu. M., Maksymovych V. M., Mandrona M. M. Generator of pseudorandom bit sequence with increased cryptographic security // Metallurgical and Mining Industry: scientific and technical jornal. Dnipropetrovsk, 2014, No. 5, pp. 25–29.

Available at: https://www.metaljournal.com.ua/assets/Journal/6-KostivY.pdf (Accessed: 15 March 2024).

6. Lakhno V., Kozlovskii V., Boiko Y., Mishchenko A., Opirskyy I. Management of information protection based on the integrated implementation of decision support systems // Eastern-european journal of enterprise technologies. Information and controlling system. Vol. 5, No. 9(89), 2017, p. 36–41. DOI: 10.15587/1729-4061.2017.111081

7. Hulak H., Kriuchkova L., Skladannyi P., & Opirskyy I. (2021). Formation of requirements for the electronic record-book in guaranteed information systems of distance learning. Paper presented at the CEUR Workshop Proceedings, 2923 , 137–142. Available at: https://ceur-ws.org/Vol-2923/paper15.pdf (Accessed: 15 March 2024).

8. Maksymovych V., Shabatura M., Harasymchuk O., Karpinski M., Jancarczyk D., Sawicki P. Development of Additive Fibonacci Generators with Improved Characteristics for Cybersecurity Needs. Appl. Sci. (2022), 12(3), 1519. pp. 1–12. https://doi.org/10.3390/app12031519

9. Maksymovych V., Shabatura M., Harasymchuk O., Shevchuk R., Sawicki P., Zajac T. Combined Pseudo-Random Sequence Generator for Cybersecurity. Sensors 2022, 22, 9700. pp.1–17. https://doi.org/10.3390/s22249700

10. Available at: https://secureframe.com/hub/soc-2/compliance-documentation (Accessed: 15 March 2024).

11. Available at: https://www.iso.org/standard/27001 (Accessed: 15 March 2024).

12. Maksymovych V., Nyemkova E., Justice C., Shabatura M., Harasymchuk O., Lakh Y., Rusynko M. Simulation of Authentication in Information-Processing Electronic Devices Based on Poisson Pulse Sequence Generators. Electronics. (2022); 11(13):2039. p.18 https://doi.org/10.3390/electronics11132039

13. Yi J., Wen Y. An Improved Data Backup Scheme Based on Multi-Factor Authentication. BigDataSecurity/HPSC/IDS 2023: pp. 187–197 Available at: https://ietresearch.onlinelibrary.wiley.com/doi/10.1049/iet-ifs.2016.0103 (Accessed: 15 March 2024).

14. Shevchuk D., Harasymchuk O., Partyka A., Korshun N. Designing Secured Services for Authentication, Authorization, and Accounting of Users (short paper). CPITS II 2023: pp. 217–225. Available at: https://ceur-ws.org/Vol-3550/short4.pdf (Accessed: 15 March 2024).

15. ARMA International, Information Classification: Getting It Right. Available at: https://www.arma.org/ (Accessed: 15 March 2024).

16. Vic (J. R.) Winkler. Securing the Cloud: Cloud Computer Security Techniques and Tactics, pages 314, 2011. Available at: https://www.amazon.com/Securing-Cloud-Computer-Security-echniques/dp/1597495921 (Accessed: 15 March 2024).

17. Karumanchi Narasimha. Data Structures and Algorithms Made Easy. Pages: 432, Year of Release: 2016. Available at: https://www.amazon.in/Data-Structures-Algorithms-Made-Easy/dp/819324527X (Accessed: 15 March 2024).

18. Watson Richard T. Data Management: Databases and Organizations. Pages 624, Year of Release: 2017. Available at: https://www.ebay.com/itm/335087377552 (Accessed: 15 March 2024).

19. Rhodes-Ousley Mark. Information Security: The Complete Reference, Second Edition, pages 896, 2012.

20. Cote Christian, Lah Matija. Professional Microsoft SQL Server 2014 Integration Services (SSIS). Pages: 912, Year of Release: 2014. Available at: https://www.amazon.com/Professional-Microsoft-Integration-Services-Programmer-ebook/dp/B00JSQ3RLG (Accessed: 15 March 2024).

21. Harenslak Bas P. (Author), Rutger de Ruiter Julian. Data Pipelines with Apache Airflow Pages: 480, Year of Release: 2021. Available at: https://www.amazon.com/Data-Pipelines-Apache-Airflow-Harenslak/dp/1617296902 (Accessed: 15 March 2024).

22. Available at: https://docs.aws.amazon.com/glue/ (Accessed: 15 March 2024).

23. Available at: https://learn.microsoft.com/en-us/azure/data-factory/ (Accessed: 15 March 2024).

24. Hoberman S. Data Modeling Made Simple: A Practical Guide for Business and IT Professionals, Pages: 314, Year of Release: 2005. Available at: https://www.amazon.com/Data-Modeling-Made-Simple-Professionals/dp/0977140008 (Accessed: 15 March 2024).

25. Data Classification: Algorithms and Applications / edited by Charu C. Aggarwal. Pages: 598, Year of Release: 2014. Available at: https://doc.lagout.org/science/0_Computer%20Science/2_Algorithms/Data%20Classification_%20Algorithms%20and%20Applications%20%5BAggarwal%202014-07-25%5D.pdf (Accessed: 15 March 2024).

METHODOLOGY FOR COLLECTING, PROCESSING, STORING, AND CLASSIFYING DATA IN ACCORDANCE WITH SOC 2 TYPE 2 REQUIREMENTS

O. Deineka, L. Bortnik

Lviv Polytechnic National University,
Information Security Department
E-mail: deinekaoleg.86@gmail.com, leonid.l.bortnik@lpnu.ua

This article explores the creation of a data classification policy in line with SOC 2 Type 2 compliance requirements. SOC 2 Type 2 is a notable certification that attests to an organization's ability to adhere to the Trust Services Criteria, including security, availability, processing integrity, confidentiality, and privacy.

The initial and crucial step in formulating a solid data security strategy is data classification, which helps organizations recognize their data and assign a sensitivity level, guiding the appropriate security measures. Data classification aims to organize and manage data in a manner that enhances its protection and aligns with the organization's overall data security strategy. In the data classification process, data security has a central role as it directly impacts the protection and management of classified data.

The design of a data classification policy for SOC 2 Type 2 compliance presents several challenges and considerations. Organizations must understand the scope of their data, align with the Trust Services Criteria, balance security with usability, provide training and awareness, conduct regular updates and reviews, define classification levels, ensure consistency, automate classification, integrate with other policies and controls, handle third-party vendors, monitor and enforce, and comply with legal and regulatory requirements.

Keywords: SOC 2 Type 2, storage standards, data classification, data storage, data security.