

**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**Факультет інформаційних технологій**

Кафедра технологій управління

Спеціальність 122 – Комп’ютерні науки,  
освітня програма «Інформаційна аналітика та впливи»

**КВАЛІФІКАЦІЙНА РОБОТА МАГІСТРА**

на тему:

**“Інформаційна аналітика впливу методик гнучких методологій  
управління проектами на успіх проектів Big Data”**

**Студентки 2-го курсу групи ІАВ-21**

Руденко Владислави Олегівни

**Науковий керівник:**

к.е.н. доц. Мезенцева Ольга  
Олексіївна

**Попередній захист:**

---

(Висновок: «До захисту в Екзаменаційній комісії»)

Завідувач кафедри  
технологій управління

\_\_\_\_\_

(підпис)

\_\_\_\_\_

(прізвище, ініціали)

\_\_\_\_\_

(дата)

**Київ – 2021**

## ЗАВДАННЯ

### КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ ІМЕНІ ТАРАСА ШЕВЧЕНКА Факультет інформаційних технологій

Кафедра технологій управління  
Освітньо-кваліфікаційний рівень Магістр  
Спеціальність 122 - Комп'ютерні науки  
Освітня програма Інформаційна аналітика та впливи

**ЗАТВЕРДЖУЮ**  
Завідувач кафедри  
професор Морозов В.В.

«\_\_\_» \_\_\_\_\_ 2021 року

### ЗАВДАННЯ НА ВИКОНАННЯ КВАЛІФІКАЦІЙНОЇ РОБОТИ

Студент: Руденко Владислава Олегівна

Група: ІАВ-21

#### **1. Тема кваліфікаційної роботи**

«Інформаційна аналітика впливу методик гнучких методологій управління проектами на успіх проектів Big Data»

Затверджена наказом по від «\_\_\_» \_\_\_\_\_ 20\_\_ р. № \_\_\_.

**2. Строк подання студентом готової роботи** – “\_\_\_” \_\_\_\_\_ 20\_\_ р.

**3. Цільова установка та вихідні дані до роботи** дослідити та надати оцінку впливу методик гнучких методологій на успіх проекту Big Data за допомогою мови програмування Python на основі даних, отриманих від проведення експертного оцінювання та статистичних даних про проекти Big Data

**4. Зміст роботи** визначити особливості проектів Big Data та застосування гнучких методологій до їх ведення, визначити і описати окремі методики гнучких методологій управління проектами як самостійні інструменти для подальшого дослідження; визначити практичне поняття успіху проекту; розробити екземпляр для заповнення експертних оцінок та провести експертне оцінювання, побудувати регресійну модель для аналізу отриманих даних та визначити ефективні методики гнучких методологій для Big Data проекту, сформувати практичні рекомендації щодо використання гнучких методологій у проектах Big Data, теоретичні основи використання гнучких методологій в проектному управлінні, особливості проектів Big Data та застосування гнучких методологій до їх ведення, визначення відмінних характеристик гнучких методологій управління проектами та їх основних методик, практичне визначення поняття та критеріїв успіху проекту, оцінювання використання методик гнучких методологій у проектах Big Data, проведення експертного оцінювання, вибір ефективних методик гнучких методологій для Big Data проекту, рекомендації і подальше дослідження, практичний досвід з впровадження гнучких методологій у Big Data команди

**5. Перелік графічного матеріалу (слайдів) 5 таблиць, 37 рисунків, 14 формул, 3 додатки, 15 слайдів презентації доповіді**

**6. Календарний план виконання роботи:**

№ з/п	Назва частин роботи	%	Виконання роботи	
			За планом	Фактично
1.	Вибір теми дипломної роботи.	3	01.10.2020	01.10.2020
2.	Протокол кафедри ТУ про затвердження тем дипломних робіт та призначення наукових керівників.	1	09.11.2020	09.11.2020
3.	Формування переліку нормативних матеріалів, літератури з проблематики дипломної роботи.	10	11.01.2021	11.01.2021
4.	Складання розгорнутого плану кваліфікаційної роботи.	4	17.01.2021	17.01.2021
5.	Ознайомлення наукового керівника з розгорнутим планом кваліфікаційної роботи. Внесення змін.	4	19.01.2021	19.01.2021
6.	Підготовка розділу 1	10	11.02.2021	11.02.2021
7.	Підготовка розділу 2	17	09.03.2021	09.03.2021
8.	Підготовка розділу 3	23	25.03.2021	25.03.2021
10.	Підготовка розділу 4	9	15.04.2021	05.04.2021
9.	Оформлення кваліфікаційної роботи. Підготовка висновків і пропозицій.	10	29.04.2021	29.04.2021
10.	Передача кваліфікаційної роботи науковому керівникові	2	04.05.2021	04.05.2021
11.	Передача кваліфікаційної роботи рецензенту для рецензування	2	05.05.2021	05.05.2021
12.	Попередній захист кваліфікаційної роботи	5	11.05.2021	11.05.2021
13.	Подача готової роботи на кафедру		20.05.2021	20.05.2021

Дата видачі завдання « 1 » жовтня 2020 р.

Керівник роботи к.е.н. доц. Мезенцева Ольга Олексіївна

\_\_\_\_\_  
(підпис)

Завдання прийняв до виконання студент групи ІАВ-21  
Руденко Владислава Олегівна

\_\_\_\_\_  
(підпис)

## ЗМІСТ

АНОТАЦІЯ .....	5
ВСТУП .....	7
РОЗДІЛ 1 ТЕОРЕТИЧНІ ОСНОВИ ВИКОРИСТАННЯ ГНУЧКИХ МЕТОДОЛОГІЙ В ПРОЕКТНОМУ УПРАВЛІННІ .....	11
1.1 Основи використання гнучких методологій у проектній діяльності .....	11
1.2 Особливості управління проектами Big Data .....	16
1.3 Використання гнучких методологій для Big Data проектів .....	25
1.4 Постановка задачі .....	28
РОЗДІЛ 2 ПРАКТИЧНЕ ВИКОРИСТАННЯ ГНУЧКИХ МЕТОДОЛОГІЙ ДЛЯ ПРОЕКТНОГО АНАЛІЗУ .....	30
2.1 Відмінні характеристики гнучких методологій управління проектами та визначення їх найпопулярніших методик .....	30
2.2 Практичне визначення поняття та критеріїв успіху проекту .....	43
РОЗДІЛ 3 ОЦІНЮВАННЯ ВИКОРИСТАННЯ МЕТОДИК ГНУЧКИХ МЕТОДОЛОГІЙ У ПРОЕКТАХ BIG DATA .....	55
3.1 Механізм оцінювання методик гнучких методологій для проектів Big Data .....	55
3.2 Вибір ефективних методик гнучких методологій для Big Data проекту .....	70
РОЗДІЛ 4 РЕКОМЕНДАЦІЇ І ПОДАЛЬШЕ ДОСЛІДЖЕННЯ .....	80
4.1 Практичні рекомендації .....	80
4.2 Практичний досвід з впровадження гнучких методологій у Big Data команді .....	87
4.3 Подальше дослідження .....	92
ВИСНОВКИ .....	94
ПЕРЕЛІК ВИКОРИСТАНИХ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ .....	99
ДОДАТКИ .....	106

## АНОТАЦІЯ

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ

ІМЕНІ ТАРАСА ШЕВЧЕНКА

Факультет інформаційних технологій

Кафедра технологій управління

Спеціальність 122 - Комп'ютерні науки,  
освітня програма "Інформаційна аналітика та впливи"

Дипломна робота магістра Руденко Владислави Олегівни

Тема роботи – «Інформаційна аналітика впливу методик гнучких методологій управління проектами на успіх проектів Big Data».

Мета дипломної роботи магістра – дослідити та надати оцінку впливу методик гнучких методологій на успіх проекту Big Data.

Об'єкт дослідження – процеси управління проектами Big Data.

Предмет дослідження – інструменти інформаційної аналітики, що описують механізми гнучких методологій в управлінні проектами Big Data.

Наукова новизна даної роботи полягає у визначенні конкретних методик гнучких методологій саме для проектів Big Data, що дозволяють досягти максимально можливого успіху для усього проекту загалом. Для аналізу було розроблено регресійну модель з використанням мови програмування, обробки та аналізу Python.

Використаними у роботі методами досліджень є методи збору та агрегації даних, метод експертних оцінок та аналіз отриманих результатів, дискретні методи, регресійні методи, статистичні методи та інструменти для обробки та аналізу даних, інструменти візуалізації даних.

Практична значимість дослідження полягає в дослідженні та визначенні методик, що впливають на успіх Big Data проекту, при використанні яких можливо досягнути покращень в управлінні проектом та вплинути на його кінцеві результати.

Дипломна робота складається зі вступу, основної частини, яка включає чотири розділи, висновків та списку використаних джерел. Всього налічує 118 сторінок та перелік посилань з 71 джерела на 7 сторінках.

Ключові слова: проект, Big Data, гнучка методологія, методика, аналіз даних.

## ВСТУП

Сучасний світ є епохою інновацій, він постійно змінюється під впливом рушійних сил. Економіка, широкомасштабні соціальні та політичні зміни, демографічна ситуація, високі технології, що з'являються на світовому ринку, а також розвиток теорії організації систем - все це сприяє появі інноваційних рішень.

Big Data, «Великі дані» ось уже кілька років є одним з ключових напрямків в ІТ індустрії. Це пов'язано з тим, що цифрові технології є невід'ємною частиною життя сучасної людини. Обсяг даних про різні аспекти життя зростає, і одночасно зростають можливості зберігання інформації.

У даний час проектна діяльність в Big Data також розширюється як в кількісному, так і в грошовому обсягах. Разом з цим відсоток неуспішних проектів все ще залишається високим, що збільшує ризики економічної діяльності в цілому (за даними звіту CHAOS Report, The Standish Group, 2015 [7]). Важливою складовою успіху є правильний вибір методології управління проектом. Саме тому ще на етапі ініціації необхідно обґрунтовано вибрати методологію і якісно запланувати методи управління і основні методики.

Однією з ключових характеристик сучасного ринку Big Data проектів є висока невизначеність і зростаюча швидкість настання змін. Ринкові тенденції пояснюють одну з головних проблем сучасних проектів - в процесі реалізації проекту змінюється і бізнес-середовище, тобто що ми робимо, і технології, використовувані передовими компаніями-конкурентами, тобто як ми це робимо. Це призводить до того, що замовник не може сформулювати детально і вичерпно свої вимоги на етапі ініціації. Специфіка поточної ситуації на ринку змушує менеджера планувати роботу таким чином, щоб була можливість вносити зміни під час реалізації проекту, що не представляється можливим в рамках класичної моделі фіксованого обсягу, бюджету і термінів, однак є ключовим аспектом для проектів Big Data.

Через прискорення змін у бізнес-процесах та постійний високий рівень невизначеності компаніям доводиться пристосовуватися до нових вимог ринку. Для вирішення цієї проблеми керівники проектів змушені докорінно переглянути підхід до управління проектами Big Data. З різних типів методологій гнучкі методології виявились найбільш придатними для поточного середовища. Історично використовувані в IT-проектах, гнучкі методології дозволяють вносити зміни під час проекту. Саме це робить їх привабливими для використання у проектах з обробки великих даних та забезпечує подальшу популярність та розвиток.

Однак, слід з обережністю використовувати гнучкі методології, оскільки і вони мають свої недоліки в порівнянні з традиційним підходом. Таким чином, проблема вибору методології ускладнюється тим, що кожна з них має свої переваги і недоліки. Залежно від типу проекту та чи інша методологія буде впливати на різні критерії успіху проекту. Разом з наявністю досліджень про вплив гнучких методологій на успіх проекту, не існує достатніх наукових досліджень на тему впливу окремих складових елементів (методик) гнучких методологій на успіх проекту. Проблема є актуальною та вимагає ретельного вивчення і проведення дослідження. **Наукова новизна** даної роботи полягає у визначенні конкретних методик гнучких методологій саме для проектів Big Data, що дозволяють досягти максимально можливого успіху для усього проекту загалом.

**Метою** даної кваліфікаційної роботи є дослідити та надати оцінку впливу методик гнучких методологій на успіх проекту Big Data.

Для досягнення мети КРМ були визначені наступні **завдання**:

- описати основи використання гнучких методологій у проектній діяльності;
- визначити особливості проектів Big Data та застосування гнучких методологій до їх ведення;
- визначити і описати окремі методики гнучких методологій управління проектами як самостійні інструменти для подальшого дослідження;

- визначити практичне поняття успіху проекту;
- розробити екземпляр для заповнення експертних оцінок та провести експертне оцінювання щодо впливу методик гнучких методологій на успіх проектів Big Data;
- інтерпретувати отримані результати експертного оцінювання;
- побудувати регресійну модель для аналізу отриманих даних та визначити ефективні методики гнучких методологій для Big Data проекту;
- сформулювати практичні рекомендації щодо використання гнучких методологій у проектах Big Data;
- визначити подальший напрямок дослідження.

**Об'єктом** дослідження виступають процеси управління проектами Big Data, **предметом** дослідження – інструменти інформаційної аналітики, що описують механізми гнучких методологій в управлінні проектами Big Data.

Використаними у роботі **методами досліджень** є методи збору та агрегації даних, метод експертних оцінок та аналіз отриманих результатів, дискретні методи, регресійні методи, статистичні методи та інструменти для обробки та аналізу даних, інструменти візуалізації даних.

Для вирішення завдань даної роботи використано мову програмування, обробки та аналізу даних Python.

**Практична значимість** дослідження полягає в дослідженні та визначенні методик, що впливають на успіх Big Data проекту, при використанні яких можливо досягнути покращень в управлінні проектом та вплинути на його кінцеві результати.

**Особистий внесок здобувача.** Усі наукові результати, які відображено у кваліфікаційній роботі, отримані автором самостійно. Результати співавторів сумісних публікацій до тексту кваліфікаційної роботи не включено. У надрукованих статтях, опублікованих у співавторстві, магістранту належить наступне: дослідження застосування різних методологій в управлінні проектами Big Data, вплив використання гнучких методологій на успіх проекту, визначення

найефективніших методик гнучких методологій в управлінні проектами Big Data.

**Апробація результатів роботи.** Автор виступав доповідачем на VI Information Technology and Interactions (Satellite): Conference Proceedings.

**Публікації.** Основні наукові положення, висновки і результати магістерської кваліфікаційної роботи знайшли відображення у 3 друкованих працях, з них: 1 стаття опублікована у іноземному фаховому виданні, 1 стаття англійською мовою, що не є перекладом із інших мов у виданні, включеному до наукометричної бази даних Scopus, 1 тези доповідей у матеріалах конференцій. За період навчання у магістратурі опубліковано 3 наукові праці:

1. The Kübler-Ross factor in managing the performance of technical and socio-economic systems / D. Lukianov, K. Kolesnikova, O. Mezentseva, V. Rudenko // Scientific Journal of Astana IT University, Vol.2, 2020, Pp. 32–43.

2. Influence Analysis of Different Management Methodologies on the Result of Big Data Projects / V. Rudenko, O. Mezentseva // 7th International conference "Information Technology and Interactions (Satellite: Conference Proceedings, December 04, 2020, Kyiv, Ukraine / Taras Shevchenko National University of Kyiv and [etc]; Vitaliy Snytyuk (Editor). Kyiv: Stylos, 2020.– Pp. 165-167.

3. Kolesnikova, K., Rudenko V., Mezentseva, O., Project management methodologies techniques in Big Data analysis. Herald of Ad-vanced Information Technology. Vol.2 No. 3

**Структура та обсяг роботи.** Кваліфікаційна робота складається зі вступу, 4 розділів, висновків. Після цього наводиться список літератури з 71 пунктів та 3 додатків. Загальний обсяг кваліфікаційної роботи становить 118 сторінок, із них 98 сторінок основного тексту, який містить 37 рисунків.

# РОЗДІЛ 1

## ТЕОРЕТИЧНІ ОСНОВИ ВИКОРИСТАННЯ ГНУЧКИХ МЕТОДОЛОГІЙ В ПРОЕКТНОМУ УПРАВЛІННІ

### 1.1 Основи використання гнучких методологій у проектній діяльності

У даний час спостерігається активне зростання розвитку гнучких методологій. Перевірені в проектах з розробки програмного забезпечення, вони стали використовуватися в інших галузях - фінансових послугах, страхуванні, уряді, охороні здоров'я, важкій промисловості тощо (згідно з доповіддю The 12th Annual State of Agile Report, 2018 [9]).

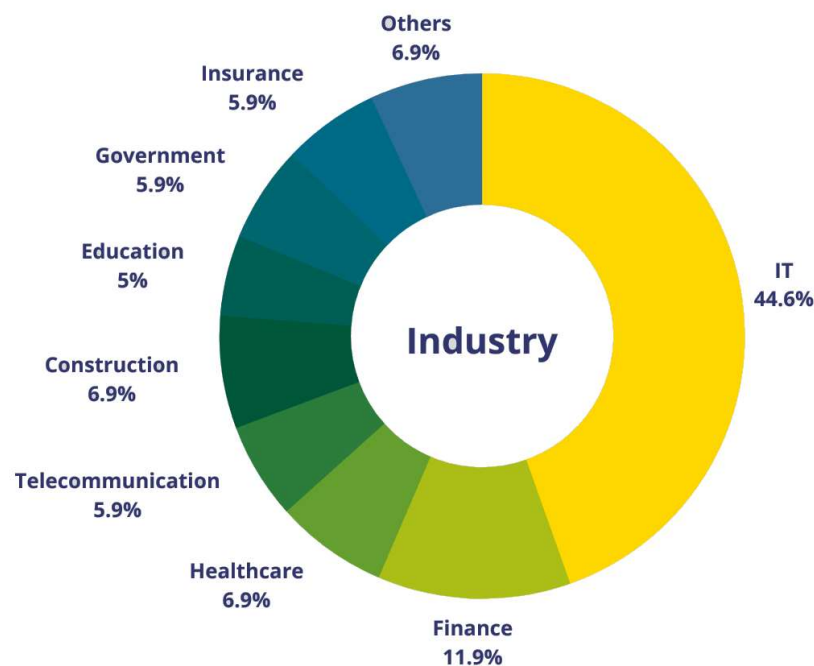


Рисунок 1.1 – Розподіл компаній, що використовують Scrum по індустріях [9]

Також усе більше провідних великих компаній роблять Agile своїм основним напрямком розвитку та основною методологією ведення проектів (див. рисунок 1.2).

Згідно зі звітом CHAOS [7], проекти, що застосовували гнучкі методології, стабільно успішніші, ніж проекти, що використовують традиційні (див. табл. 1.1):

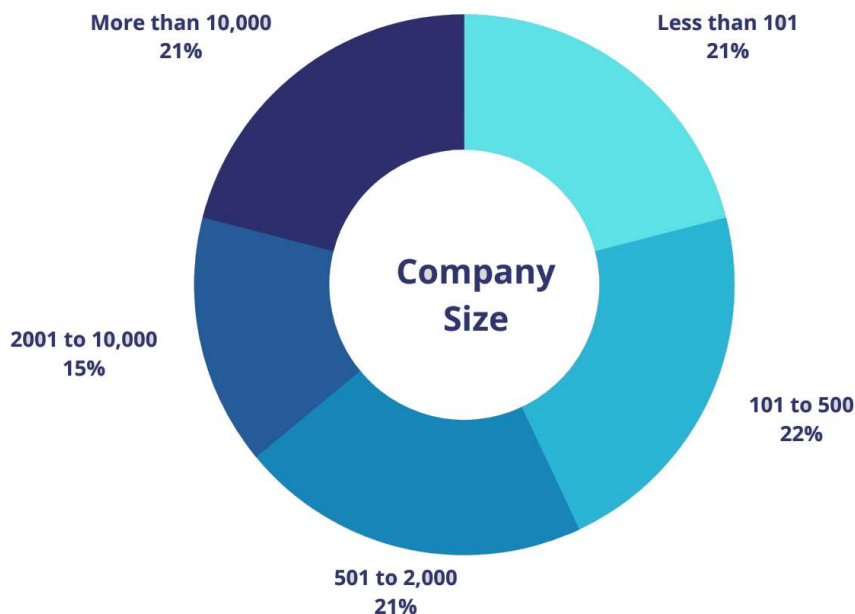


Рисунок 1.2 – Розподіл розміру компаній, що використовують Scrum [9]

Таблиця 1.1 - Кількість успішних і неуспішних проектів в залежності від методології і розміру, %

Розмір проекту	Метод	Успішний проект,%	Неуспішний проект,%
Проекти всіх типів	Agile	39%	9%
	Waterfall	11%	29%
Великі	Agile	18%	23%
	Waterfall	3%	42%
Середні	Agile	27%	11%
	Waterfall	7%	25%
Дрібні	Agile	58%	4%
	Waterfall	44%	11%

Описуючи «гнучкість» проектів, автори поділяють підходи на традиційні, класичні («важкі») та гнучкі («легкі») [10]. Традиційні методології базуються на лінійному підході. У цьому випадку план є окремим етапом, а реалізація проекту здійснюється за схемою планування - реалізація - завершення. У свою чергу, гнучкі методології використовують ітераційний підхід і призначені для легкого врахування змін, що виникають. Замість одного кроку планування використовується планування "за потребою". Це дозволяє команді адаптуватися до умов високої невизначеності та можливості змін, що є цілком застосовним і навіть необхідним у проектах великих даних, де все пов'язано з невизначеністю та змінами.

Щоб зрозуміти, які проблеми в управлінні проектами великих даних повинна вирішувати гнучка методологія у порівнянні з традиційною, представляється очевидним вивчити передумови виникнення та історії розвитку гнучких методологій.

Хеллгрєн та Маанінен-Олссон [11] зазначають, що в проектах рано чи пізно найчастіше з'являються «відхилення» - ситуації, що призводять до відхилень від початкового плану внаслідок змін різного характеру та розміру. Завданням керівника проекту є планування ресурсів для цих відхилень та вироблення підходу до боротьби з ними. В інтерв'ю, проведеному Коллієром та Уорреном [12] з 31 керівником проектів з 10 галузей, було виявлено, що традиційні підходи не працюють у швидко мінливих умовах, оскільки вони не можуть впоратися з трьома типами змін у проекті:

- 1) цілі;
- 2) матеріали, ресурси, техніка та інструменти;
- 3) взаємозв'язок з іншими продуктами, послугами та проектами.

У 1970 році WW Рouse видає «Управління розвитком великих програмних систем». З роботи була помилково виведена каскадна модель розробки ПЗ. Відповідно до статті, успішний проект складався з 7 етапів (рис. 1.3):



Рисунок 1.3 - Етапи проекту згідно "каскадної моделі".

Розробник не може перейти до наступної фази, не виконавши попередню. Це підвищує прозорість процесу та його ефективність [13]. Однак у рамках проекту Big Data з високим рівнем невизначеності та великим обсягом неможливо опрацювати кожну фазу одночасно.

У статті, що описує підхід до розробки продуктів Microsoft [14], Кузумано та Селбі зазначають, що Waterfall (одна з каскадних методологій) поступово втрачає популярність, оскільки компанії, як правило, створюють кращі продукти, якщо вони мають можливість змінювати технічні характеристики та дизайн, отримувати зворотній зв'язок із замовниками та постійно тестувати окремі компоненти по мірі розвитку продуктів. З цієї причини в наступному, четвертому виданні РМВОК було вирішено додати альтернативний ітераційний метод управління проектами [15]. Необхідно відзначити, що в дійсності в своїй роботі У.У. Ройс писав про те, що «каскад» необхідно пройти кілька разів для повного покриття вимог і проходження тестування, але спочатку цій тезі не було приділено належної уваги.

Першими гнучкими методологіями вважаються ті, які засновані на принципі ітеративності, наприклад rolling wave. Так, суть підходу Rolling Wave planning полягає в тому, що планування відбувається на протязі всього проекту [16]. На початку проекту загальний обсяг розбивається на компоненти, що описуються максимально докладно. Далі складається детальне планування

розкладу на перші етапи проекту (наприклад, на дві фази з чотирьох), для інших фаз - на верхньому рівні. У міру наближення до третьої і четвертої хвиль планування коригується і деталізується. З одного боку, така робота дозволяє почати активності за проектом, не чекаючи всіх деталей проекту. З іншого боку, при такому підході обсяг проекту повинен бути точно визначений.

Пізніше знання про ітеративні підходи були з'єднані 17 незалежними практиками в одному джерелі - Agile-маніфесті (далі - Маніфест). Це була перша спроба задокументувати знання про гнучкі методології, тому датою створення Agile прийнято вважати 11-13 лютого 2001 року - дата створення Маніфесту в електронному вигляді. Відповідно до даного збірника правил і принципів, розробка за методологією Agile повинна відповідати таким умовам (Manifesto for Agile Software Development, 2001):

1. Люди і взаємодія важливіше процесів та інструментів.
2. Працюючий продукт важливіше вичерпної документації.
3. Співпраця з замовником важливіше узгодження умов контракту.
4. Готовність до змін важливіше відповідності попередньому плану.

Методики Agile розроблені з використанням мінімуму документації для підтримки гнучкості процесу та швидкості реакції на навколишнє середовище порівняно з традиційним підходом. Це особливо важливо при роботі з великими даними. Важливо зазначити, що Agile Manifesto не встановлює чітких правил управління проектами, як це роблять методології, він не містить таких описів, як управління процесами, методи планування, ключові ролі та документи. Але в той же час Маніфест пропонує 12 принципів, на яких повинен базуватися конкретний підхід до управління проектами.

Для управління таким процесом з високим рівнем невизначеності розроблено безліч гнучких методологій та методів, кожна з яких продовжує або доповнює інші, такі як: Scrum, Екстремальне програмування (XP - Extreme Programming), Ощадлива розробка програмного забезпечення (Lean Software Development), Розробка, керована функціональністю (FDD - feature driven development), Kanban, Адаптивна розробка (ASD) тощо [16]. Далі будуть

детально описані найбільш використовувані з них с метою відбору найбільш популярних методик гнучких методологій.

Гнучкі методології були створені для передачі певної сили управління від керівника проекту команді проекту. Команди будуть залучені до проекту з етапу самоорганізації та зможуть перепланувати завдання під час проекту. Такий підхід підвищує продуктивність праці, дозволяє працівникам вчитися, застосовувати інновації та підвищує задоволеність роботою [19]. У традиційних методологіях "водоспаду" роботою керують менеджери, що створює жорстку ієрархію та розподіл ролей у колективі [20]. Якщо перенести цю тенденцію на проекти з великими даними, стає зрозумілим, що включення аналітиків на фазу планування матиме лише позитивний вплив на результат усього проекту, оскільки це дозволить більш точно визначити терміни проекту та визначити більш сильну гіпотезу на більш ранній даті.

Згідно зі звітом VersionOne за 2018 рік [9], найбільш використовуваними гнучкими підходами є Scrum (54%), Гібридні методології (14%), Scrum / XP Hybrid (10%), ScrumBan (8%) тощо. Scrum, найпопулярнішими є гібридні методології, що ще раз підтверджує актуальність дослідження впливу не однієї методології на успіх проекту Big Data, а методик гнучких методологій. Той самий список найбільш популярних методологій був знайдений у дослідженні Vallona et al. [19]: Scrum - 60%, Kanban - 11%, XP & Scrum - 16%.

## **1.2 Особливості управління проектами Big Data**

Проект, як об'єкт управління, має набір функцій, які вимагають використання спеціальних прийомів і методів управління ним. Протягом останніх сорока років управління проектами стало окремою професійною галуззю та самостійною дисципліною, оснащуючи керівників проектів технологіями та інструментами для планування, моніторингу та координації реалізації проектів.

Проекти Big Data є окремим та одним з ключових напрямків в ІТ-галузі. Більшість експертів сходяться на думці, що прискорення зростання даних є об'єктивною реальністю. Соціальні медіа, мобільні пристрої, дані з вимірювальних пристроїв, ділова інформація - лише кілька джерел, які можуть генерувати гігантські обсяги інформації. Згідно з дослідженням IDC Digital Universe [14], опублікованим у 2012 році, протягом наступних 8 років обсяг даних у світі досягне 40 Zb (зеттабайт), що еквівалентно 5200 ГБ на кожного жителя планети. За період 2018 року цей показник вже досяг 33 Zb.

Парадигма великих даних визначає три основні типи завдань, які дають цим проектам перевагу:

1. Зберігання та управління сотнями терабайт або петабайт даних, які звичайні реляційні бази даних ефективно не використовують.
2. Організація неструктурованої інформації, що складається з текстів, зображень, відео та інших типів даних.
3. Аналіз великих даних, який піднімає питання про те, як працювати з неструктурованою інформацією, створення аналітичних звітів та запровадження прогнозних моделей.

Крім того, використання технологій Big Data є актуальним для рішень класу забезпечення доходів (RA), призначених для автоматизації діяльності компаній. Сучасні системи забезпечення доходів включають засоби виявлення невідповідностей та поглибленого аналізу даних, що дозволяють своєчасно виявляти можливі втрати або спотворення інформації, що може призвести до зниження фінансових результатів. На цьому тлі українські компанії, підтверджуючи попит на технології Big Data на внутрішньому ринку, зазначають, що факторами, що стимулюють розвиток Big Data в Україні, є зростання даних, прискорення прийняття управлінських рішень та вдосконалення їх якості.

Однак сьогодні аналізується лише 0,5% накопичених цифрових даних, незважаючи на те, що об'єктивно існують загальногалузеві завдання, які можна вирішити за допомогою аналітичних рішень класу Big Data. Розвинені ІТ-ринки

вже мають результати, які можна використовувати для оцінки очікувань, пов'язаних із накопиченням та обробкою великих даних.

Одним з основних факторів, що заважає реалізації проектів великих даних, крім високої вартості, є проблема вибору даних, які підлягають обробці: визначення, які дані потрібно витягувати, зберігати та аналізувати, а які не слід брати до уваги.

Багато представників бізнесу зазначають, що труднощі з реалізацією проектів Big Data пов'язані з нестачею фахівців - маркетологів та аналітиків. Норма рентабельності інвестицій у великі дані безпосередньо залежить від якості роботи працівників, які займаються глибокою та прогноною аналітикою. Величезний потенціал даних, що вже існують в організації, часто не можуть ефективно використовувати самі маркетологи через застарілі бізнес-процеси або внутрішні правила. Тому проекти Big Data часто сприймаються бізнесом як складні не лише при реалізації, але й при оцінці результатів: цінності зібраних даних.

Проблеми безпеки даних, що надходять із зовнішніх джерел, повинні мати рішення, що відповідають обсягу зібраної інформації. Оскільки методи аналізу великих даних поки що розробляються лише слідом за зростанням обсягу даних, важливу роль відіграє властивість аналітичних платформ використовувати нові методи підготовки та агрегування даних. Це свідчить про те, що, наприклад, дані про потенційних клієнтів або величезне сховище даних з історією кліків на сайтах інтернет-магазинів можуть бути цікавими для вирішення різних проблем.

Тому в даний час проектна діяльність у Big Data значно зростає, проте кількість невдалих проектів все ще залишається високою. Одним з важливих критеріїв успіху є вибір правильної методології управління проектами. Вже на етапі ініціювання необхідно обґрунтовано вибрати методологію та якісно спланувати методи управління та основні заходи.

Незалежно від кінцевої мети та особливостей реалізації, команда будь-якого проекту великих даних виконує всі процеси відповідно до стандарту

CRISP-DM, від формування бізнес-вимог до впровадження програмного рішення на основі розроблених аналітичних моделей, включаючи використання машинного навчання (Machine Learning). Згрупувавши етапи розробки програмного забезпечення з етапами CRISP-DM, ми можемо виділити чотири спеціалізовані категорії проектів великих даних, в яких беруть участь різні фахівці:

- бізнес, до якого входять спеціалісти з предметних дисциплін (експерти, потенційні користувачі), посередники між проблемами та технічними рішеннями (аналітики, консультанти), а також менеджмент (замовник, керівник проекту);

- дані (Data Professionals: аналітики, архітектори, дослідники та інженери даних) - це люди, відповідальні за збір, перетворення, управління інформаційними потоками та моделями, а також за вилучення корисної ділової інформації з масивів великих даних та моніторинг їх життєвого циклу;

- впровадження - розробники програмного забезпечення (програмісти, тестувальники, дизайнери UI / UX) та програмна документація (технічні автори), а також спеціалісти з розгортання програм з функціями системних адміністраторів та спеціалісти з інформаційної безпеки;

- експлуатація - технічна підтримка користувачів, моніторинг продуктивності програми та релевантності даних, забезпечення інформаційної безпеки, включаючи захист від взлому та витоків даних.

Слід звернути увагу, що фаза впровадження - це фактичний кінець проекту, оскільки тут результат доставляється до кінцевого терміну. Операція - це вже процес, який виконується регулярно. Тому важливо на постійній основі диференціювати тимчасові ролі учасників проекту від тих, хто бере участь у процесі.

За своєю суттю управління проектами з обробки великих даних покладається на загальні методології управління проектами, проте не всі з них можуть бути успішно застосовані в повній мірі. Оскільки це відносно нова

галузь, для великих даних може знадобитися щось нове або, принаймні, комбінація стандартних підходів.

Оскільки Data Science, і зокрема Big Data, все ще розвивається і не має прямої форми, не може бути однозначної відповіді на питання, яка методологія найкраще працює в таких проектах.

Загалом можна виділити наступні методології:

1. CRISP-DM як традиційний підхід в управлінні проектами в галузі Data Science.
2. Waterfall як традиційний підхід.
3. Scrum як Agile підхід.
4. Kanban як Agile підхід.

Це далеко не повний перелік методологій, що використовуються в проектах з обробки великих даних. Проте інструменти, які базуються на вказаних вище підходах, допоможуть визначити ключові моменти в яких може бути корисним кожен підхід, щоб розробити модель, яка буде працювати для конкретного проекту по Big Data.

Стандарт CRISP-DM включає шість ітеративних фаз в управлінні проектами з обробки даних: розуміння бізнес-проблеми, розуміння та отримання даних із різних джерел, підготовка даних, моделювання даних при побудові та оцінці моделі, власне оцінка, що включає візуалізацію і спілкування, та розгортання і технічне обслуговування з остаточними звітами та оглядом проекту. Ця методологія має гнучкий циклічний характер та цілеспрямований підхід, проте вона не працює для команд і зовсім не охоплює питань спілкування. CRISP-DM можна назвати послідовністю робіт, що необхідні для виконання в проектах Data Science.

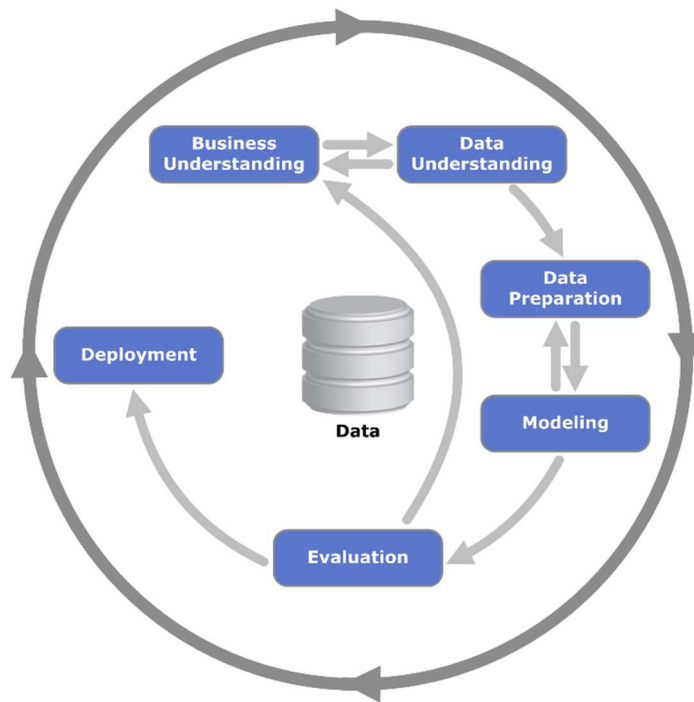


Рисунок 1.4 – Етапи робіт за стандартом CRISP-DM [71]

Цей список робіт можна мати на увазі при керуванні Big Data проектами, проте він є надто загальним, тому варто розглянути також інші інструменти.

Такий підхід як Waterfall дає чітку послідовну картину всіх завдань, які були визначені з початку проекту. Проект або окремі його фази розбиваються на більш дрібні частини, які залежать один від одного. Зміни в цьому підході не передбачаються, однак, у проектах Big Data вони можуть відбутися. Ця методологія не працює для процесу обробки даних. Однак, може бути корисним для планування використання деяких інструментів методології, наприклад, діаграми Ганта. Отже, основним мінусом Waterfall є неохоплення нею управління змінами, що не поєднується з проектами обробки великих даних.

Основною проблемою проектів Data Science взагалі є нерозуміння між ними та цілями бізнесу. У проектах Big Data додаються ще такі специфічні проблеми як величезний об'єм даних та неперервні зміни. З цими проблемами можуть справитися Agile методології. Scrum це один із найпоширеніших у світі підходів Agile, у якому великі проекти поділяються на менші фази, що називаються спринтами і тривають від 1–2 тижнів до 1–3 місяців. Кожен спринт має фіксовані часові рамки і повинен досягти результатів, встановлених на

зустрічах. Scrum значною мірою орієнтований на відгуки клієнтів. Він адаптивний та гнучкий, з великим ступенем автономності, що з точки зору Data Science дозволяє оптимізувати передбачуваність. Проте на відміну від Scrum в Software Engineering, де є постійний інкремент для демонстрації результатів, в Big Data немає матеріалу збору зворотного зв'язку. Вимоги на спринт змінюються. У роботі з даними є ітеративність між фазами підготовки даних (Data preparation) і моделювання (Modeling), де все може змінитися.

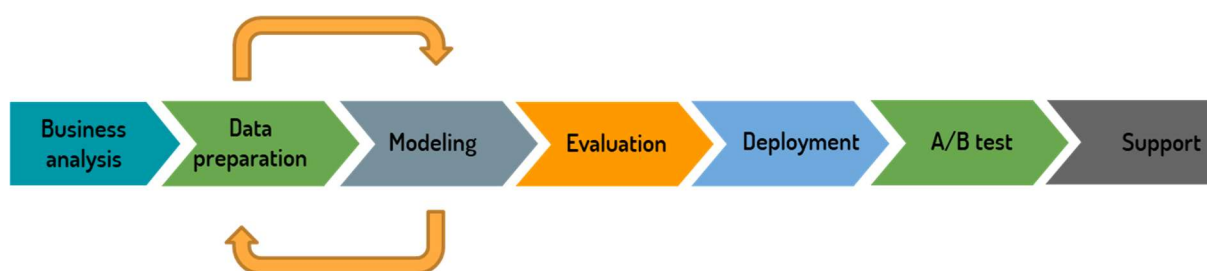


Рисунок 1.5 – Етапи роботи з даними у Big Data проектах в Scrum

Наприклад, з'явилася нова гіпотеза, підготували дані, перевірили гіпотезу, і все, що заплановано до кінця спринту, вже не має ніякої цінності. Все потрібно перероблювати, тому що концепція змінилася. Якщо врахувати таку особливість та трохи змінити класичний підхід Scrum, то він може вдало працювати на проектах Big Data. Безперервний потік величезного обсягу інформації з інтернету, корпоративних систем або приладів потрапляє під визначення Big Data при великих швидкостях завантаження або накопичення. Швидкість реагування та висока частота представлення результату є одним з 12 основних принципів Agile Manifesto.

Розглянемо інший Agile підхід – Kanban. Це методологія, яка використовує дошку як проект, а картки - як завдання. Традиційна дошка канбан включає три колонки - Готово, Виконується та Зробити. Що стосується проектів з обробки даних, можна додати додаткові стовпці: підготовка даних, в розробці, кодування, тестування тощо. Підхід робить більший акцент на незавершеному виробництві без посилення на дати та ролі, робота у ньому сильно візуалізується.

Мінуси методології - відсутність акценту на датах та термінах - можна, навпаки, назвати плюсами для проектів з обробки великих даних.

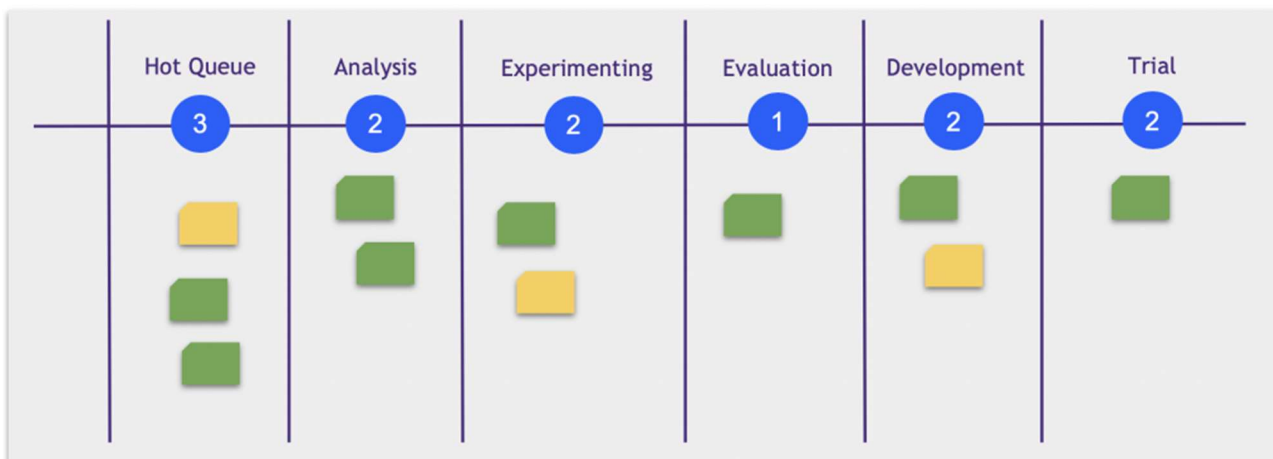


Рисунок 1.6 – Робота над проектами Big Data по Kanban [31]

Зауважте, що також можливе об'єднання декількох фаз в одну, наприклад, моделювання і підготовка даних. Це робиться для більшої зручності: у цій варіації робота з даними та експерименти відбуваються в один і той же момент.

У цьому процесі важливо постаратися вийти в продакшн рано для того, щоб провести тестування. Це складний момент, тому що аналітики даних часто хочуть дотримуватися waterfall-стилю і вважають, що потрібно спочатку відладити модель, а потім переносити її прод-середовище. Цей підхід не дуже добре працює, тому що завжди можна знайти недоліки в роботі: немає даних, не можна перенести їх на прод, недостатньо потужності на серверах і т.д. Якщо це з'ясується в пізній стадії життєвого циклу, то доводиться занадто багато переробляти.

Важко знайти одну методологію, яка б чудово працювала протягом життєвого циклу проекту в галузі Big Data. Проекти з обробки великих даних можуть поєднувати в управлінні ними декілька інструментів різних методологій, і це абсолютно природно. Варто дослідити, які саме інструменти впливають на успіх Big Data проектів та синтезувати їх у єдиний підхід, що зможе позитивно відобразитися на результатах проекту.

Поточна ситуація на ринку змушує планувати роботу таким чином, щоб була можливість вносити зміни під час реалізації проекту, що неможливо в рамках класичної моделі фіксованого обсягу, бюджету та часу, але є ключовим аспектом для проектів великих даних. З різних типів методологій, гнучкі методології виявились найбільш придатними для мінливого середовища. Історично використовувані в ІТ-проектах, гнучкі методології дозволяють змінюватись під час проекту. Саме це робить їх привабливими для використання у проектах Big Data.

Бізнес формулює проблему, мету, завдання та обмеження ресурсів проекту Big Data: терміни, бюджет, вимоги до продукту та показники оцінки корисності. Висока участь бізнесу є необхідною умовою підходу Agile, тому представники бізнесу також беруть участь у підготовці, впровадженні та експлуатації даних. Від бізнесу в команді проекту Big Data є такі профільні ролі:

- експерт - фахівець з предметних областей, який може бути потенційним користувачем розробленого рішення;
- бізнес-аналітик, який разом із експертом та замовником виявляє проблему та формалізує вимоги до продукту для її вирішення, перекладаючи бізнес-побажання на технічну мову для фахівців з обробки даних та розробників, іноді консультуючи щодо використання готової продукції, беручи до уваги враховувати специфіку процесів застосування;
- замовник (спонсор) - особа, відповідальна за походження проекту, що стимулює його реалізацію та вирішення головної бізнес-проблеми, зазвичай фінансує проект та оцінює цінність кінцевих результатів з точки зору його корисності для бізнесу;
- керівник проекту - це куратор усіх етапів проекту, який забезпечує своєчасне та якісне виконання своїх завдань.

Важливо також відзначити навички, якими повинен володіти менеджер проекту великих даних, які поділяються на три блоки:

- бізнес-блок - навички спілкування з клієнтом, виявлення вимог, формування бачення рішення;

- логічний блок - основні концепції науки про дані: статистика, машинне навчання, штучний інтелект;

- технологічний блок - навички, необхідні для керівництва впровадженням моделі як повного компонента.

### **1.3 Використання гнучких методологій для Big Data проектів**

Можна виділити наступні переваги використання Agile у проектах великих даних:

#### **1. Планування та встановлення пріоритетів.**

Починаючи із залучення зацікавлених сторін, методологія Agile надає аналітикам даних змогу визначати пріоритети та створювати дорожні карти на основі вимог та цілей.

Це також дозволяє технічним командам надати зацікавленим сторонам огляд та розуміння загальних витрат, пов'язаних з кожною поставленою метою. Таким чином, весь процес створює кращий взаємозв'язок між аналітиками даних та зацікавленими сторонами, створюючи постійні канали зв'язку.

#### **2. Співвідношення великих даних з технікою.**

Agile - це не тільки робота з програмним забезпеченням та моделями; мова також йде про узгодження аналітиків даних з рештою організації. Часом між інженерами та аналітиками даних можуть виникати розбіжності, коли аналітики чекають на розгортання моделі, тоді як інженери цікавляться, що аналітики будуть робити з прикладними дослідженнями та аналізом даних.

У цьому сценарії Agile запобігає розриву між обома командами, координуючи свої цілі. Причина полягає в тому, що Agile методології мають справу з непередбачуваними реаліями при масштабуванні програм для аналізу вихідних даних.

#### **3. Більше досліджень, менше розвитку.**

На відміну від розробки програмного забезпечення, проекти великих даних не можуть бути написані або розроблені з самого початку, оскільки

неможливо заздалегідь знати найкращі практики для проекту. Як правило, кожен проект у великих даних вимагає від вас різних шляхів та випробування різних методів. Як такі, ці проекти ітеративні, тому Agile, як правило, ідеально підходить для проектів з великими даними.

#### 4. Постійне розгортання моделі.

Коли компанії використовують такі підходи, як постійне впровадження, вони швидко впроваджують нові функціональні можливості програм та виробничі зміни. У традиційних моделях даних це розгортання є багатоетапним процесом, який в кінцевому підсумку переходить до інженерів. Потім інженери переписують та тестують модель даних перед її розгортанням. Весь цей процес займає місяці після первинного впровадження.

З часом компанії зрозуміли, що аналітики даних обмежені потужністю, яку мають локальні машини, і не можуть навчати моделі, які потрібно впроваджувати у виробництво. Використовуючи методики Agile, провідні фірми в даний час будують платформи машинного навчання, які обмінюються навчальними даними для перепідготовки та розгортання на моделях за допомогою API.

#### 5. Створення цінності.

Що стосується планування та створення цінності з необроблених даних до ітеративних прогнозів, команди Big Data можуть отримати допомогу від пірамід цінності даних. Він забезпечує концептуальну основу для створення візуалізації прогресу проекту. Використовуючи піраміди цінності даних, команди великих даних можуть логічно представляти послідовні результати. Ця піраміда цінності даних є однією з особливостей, запропонованих методологією Agile. Таким чином, з кожним життєвим циклом розробка стає кращою.

Ось короткий опис рівнів піраміди та пов'язаних артефактів науки про дані:

- записи: відображення основних даних, з яких слід будувати моделі;
- основні діаграми: можливість узагальнити основні властивості даних;

- діаграми з кореляціями та відносинами: можливість продемонструвати взаємозв'язки та взаємозв'язки між даними;

- моделі, які можуть передбачати речі: можливість робити прогнози, що стосуються ділових проблем за допомогою нових даних.

Складності для Agile у роботі з Big Data.

#### 1. Дослідження та гнучкість.

Впровадження Scrum та Agile може бути не для кожної команди Big Data. Дослідження - це мистецтво та наука, і іноді для цього потрібно більше творчості, ніж строгість процесу розвитку. Дослідження науки про дані вимагає творчих методів вирішення проблем з використанням настанов, а не суворих правил.

Немає правильного способу проведення та управління цим дослідженням, оскільки кожен проект вимагає перевірки різних методів. До того ж цілі не завжди є настільки чіткими. Відповідь на одне питання може призвести лише до нових питань, що може призвести до того, що аналіз триватиме вічно.

#### 2. Синхронізація моделі за участю користувача.

Під час перевірки результатів будь-якого алгоритму зазвичай існує безліч різних рівнів коректності та/або точності. Наприклад, часто легко отримати модель із точністю 70%, 80% або навіть 95%. Однак, щоб отримати точність останніх 5–25%, можуть знадобитися тижні чи місяці роботи.

Agile методологія віддає перевагу реальним рішенням. Проблема полягає в тому, що команда з обробки даних може стримувати спритний розвиток команди розробників програмного забезпечення, оскільки для перегляду можливих моделей потрібно так багато часу. Постійне налаштування моделі іноді може стримувати відчутний прогрес.

Робота з великими даними забезпечує цінність завдяки ідеям та моделям, які вони можуть представляти. Для цього важливо дозволити командам працювати над дослідженнями спільно та ітеративно з усіма зацікавленими сторонами. Якщо витрачається занадто багато часу для того, щоб усі зацікавлені

сторони домовились про те, що таке кінцевий продукт, то продукт ніколи не з'явиться. Тому потрібна Agile методологія.

#### **1.4 Постановка задачі**

Проблема вибору методології ускладнюється тим, що кожна з них має свої переваги та недоліки. Залежно від типу проекту, та чи інша методологія впливатиме на різні критерії успішності проекту. Agile методологія з часом розвивалась, пропонуючи найбільш відповідні методи для кількох областей. Йдеться не лише про впорядкування життєвого циклу розвитку великих даних. Йдеться про узгодження роботи команди з різними зацікавленими сторонами, забезпечення правильного зворотного зв'язку для узгодження з бізнес-цілями.

Поряд із наявністю досліджень щодо впливу гнучких методологій на успіх проекту, недостатньо наукових досліджень щодо впливу окремих складових елементів (методик) гнучких методологій на успіх проекту. Проблема вимагає ретельного вивчення та дослідження.

Відповідно, для того, щоб виявити методики гнучких методологій, які гарантовано матимуть позитивний вплив на успіх проектів великих даних, необхідно враховувати особливості їх планування та розвитку, умови швидко зростаючого обсягу інформації і невизначеність, і визначити, які методи будуть найкраще працювати в проектах цього типу. Для цього пропонується провести емпіричне дослідження щодо впливу методик гнучких методологій на успіх проекту з обробки великих даних. Це допоможе знайти, які з основних Agile методик є найбільш ефективними. Важливо зазначити, що більшість методик методологій дійсно можна використовувати окремо, не втрачаючи своєї корисності.

Дослідження будемо проводити на основі даних опитування реальних менеджерів та учасників команд проектів Big Data, оскільки інших даних поки що недостатньо. Після опитування необхідно побудувати модель регресії. Побудова регресійної моделі дозволяє розрахувати статистичну залежність між

досліджуваними параметрами, а саме між успіхом проекту та використанням різних Agile практик. Можна виділити деякі основні обмеження моделі: модель не пояснює наявність причинно-наслідкового зв'язку, а лише показує наявність статистичного зв'язку, якщо такий є. Для встановлення причин і наслідків результати дослідження інтерпретуються на основі існуючих теоретичних основ щодо даної теми.

Цей підхід дозволить знайти важливі залежності в управлінні проектами великих даних. Визначення методів, що впливають на успіх проекту Big Data, є важливим кроком у розумінні правильної організації роботи з ними, поліпшення якості та швидкості послуг, що надаються бізнесом, а також злагодженої командної роботи.

### **Висновки до Розділу 1**

У цьому розділі було описано принцип використання гнучких методологій у проектній діяльності, успішність та особливості цього підходу, а також перевага над послідовним підходом розробки. Також були наведені особливості в управлінні проектами Big Data та як Agile методології можуть бути застосовані для них. Можна зробити висновок, що у проектах з такою кількістю невизначеності, як у проектах Big Data, застосування гнучких методологій чи принаймні їх методик буде дуже вдалим.

Було також описано особливості Agile-команди у проектах Big Data, та визначено основні принципи за якими ці команди можуть успішно працювати.

У розділі була надана розгорнута постановка задачі, над якою буде надалі вестись робота у кваліфікаційній роботі магістра. Відповідно до неї у подальших розділах необхідно буде виявити основні методики гнучких методологій, провести експертне опитування та оцінити вплив методик на успіх проектів Big Data.

## РОЗДІЛ 2

### ПРАКТИЧНЕ ВИКОРИСТАННЯ ГНУЧКИХ МЕТОДОЛОГІЙ ДЛЯ ПРОЕКТНОГО АНАЛІЗУ

#### 2.1 Відмінні характеристики гнучких методологій управління проектами та визначення їх найпопулярніших методик

##### **Відмінні характеристики методології Scrum.**

Scrum - метод управління проектами, заснований на принципах Agile. Спочатку Scrum був розроблений для галузей, пов'язаних з розробкою ПЗ, проте в останнє десятиліття підхід успішно застосовується в інших областях і індустріях (Wykowski, Wykowska, 2018).

##### **Історія Scrum.**

Метод був вперше представлений Хіротака Такеші і Ікудзіро Нонака в статті The New Product Development Game в журналі Harvard Business Review в 1986 [22]. Автори описують Scrum як систему управління проектом по розробці нового комерційного продукту, яка дозволяє підвищити швидкість і гнучкість процесу реалізації. Термін «Scrum» автори беруть зі спортивної гри в регбі. Правила гри влаштовані таким чином, що команда складається з гравців з різною унікальною роллю, інакше кажучи, створюється кроссфункціональна команда. Завдяки високій ефективності комунікації між гравцями, унікальності умінь кожного і об'єднанню в команду однодумців, які працюють над одним продуктом, процес «гри» (в нашому випадку мова йде про управління проектами), стає максимально швидким і гнучким.

##### **Опис Scrum.**

Scrum спочатку був описаний в більшій мірі як структура для управління проектами, а не як окрема методологія з докладно описаними процесом. Таким чином, Scrum можна використовувати на багатьох проектах з ітеративними або інкрементальними фазами. Кен Швабер, Майк Бідл і Джефф Сазерленд внесли найбільший вклад в розвиток даної методології [23]. За останнє десятиліття

Scrum отримав найбільшу популярність в порівнянні з іншими методологіями, тому що разом з лаконічністю і простотою пропонованих правил, методологія успішно співпрацює з іншими методологіями, як гнучкими, так і класичними.

Ключовими факторами, необхідними для успішного використання Scrum, є:

1. Прозорість - процес повинен бути повністю доступний для всіх учасників проекту.
2. Незалежне обстеження - скрам-майстри повинні часто обстежувати артефакти спринту, щоб попереджати проблеми і ризики в проекті.
3. Адаптивність - оскільки незалежний спостерігач буде знаходити «сірі зони», які не проаналізовані і не включені в обсяг робіт на першому етапі, процес повинен бути гнучким і самі команди повинні гнучко підходити до прийняття таких запитів на зміну.

До основоположних правил методу можна віднести [24]:

- поділ організації на невеликі, крос-функціональні команди, що самоорганізуються;
- поділ проекту на так званій «беклог» - список невеликих, самостійних функціональних бізнес-процесів або підпроцесів, що мають соціальну значимість з точки зору бізнесу; позиції списку упорядковано відповідно до пріоритету важливості для бізнесу і трудомісткості розробки і реалізації;
- поділ часу проекту на короткі етапи, які не перевищують одного місяця; по закінченню кожного міні-етапу, званого «спринтом», відбувається демонстрація, в ході якого користувач отримує реалізовану одну або кілька одиниць беклога і дає команді розробки перший зворотний зв'язок;
- оптимізація плану релізу і уточнення пріоритетів зі стейкхолдерами за результатами минулих ітерацій;
- оптимізація процесу шляхом проведення ретроспективи - церемонії, при якій команда обговорює аспекти, пов'язані з роботою всередині команди, а

саме: ефективність взаємодії, виявлені проблеми в комунікації і шляхи їх поліпшення.

Процес реалізації частини Беклога в рамках спринтів зображений на рисунку 2:

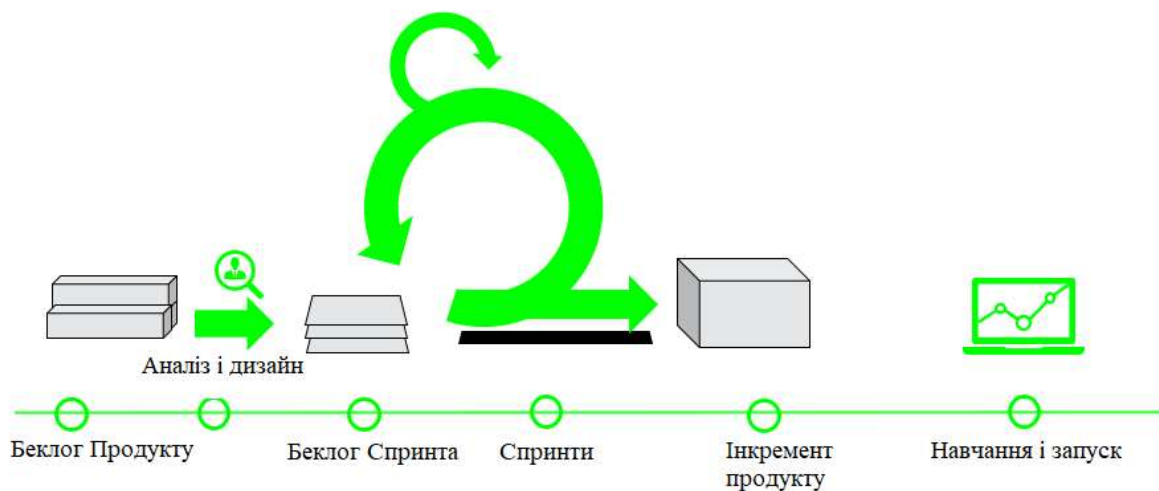


Рисунок 2.1 - Процес розробки за методологією Scrum

### **Обмеження Scrum.**

З одного боку, підхід Scrum виграє перед іншими методологіями за рахунок своїх переваг, описаних вище. Однак, важливо пам'ятати про недоліки або обмеження методу. Частина з них була описана в статтях Акіфа і Маджид [25] і Турка, Франса і РУМП [26]. Метод не працює ефективно в наступних випадках:

1. Продукт тісно пов'язаний з legacy-системами або зобов'язаний пройти перевірку Відділу технічного контролю. У Scrum методі частина продукту повинна бути реалізована і протестована повністю в рамках однієї ітерації. Якщо продукт має на увазі тривале інформаційне або технічне тестування, застосування класичних методологій буде більш виправдано.

2. Продукт володіє безліччю зовнішніх залежностей. Як вже було зазначено, функціонал розробленого продукту в рамках одного спринту повинен бути незалежним від інших спринтів або проектів в компанії. Наявність глибоких зовнішніх взаємозв'язків ускладнює успішну реалізацію продукту в

установлений строк, функціонал може виявитися неповним через недостатнє аналізу та обліку взаємозв'язків.

3. Учасники команди мають вузько-спеціалізованими навичками і вміннями. У Scrum-командах учасники дійсно мають унікальну специфікацією (аналітики, розробники, тестувальники і т.д.), однак їх функції повинні бути частково замінними, оскільки процес роботи по Scrum має на увазі тісну співпрацю учасників, в тому числі і виконання завдань поза своєю спеціалізацією. Якщо в команді потрібні вузько-кваліфіковані учасники, то, як вони, так і інша команда повинні бути готові швидко освоїти і вміти застосувати інший навик.

4. Команда географічно розподілена по кільком локаціях. Scrum підхід має на увазі тісну комунікацію і постійна взаємодія між учасниками. Останні досягнення в інструментах управління проектами допомогли скоротити негативний вплив даного бар'єру (наприклад, створення ПЗ, що дозволяє створити онлайн-дошку зі Scrum-завданнями). Однак, автори Маніфесту стверджують, що спілкування віч-на-віч і раніше найчастіше вдаються для даного підходу.

#### **Відмінні характеристики методології Kanban.**

Kanban є іншим методом з управління проектами розробки. У цьому методі фокус полягає у виконанні завдань «точно в строк».

#### **Історія Kanban.**

Вперше метод був застосований на фабриці «Тойота» в 1962 році і є однією з підсистем виробничої системи «Тойоти» (Toyota Production System). Таичи Воно, інженер, розробив Kanban для збільшення продуктивності на заводі по зборці автомобілів. Назва системи управління походить від системи організації на дошці карток, які фіксують переміщення продукції на заводі. При цьому кожна картка була сигналом, скільки і якого матеріалу потрібно на наступній стадії або в наступному відділі. Система дозволила ефективніше контролювати процес складання автомобілів, уникнути простоїв або нестачі

матеріалів на кожному етапі, скоротити життєвий цикл реалізації продукту і здешевити процес складання при збереженні того ж рівня якості.

### **Опис Kanban.**

Такий ефект досягається за рахунок правильної пріоритизації завдань і визначення тривалості кожного завдання в загальному потоці [27]. Методологія дозволяє, з одного боку, приділити найбільшу увагу завданням з високою ймовірністю появи ризику, з іншого боку, - оперативно управляти іншими завданнями проекту. В універсальному сенсі підхід Kanban фокусується на тому, щоб виконати потрібну задачу в необхідний термін з урахуванням можливостей і кваліфікації співробітників і доступності необхідних ресурсів. Проект починається з поділу обсягу на так звані компоненти - частини фіксованого обсягу, згруповані за ознакою бізнес-цінності [27].

Принципи Kanban полягають в наступному:

1. Візуалізація потоку завдань дозволяє учасникам в будь-який момент часу бачити цілком прогрес проекту і статус завдань.

2. Обмеження за кількістю завдань на одній стадії дозволяє скоротити кількість паралельно виконуваної роботи і прискорити час пересування завдання від першого статусу до останнього. Ті завдання, які не потрапили під допустиме обмеження, повинні бути пріоритезовані з точки зору виконання термінів і збільшення ефективності процесу.

3. Контроль потоку завдань допомагає стежити за завантаженням ресурсів. При цьому візуалізація показує «вузьке горло» проекту і сигналізує команді, чиї ресурси при необхідності можна перевести на іншу задачу, щоб відновити необхідну швидкість виконання завдань.

4. Збільшення «пропускної спроможності» команди відбувається за рахунок збільшення ефективності взаємодії. На регулярних зустрічах відбувається аналіз стану дошки, обчислення середньої швидкості виконань завдання і обговорюються шляхи збільшення цієї швидкості. За підсумками зустрічей кожен учасник розуміє, як можна швидше і ефективніше виконати свій пул завдань.

5. Визначення обсягу роботи і очікувань на початку проекту дозволяє уніфікувати правила взаємодії гравців на етапі планування і розбиття проекту за завданнями. Передумовами використання методу є фіксований термін, обсяг і бюджет, що дозволяє прогнозувати ресурси, завдання та обмеження за обсягом завдань на кожному етапі, тобто основні параметри системи управління проектом.

6. Підвищення якості роботи під час проекту досягаються за рахунок загальних експериментів команди. Під час проекту учасники, регулярно аналізуючи прогрес проекту, пропонують методи щодо підвищення ефективності і швидкості реалізації продукту. Саме візуалізація всього проекту в одному місці дає можливість учасникам команди домогтися синергетичного ефекту змін в потоках завдань різних фахівців або відділів.

#### **Обмеження Kanban.**

Автори виділяють такі причини можливого неуспішного використання підходу Kanban [28]:

1. Нестабільний попит. Оскільки терміни і обсяг завдань в методі складаються на основі прогнозованого попиту, при відсутності коректної інформації вся організація роботи може виявитися невірною, що призведе до зриву термінів або невиконання потрібного обсягу.

2. Нестабільність часу виконання завдання. Від кожного завдання залежить термін виконання інших завдань або проходження етапу в цілому. При різному часі виконання завдання проектна команда не може комплексно аналізувати ефективність процесу і шукати шляхи його прискорення.

3. Не стандартизовані операції. Як і в описаному вище випадку, не уніфіковані операції перешкоджають плануванню термінів і обсягу виконання того чи іншого завдання.

4. Складність забезпечення узгодженості між різними етапами продукції / різними відділами. Система Kanban має на увазі можливість швидкої передачі матеріалу між різними етапами життєвого циклу продукту. У разі, коли вся організація не живе по системі Kanban, проект буде зазнавати труднощів

управління часом, наприклад, коли необхідну частину продукту можна отримати тільки з верстата, який запускається раз в два місяці зважаючи на складність налаштування.

5. Невизначеність поставки «вступних» матеріалів. При відсутності стабільності подачі необхідних для виконання проекту даних, матеріалів або ресурсів, виконати роботу з урахуванням прогнозованого попиту не представляється можливим.

6. Наявність жорстких часових рамок. Незважаючи на установку «виконання роботи в строк», підхід Kanban не має на увазі наявність термінів для кожного завдання і суворого терміну для кінцевого продукту, оскільки використовується прогнозований попит.

7. Низька комунікація між учасниками або непоінформованість про чужі функції. Підхід будується на тому, що відповідальні за певний етап створення продукту ефективно комунікують між собою і в разі виникнення позапланових ситуацій можуть спільно вжити заходів щодо усунення проблеми.

### **Відмінні характеристики методології Екстремальне програмування (Extreme Programming, XP).**

На відміну від описаних вище методологій, дана методологія призначена тільки для проектів з розробки ПЗ. Метою її застосування є підвищення продуктивності і якості розробки. Запити на зміни вважаються невід'ємною частиною проекту, тому система має на увазі ітеративний підхід. Таким чином, з одного боку, підвищується довіра замовника до команди розробки, оскільки після кожної ітерації демонструється версія готового, функціонуючого продукту, і з іншого боку, команда розробки з кожною ітерацією може збільшувати якість вже розробленого програмного продукту.

### **Історія XP.**

Історія методології починається з Кента Бека, Уорда Каннінгема і Мартіна Фаулера [29]. Команда була найнята з метою оптимізації процесу розробки в проекті по оплаті праці в компанії Chrysler Comprehensive Compensation System (C3). Продукт поставлявся невчасно і низької якості,

потрібні були термінові зміни управління проектом. Нововведення в системі управління почалося з того, що замість встановленого терміну об'єднання і завантаження коду раз в три тижні команді розробників було запропоновано самим визначити дату поновлення, виходячи з потреб проекту.

### **Опис XP.**

Методологія Екстремального програмування заснована на п'яти принципах: простота, готовність до критики, комунікація, повага і сміливість. Правила Екстремального програмування складаються з 29 правил, розділених на чотири пункти - планування, управління, тестування, проектування і розробка коду. Необхідно відзначити, що багато методик методології повторюють методики підходу Scrum, однак процес реалізації проекту відрізняється.

Для цілей роботи є корисним вивчити основні 12 практик Екстремального програмування, розділені на чотири блоки:

1. Короткий цикл отримання зворотного зв'язку:
  - гра в планування;
  - парне програмування;
  - розробка через тестування;
  - швидкий зворотний зв'язок.
2. Безперервний процес розробки:
  - неперервна інтеграція;
  - рефакторинг;
  - невеликі, але часті релізи.
3. Розуміння, що поділяється усіма:
  - стандарти написання коду;
  - колективне володіння кодом;
  - простий дизайн;
  - системна метафора.
4. Благополуччя розробника:
  - фіксований робочий день, відсутність перепрацювань.

Процес розробки за методологією зображений на рисунку 2.2:

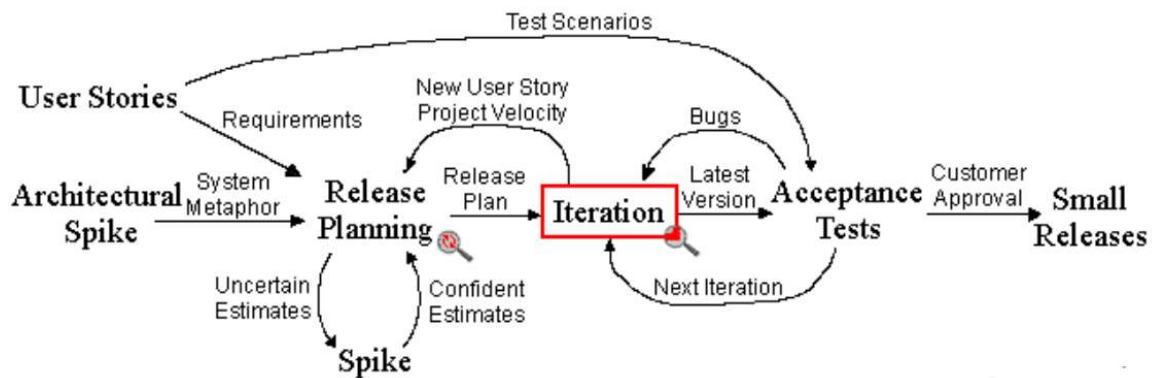


Рисунок 2.2 - Процес управління проектом по методології XP [55]

### Обмеження XP.

Незважаючи на доведену ефективність, існує критика деяких аспектів даного підходу. Істотна частина критики відноситься до Agile методів в цілому, наприклад, вимога залученості учасників, високий рівень кваліфікації кожного розробника [30]. До унікальних обмежень структури XP можна віднести наступні [55]:

1. Підхід веде до «розростання» початкового об'єму вимог (score creep), що призводить до невиконання термінів і / або бюджету проекту.
2. Вимоги до системи представлені у вигляді автоматичних acceptance test (вид тестів для фінального тестування продукту перед запуском), а не у вигляді специфіковані узгоджених документів.
3. З огляду на коротких ітерацій і необхідності мати працюючий функціонал по закінченню терміну, розробники будують свою роботу таким чином, щоб побудувати програмний продукт, що покриває максимально можливим функціонал мінімальними зусиллями розробника. При цьому код ускладняється і доповнюється тільки в разі не пройдених acceptance test. Цим система нагадує підхід code-and-fix, який веде до великих трудових витрат, а не лише невелика зміна функціоналу за запитом від замовника після демонстрації, як в методології Scrum. Іншими словами, одна і та ж область програмного продукту може бути розроблена кілька разів, що значно довше, ніж проходження заздалегідь узгодженим планом.

4. Представник замовника може стати «вузьким горлечком» на проекті. З одного боку, без належного навчання, він / вона буде змушений / а працювати в незнайомому середовищі, що складається з розробників, що володіють в основному технічними вміннями. З іншого боку, є ризик мікро-менеджменту команди розробки з боку представника бізнесу, при цьому кваліфікації замовника не буде вистачати для ефективного розподілу ресурсів і роботи з розробки ПЗ. Більш того, вимога щодо постійного залучення в проект кінцевого користувача або представника бізнесу істотно збільшує витрати проекту.

#### **Відмінні характеристик гібридних методологій.**

Далі будуть описані гібридні методології ScrumBan і XP Hybrid, які показали свою ефективність і використовуються в тій же мірі, як і інші гнучкі методології.

#### **ScrumBan.**

Підхід ScrumBan є гнучкою методологією управління проектами, заснованою на підходах Scrum і Kanban. Спочатку система була вироблена як перехідна між Scrum і Kanban для команд, що працюють по Scrum і бажаючих освоїти концепти методологій Lean і Kanban [25]. Від «перехідної» методології ScrumBan був розвинений до повноцінного підходу. Згідно з концепцією, команда, що працює по методи Scrum, використовує Kanban техніки для регулярної оцінки ефективності своєї діяльності і постійного збільшення продуктивності. Приклад візуалізації процесу зображений на рисунку 2.3:

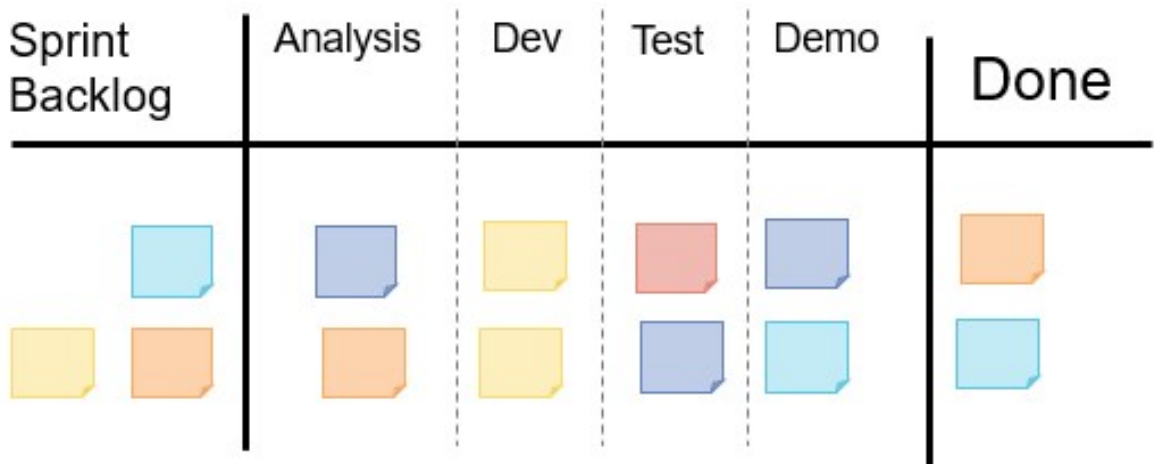


Рисунок 2.3 - Приклад Kanban-дошки для гібридної моделі

До основних правил методології відносяться [32]:

1. Короткі ітерації тривалістю в кілька тижнів, не більше трьох.
2. Планування здійснюється на вимогу на підставі того, скільки позицій на дошці завдань знаходяться в статусі «Зробити».
3. Пріоритизація завдань проходить кожну сесію планування.
4. Планування «по кошиках» (Bucket size planning) має на увазі наявність трьох видів завдань:
  - довгострокові (1 рік) присвячені стратегічним цілям компанії, наприклад, запуск нового продукту, освоєння нового ринку і т.д.;
  - середньострокові (6 місяців) включають в себе ті довгострокові завдання або проекти, за якими вже складено детальний план;
  - короткострокові (3 місяці) складені із завдань нижчого рівня, реалізація яких запланована на найближчий час і до виконання яких проектна команда вже може приступити.
5. Для детального планування використовується Канбан-дошка зі статусами «Зробити», «В роботі» і «Виконано». При необхідності статус «В роботі» може бути розділений на більш дрібні підетапи. Як і в методології Kanban, статус «В роботі» має обмеження за кількістю завдань.
6. Немає жорстких вимог до складу команди проекту, при цьому, на відміну від системи Scrum, команди можуть бути менш крос-функціональними.
7. Завдання призначаються не керівником команди і не менеджером проекту, а самою командою. Кожен учасник вибирає, яке завдання він візьме зі списку «Зробити». Такий підхід гарантує, що учасники будуть рівномірно завантажені і процес буде проходити рівно і ефективно. Однак, такий принцип працює тільки в разі сильної залученості кожного учасника команди і відсутності «проблеми безбілетника».
8. «Замороження функцій» (Feature freeze) проходить за деякий час до настання дедлайну проекту. Це означає, що ніякі додаткові функції не можуть бути розроблені або ніякі додаткові окремі бізнес-процеси не можуть бути взяті в роботу. Доробки здійснюються тільки по незакінченим завданням.

9. Сортування залишившихся задач (Triage) здійснюється після «замороження функцій». На цьому етапі проектний менеджер вирішує, які завдання повинні бути завершені в першу чергу, а по яким робота може бути не завершена. Це гарантує, що найбільш важливі і необхідні елементи будуть завершені вчасно.

### **XP Hybrid Scrum.**

Підхід XP Hybrid Scrum складно назвати повноцінною методологією з урахуванням наявної літератури. В даний час система являє собою набір практик, які використовуються спочатку в методології Scrum і XP. Необхідно зауважити, що підходи гармонійно і ефективно доповнюють один одного, оскільки Scrum - підхід до управління проектом в цілому, визначає рамки управління термінами, командою і роботами, в той же час XP - система управління розробкою програмного продукту.

Не існує узгодженої формули комбінації між практиками Scrum і XP. Один з підходів був запропонований Муштаком [33]. Полягає він у тому, що загальні рамки проекту - церемонії, артефакти, ролі - визначаються методологією Scrum. В рамках спринту розробка здійснюється за правилами методології XP: з використанням парного програмування, розробки через тестування, високого залучення користувача, рефакторінга і т.д. Такий метод дозволяє, з одного боку, подолати «слабкі місця» методу XP, а саме: управління проектом середніх і великих розмірів, недолік практик менеджменту (фокус на процесі розробки) і залежність від відкриття користувача, що підвищує ризик невдачі проекту. З іншого боку, даний підхід вирішує проблему застосування «чистого» Scrum - процес розробки недостатньо детально організований і не враховуються специфічні риси програмного продукту.

Як підсумок можна визначити список методик, властивих найбільш популярним гнучкими методологіями (див. таблицю 2.1):

Таблиця 2.1 – Основні методики гнучких методологій

Методологія	Список методик
Scrum	<ul style="list-style-type: none"> <li>- беклог проекту</li> <li>- користувацькі історії</li> <li>- спринт</li> <li>- стенд-ап</li> <li>- планування спринту</li> <li>- рефайнмент сесія</li> <li>- оцінка КІ в балах</li> <li>- Scrum-ролі</li> <li>- ретроспектива</li> </ul>
Kanban	<ul style="list-style-type: none"> <li>- канбан-дошка</li> <li>- обмеження за кількістю завдань на одній стадії</li> <li>- визначення обсягу роботи і очікувань на початку проекту</li> </ul>
XP	<ul style="list-style-type: none"> <li>- гра в планування</li> <li>- парне програмування</li> <li>- розробка через тестування</li> <li>- швидкий зворотний зв'язок</li> <li>- неперервна інтеграція</li> <li>- рефакторинг</li> <li>- невеликі, але часті релізи</li> <li>- стандарти написання коду</li> <li>- колективне володіння кодом</li> <li>- простий дизайн</li> <li>- системна метафора</li> <li>- фіксований робочий день, відсутність переробок</li> </ul>

Методологія	Список методик
ScrumBan	<ul style="list-style-type: none"> <li>- короткі ітерації (1-2 тиж.)</li> <li>- планування на вимогу</li> <li>- планування «по кошиках»</li> <li>- «замороження функцій»</li> <li>- сортування залишившихся задач</li> </ul>
XP Hybrid Scrum	- поєднання елементів Scrum і Kanban

## 2.2 Практичне визначення поняття та критеріїв успіху проекту

Інтенсивний розвиток вивчення того, що варто включати в поняття успіху проекту обумовлена тим, що тема несе не тільки академічне значення, а й практичну значимість. Щоб виміряти і порівняти ефективність використання тієї чи іншої методології, необхідно визначити показники, за якими буде в підсумку оцінений внесок в успіх проекту. Видається очевидним, що для цього необхідно зрозуміти, що входить в поняття «успіх».

У літературі існує безліч інтерпретацій даного визначення, що веде до різних використовуваних індикаторів успіху. При цьому вчені до цих пір не прийшли до консенсуса про те, як вимірювати успіх проекту [34]. Відомо, що в проектах різних типів будуть різні найбільш важливі критерії успіху [35, 36, 37]. Критерії успіху можуть варіюватися в залежності від типу проекту, його складності, фази життєвого циклу, індустрії, національності, організаційної структури [38, 39]. Також варто відзначити, що різні стейкхолдери мають різне уявлення і очікування щодо успішності проекту, оскільки можуть мати різне бачення критеріїв успіху та результативності [40].

## **Розвиток поняття «успіх проекту» до 2000-х.**

Традиційний підхід до вивчення успіху проекту фокусується на тому, що проект повинен бути виконаний в установленій якості і/або обсязі, в запланований термін і в рамках запланованого бюджету, тобто в рамках «проектного трикутника», або «залізного трикутника» [41, 42]. Однак, в період 1980-1990-х років акцент в дослідженнях був на розвиток системи або схеми успіху проекту з включенням не тільки відчутних параметрів, але і не відчутних, більш складно піддаються обчисленню [38]. В цей час обмежене академічне поняття успіху проекту не відповідало корпоративній реальності, з якими стикалися проектні менеджери. В аналіз були включені показники таких вимірів, як задоволеність клієнта і довгострокові результати проекту [43, 44]. Так, цілі та результати проекту повинні підтримувати стратегічні цілі компанії. Шенар і ін. (2001) пишуть, що тепер проектні менеджери стають новими стратегічними лідерами, хто повинен нести повну відповідальність за результати проекту. Є кілька прикладів проектів в ІТ та інших індустріях, в яких цілі проекту були досягнуті, проте в довгостроковому періоді проект привів до великих втрат в контексті всієї організації [35]. Іншими словами, успішний менеджмент може привести до успіху проекту, проте, проект може провалитися незалежно від проектного менеджменту [45].

У період 1999-2000-х в аналіз були включені такі поняття, як внесок в бізнес стратегію і розвиток особистісний, команди і організації в цілому [38].

Підводячи підсумки періоду, можна зробити висновок, що в цей часовий проміжок формується база поняття успіху проекту і визначається напрямок майбутнього розвитку теорії, а саме включення в аналіз успіху проекту показників: бенефіти для стейкхолдерів і організації в цілому, стратегічні цілі організації та її клієнтів і вигоди і внесок в розвиток співробітників компанії, починаючи з проектною командою.

## **Сучасне поняття «успіх проекту».**

Важливо відзначити, що в періоді з 2001 до 2005 критерії успіху проекту визначалися на теоретичному рівні і мала кількість теорій було перевірено на

практиці [44, 37, 46]. Після 2015 року більшість сформованих систем визначення успіху проекту, підтверджених емпіричними даними, будувалося в певному концепті. Це говорить про те, що проекти повинні бути вивчені в контексті [47; 45].

Щоб подолати рамки класичного «трикутника» проекту, дослідники популяризували і розвивали теорії, що сфокусовані на вигоді, яку проект приніс стейкхолдерам і організації. Успіх проекту може бути визначений його ефективністю і результативністю [47]. Виконання термінів, бюджету та обсягу показують ефективність проекту, в той час як результативність може бути виміряна за допомогою бенефітів для стейкхолдерів і організації. Про важливість даних вимірювань писав ще Аткинсон (1999), один з перших дослідників, що займаються концепцією успіху проекту. Висновки дослідження не мали емпіричного обґрунтування, однак уже було готово пояснення того, чому деякі проекти, виконані в строк і в рамках бюджету, закінчувалися невдачею.

Незважаючи на те, що більшість вимірів проекту в тій чи іншій мірі зачіпає вимірювання проекту, пов'язані з часом, обсягом і бюджетом, одним з найбільш важливих критеріїв успіху проекту є виконання очікувань замовника [37]. Баккаріні (1999) розділяє успіх проекту на два виміри - успіх проектного менеджера (термін, бюджет, обсяг) і успіх продукту - ефект від результату проекту. Він пропонує нову структуру, в якій є вхідні і підсумкові дані для успіху управління проектом і цілі - для успіху продукту проекту.

Шенар (2001) групує показники проекту по чотирьом вимірам:

- 1) ефективність проекту;
- 2) вплив на замовника;
- 3) прямий ефект на успіх організації;
- 4) стратегічна цінність.

Вибір параметрів успіху проекту залежить від типу проекту. Для проектів з низьким рівнем невизначеності, де важлива ефективність, успіх в більшій мірі визначається часом і бюджетом. Коли рівень невизначеності високий і низька

результативність в короткостроковому періоді може бути компенсована довгостроковими бенефітами, інші критерії успіху виявляться доречними. Структура, запропонована автором, була емпірично перевірена і використана в подальших роботах [47].

Інша концепція була запропонована в роботі Діалло і Туїлльєра (2004). Автори розділили критерії успіху на три виміри:

- 1) профіль, ймовірність отримання додаткового фінансування, отриманий результат і репутація;
- 2) бюджет, терміни і цілі;
- 3) тривалість, внесок, задоволеність.

Автори спробували вивести універсальне правило, яке перевірили і підтвердили на декількох проектах, що відрізняються за секторами і рівнем завершеності, в країнах Африки.

Двір і ін. (2003) знаходять чотири складових успіху проекту:

- 1) виконання плану;
- 2) бенефіти для кінцевого користувача;
- 3) бенефіти для підрядника;
- 4) успіх проекту в цілому.

При цьому автори звертають увагу, що проекти, що сприймаються як успішні, повинні бути успішними з точки зору стейкхолдерів.

Починаючи з 2005 року, фокус в академічній літературі змінюється з теоретичних робіт по визначенню успіху проекту на емпіричні роботи, які вивчають критерії успіху в контексті певного проектного типу. Необхідно відзначити, що різні автори в своїх роботах використовували різні поняття «тип проекту» в залежності від його тривалості, галузі, рівня завершеності і т.д. Мое і Канг (2008) тестують нову модель визначення успіху проекту в контексті міжнародних проектів по нерухомості, компілюють різні критерії успіху для різних стадій життєвого циклу проекту і коригують критерії на індивідуальні особливості секторів і властивостей замовника. Аль-Тмімі і ін. (2011) тестують

концепцію успіху проекту в залежності від успіху менеджменту, проекту і ринку для будівельних проектів в Малайзії.

У недавньому дослідженні Кан і ін. (2013) розвивають і успішно перевіряють модель чинників успіху проектів публічного сектора в Пакистані з урахуванням аналізу академічної літератури за останні 40 років. Їх модель пропонує баланс між «жорсткими» і «м'якими» чинниками і критеріями успіху проекту. У підсумку автори використовують такі вимірювання:

- 1) ефективність проекту;
- 2) бенефіти для організації;
- 3) внесок проекту;
- 4) задоволеність замовника;
- 5) майбутній потенціал.

У кількісному дослідженні Панкратц і Бастен (2014 року) провели інтерв'ю з 11 менеджерами, що займаються впровадженням інформаційних систем. Автори виділили вісім критеріїв успіху проекту:

- 1) відповідність бюджету;
- 2) відповідність термінів;
- 3) виконання вимог по функціональності;
- 4) виконання бізнес-вимог;
- 5) ефективність проекту;
- 6) задоволеність замовника;
- 7) задоволеність підрядника;
- 8) використання замовником кінцевого продукту.

Світ і Піннінгтон (2014 року) адаптували модель Шенар і ін. (2001), додавши в модель новий вимір - вплив на команду. Далі автори протестували модель на проектноорієнтованих організаціях в Арабських Еміратах. Ця структура була використана в роботі Мартенса і Карвало (2016) з включенням нового критерію - стійкості.

До сих пір автори не прийшли до єдиної думки про те, що становить вичерпний список факторів успіху проекту. Більш того, більшість робіт

засновані на теоретичних дослідженнях, а не на емпіричних [42]. Незважаючи на те, що єдиної думки немає, автори погодилися, що вивчати і доповнювати ці фактори критично. Кан і ін. (2013) пишуть, що в залежності від того, що сама організація розуміє під успіхом проекту, залежить успішність проектів. Визначення критерію результату проекту як важливого підвищує ймовірність, що до кінця проекту цей критерій буде виконаний [39]. У той же час формально певний критерій успіху підвищує ефективність розподілу ресурсів і покращує результат [49]. Аналіз проектного успіху також привносить цінність в управління знаннями в проектній діяльності компанії [50]. Саме з цієї причини дослідники прагнуть розробити теоретичний концепт, який згодом запропонує найбільш корисні поліпшення для методології управління проектом. Такий аналіз дозволяє зрозуміти, як, наприклад, особистість менеджера [51], трансформаційне лідерство [52], задоволеність роботою і довіра впливають на успіх проекту. Крім цього, дослідження в області допомагає зрозуміти, що стоїть за провалом проекту [35, 20]. Саме з цієї причини дослідники прагнуть розробити теоретичний концепт, який згодом запропонує найбільш корисні поліпшення для методології управління проектом. Крім цього, дослідження в області допомагає зрозуміти, що стоїть за провалом проекту [35, 20].

Загальна модель визначення успіху проекту повинна пропонувати менеджменту керівництво те, на які критерії необхідно звернути увагу, щоб забезпечити успіх проекту. Наявність різних підходів веде до неузгодженості аналізів і може привести до протилежних результатів [34]. Крім того, не обґрунтовано вимоглива і велика система критеріїв успіху проекту може привести до завищених очікувань з боку стейкхолдерів, що веде до конфлікту між замовником і менеджментом проекту [53].

Як підсумок, можна сформулювати базу критеріїв успішного проекту, що визначили автори, та вивести загальну таблицю критеріїв успіху проекту:

Таблиця 2.2 – Критерії успішного проекту за аналізом літературних джерел

Автори	Критерії
Аткінсон, Керзнер	<ul style="list-style-type: none"> <li>- проект повинен бути виконаний в рамках «проектного трикутника»</li> <li>- бенефіти для стейкхолдерів і організації</li> </ul>
Бакаріні, Шенар, Двір	<ul style="list-style-type: none"> <li>- задоволеність клієнта</li> <li>- довгострокові результати проекту</li> </ul>
Бакаріні	<ul style="list-style-type: none"> <li>- виконання очікувань замовника</li> </ul>
Шенар	<ul style="list-style-type: none"> <li>- ефективність проекту</li> <li>- вплив на замовника</li> <li>- прямий ефект на успіх організації</li> <li>- стратегічна цінність</li> </ul>
Мюллер, Югдев	<ul style="list-style-type: none"> <li>- внесок в бізнес стратегію</li> <li>- розвиток особистісний, команди і організації в цілому</li> </ul>
Діалло, Туїлльєра	<ul style="list-style-type: none"> <li>- профіль, ймовірність отримання додаткового фінансування, отриманий результат і репутація</li> <li>- бюджет, терміни і цілі</li> <li>- тривалість, внесок, задоволеність</li> </ul>
Двір, Раз, Шенар	<ul style="list-style-type: none"> <li>- виконання плану</li> <li>- бенефіти для кінцевого користувача</li> <li>- бенефіти для підрядника</li> <li>- успіх проекту в цілому</li> </ul>
Кан, Чі, Чу, Пісарски	<ul style="list-style-type: none"> <li>- ефективність проекту</li> <li>- бенефіти для організації</li> <li>- внесок проекту</li> <li>- задоволеність замовника</li> <li>- майбутній потенціал</li> </ul>

Автори	Критерії
Панкратц, Бастен	<ul style="list-style-type: none"> <li>- відповідність бюджету</li> <li>- відповідність термінів</li> <li>- виконання вимог по функціональності</li> <li>- виконання бізнес-вимог</li> <li>- ефективність проекту</li> <li>- задоволеність замовника</li> <li>- задоволеність підрядника</li> <li>- використання замовником кінцевого продукту</li> </ul>
Світ, Піннінгтон	- вплив на команду
Мартенса, Карвало	- стійкість

### **Вплив використання гнучких методологій на успіх проекту Big Data.**

Видається очевидним, що при наявності такої кількості різних підходів в менеджменті для визначення критеріїв успіху проекту, для аналізу впливу гнучких методологій підхід необхідно уточнити на підставі релевантної академічної літератури. Спочатку гнучкі методології застосовувалися в проектах по розробці програмних продуктів. З цієї причини вплив використання методології на успіх проекту в цій галузі найбільш докладно вивчено, чого не можна сказати про використання методологій у проектах Big Data. Однак навіть в рамках однієї галузі існує широка варіативність у визначенні успіху проекту по розробці ПО [54]. Розглянемо дослідження визначень критеріїв успішності проектів на прикладі індустрії розробки програмного забезпечення, оскільки насправді для проектів великих даних цей аспект все ще недостатньо вивчений.

Для подальшого аналізу необхідно виявити, на які характеристики успіху проекту звертали увагу автори, коли досліджували вплив гнучких методологій на успіх проекту. У першу чергу, основу критеріїв успіху проекту становить «проектний трикутник» [54, 20].

Робота Чой і Чао [20] є однією з основних робіт у визначенні критеріїв успіху проекту в проектах по розробці ПЗ з використанням гнучких методологій. У роботі автори досліджували як фактори успіху, так і чинники провалу проекту. Автори визначили 19 чинників провалу і 36 факторів успіху. Успіх проекту визначається за чотирма вимірами: обсяг (виконання всіх вимог проекту), час (здача проекту вчасно), вартість (дотримання рамок бюджету і плану) і якість (наприклад, задовільний кінцевий продукт або результат проекту). Фактори провалу згруповані так само за чотирма категоріями: організація, люди, процес і технологія.

При використанні гнучких методологій виявляється критичним як управління результативністю (*performance management*), так і орієнтація на взаємини в команді і соціальний контекст. Згідно Шефіді і Лемат (2013), практики, що використовуються при розробці ПЗ, повинні збігатися з методологією менеджменту і процесами управління в компанії (управління проектами, бізнес-аналіз, бізнес-процеси, вибудовування ІТ) на різних рівнях (розробка ПО, проект, програма і т.д.) [21]. Автори використовують для цього трьохрівневу ієрархічну модель - технічна, організаційна та зовнішня частини.

Шефід і Лемат [21] виділили наступні фактори успіху проекту, пов'язані з оточенням проекту: організаційна культура (консервативна або підприємницька), стабільність організації та індекс дистанції влади. Серед факторів оточення проекту автори виділяють гнучкість, підтримувану топ-менеджментом компанії, рівень підприємництва і схильність до ризику. Таким чином, фактори відображають корпоративну культуру організації. Проектний фактор, виражений в посиленні проектної команди, включає в себе рівень гнучкості замовника, близьке співробітництво з замовником і повноваження, передбачувані процедурою.

Автори пишуть, що бар'єри для успішного використання гнучких методологій - соціальні і культурні фактори, а не технічні, які визначаються завданнями проекту. В успішних проектах гнучкість розробки програмного забезпечення узгоджується з факторами проекту і проектного оточення. Іншими

словами, не можна застосовувати один і той же підхід до всіх проектних команд і проектів.

Станкович та ін. (2013) своєму дослідженні виявили три фактори, що впливають на успіх проекту, - визначення процесу в проекті, природа проекту, тип проекту.

Таким чином, підсумуємо результати літературного аналізу. У наступній таблиці 2.3 наведено критерії успіху проекту, на які мають вплив гнучкі методології управління проектами:

Таблиця 2.3 – Критерії, на які мають вплив гнучкі методології

Автори	Критерії
Чой, Чао	<ul style="list-style-type: none"> <li>- сильна підтримка виконавчої влади</li> <li>- доручений спонсор або менеджер</li> <li>- кооперативна організаційна культура замість ієрархічної</li> <li>- усна культура, що надає високу цінність особистого спілкування</li> <li>- організації, де гнучка методологія є універсально прийнятою</li> <li>- колокація всього колективу</li> <li>- помешкання з належним робочим середовищем у гнучкому стилі</li> <li>- система винагород, що підходить для Agile</li> <li>- члени команди з високою компетентністю та досвідом</li> <li>- члени команди з великою мотивацією</li> <li>- менеджери, обізнані у Agile процесі</li> <li>- менеджери, які мають легкий або адаптивний стиль управління</li> <li>- злагоджена, самоорганізована робота в команді</li> <li>- хороші стосунки з клієнтами</li> <li>- слідування Agile орієнтованому управлінню вимогами</li> </ul>

Автори	Критерії
	<ul style="list-style-type: none"> <li>- слідування за Agile процесу управління проектами</li> <li>- слідування гнучкому процесу керування конфігурацією</li> <li>- сильний фокус на спілкування з щоденним очним спілкуванням</li> <li>- дотримання регулярного робочого графіка - відсутність понаднормових робіт</li> <li>- сильна прихильність та присутність клієнтів</li> <li>- клієнт, що має повноваження</li> <li>- наперед чітко визначені стандарти кодування</li> <li>- простий дизайн</li> <li>- сувора діяльність з рефакторингу</li> <li>- правильна кількість документації</li> <li>- регулярна поставка програмного забезпечення</li> <li>- перш за все надаємо найважливіші функції</li> <li>- правильне тестування інтеграції</li> <li>- відповідна технічна підготовка команди</li> <li>- природа проекту не є критичною для життя</li> <li>- тип проекту, що має змінний обсяг з появою термінових нових вимог</li> <li>- проекти з динамічним, прискореним графіком</li> <li>- проекти з невеликою командою</li> <li>- проекти без кількох незалежних команд</li> <li>- проекти з попередньою оцінкою витрат</li> <li>- проекти з попереднім аналізом ризиків</li> </ul>
Шефід, Лемат	<ul style="list-style-type: none"> <li>- організаційна культура (консервативна або підприємницька)</li> <li>- стабільність організації</li> </ul>

Автори	Критерії
	<ul style="list-style-type: none"> <li>- індекс дистанції влади</li> <li>- гнучкість, підтримувана топ-менеджментом компанії</li> <li>- рівень підприємництва</li> <li>- схильність до ризику</li> <li>- рівень гнучкості замовника</li> <li>- близьке співробітництво з замовником</li> <li>- повноваження замовника, передбачувані процедурою</li> </ul>
Станкович	<ul style="list-style-type: none"> <li>- визначення процесу в проекті</li> <li>- природа проекту</li> <li>- тип проекту</li> </ul>

### Висновки до Розділу 2

У цьому розділі було описано відмінні особливості найбільш популярних гнучких методологій, таких як Scrum, Kanban, XP та гібридні методології, їхні переваги та недоліки. Також було визначено основні методики даних методологій, ефективність яких далі буде піддаватися експертному оцінюванню.

Крім цього розділ був присвячений вивченню питання про те, що входить в поняття проектного успіху та що впливає на нього. Були виведені основні критерії, що впливають на успіх проекту.

На закінчення ми маємо, що для створення робочої системи для оцінки успіху проектів великих даних та перевірки обґрунтованості емпіричних даних необхідно враховувати індивідуальні особливості цього типу проекту та вказувати критерії успіх проекту. Існує два напрямки розвитку досліджень у галузі впливу гнучких методологій на успіх проекту: з одного боку, це поглиблення в практичну область та конкретизація теорії щодо певного типу проекту, з іншого боку, створення універсальної теорії, яка визначає критерії успіху та ідентифікацію методик гнучких методологій, гарантує позитивний вплив на успіх проекту.

## РОЗДІЛ 3

### ОЦІНЮВАННЯ ВИКОРИСТАННЯ МЕТОДИК ГНУЧКИХ МЕТОДОЛОГІЙ У ПРОЕКТАХ BIG DATA

#### 3.1 Механізм оцінювання методик гнучких методологій для проектів Big Data

У якості методів дослідження було обрано метод експертних оцінок і побудову регресійної моделі.

Для збору даних використано індивідуальне експертне опитування у вигляді аналізу експертних оцінок [70], що передбачає індивідуальне заповнення експертом розробленого формуляра, за результатами якого проводиться всебічний аналіз проблемної ситуації і виявляються можливі шляхи її вирішення. Свої міркування експерт виносить у вигляді окремого документа.

Експертне оцінювання вирішує наступні завдання:

- 1) збір даних по експертам з різних індустрій і різними проектними ролями;
- 2) економія часу для експерта з допомогою використання форми для заповнення оцінки призводить до зростання кількості в отриманих відповідях;
- 3) формалізованість результатів;
- 4) мінімізація впливу середовища на відповіді експерта.

З головних недоліків експертного оцінювання можна виділити суб'єктивізацію відповідей і несамотійне їх надання. Щоб нівелювати дані мінуси, були зроблені наступні дії:

- 1) форма для заповнення експертної оцінки попередньо була випробувана як на представниках академічного середовища, так і в практиках, формулювання були уточнені;
- 2) через вузьку специфіку опитування експерти були обрані особисто, обов'язковою умовою була наявність практики в області;

3) експерт міг в будь-який момент отримати роз'яснення по кожному з пунктів форми.

Побудова регресійної моделі дозволяє обчислити статистичну залежність між досліджуваними параметрами, а саме успіхом проекту і використанням різних практик гнучких методологій.

З головних обмежень моделі можна виділити те, що модель не пояснює наявність причинно-наслідкового зв'язку, а тільки показує наявність статистичної залежності, якщо така є. Для встановлення причини і слідства, результати дослідження інтерпретуються на основі наявної теоретичної бази по темі, в якій в тому числі велику увагу приділено тому, як та чи інша методика може вплинути на успіх проекту.

У період березень - квітень 2021 року був проведено експертне оцінювання на тему «Використання гнучких методик в управлінні проектами Big Data». Оцінювання було присвячено пошуку впливу окремих методик гнучких методологій на успіх проекту на практичних прикладах.

Одна форма для заповнення експертної оцінки покривала один проект і складалася з розділів:

- відомості про експерта;
- відомості про компанію, для якої здійснювався проект;
- відомості про проект;
- використання гнучких методологій в управлінні проектами;
- успіх проекту.

Докладний зміст форми можна побачити в Додатку А.

Для участі в оцінюванні залучалися експерти з різних компаній різних галузей, але обов'язково з практикою в проектах Big Data і гнучких методологіях. Внаслідок вузької тематики дослідження мета зі збору форм становила 25-30 штук. Консолідована статистика з відповідей представлена далі.

Загалом було залучено 30 експертів. Збір інформації проходив шляхом заповнення друкованої або електронної форми. Пошук експертів відбувався шляхом залучення до опитування колег із бази практики, місця роботи, а також

з допомогою ресурсу LinkedIn. У цьому ресурсі є можливість спілкуватися з експертами у різних професійних галузях з усього світу, тому було створено звернення до деяких колег, які взяли участь у оцінюванні.

Усі експерти мають вищу профільну освіту, більшість з них також має досвід у роботі більше двох років та сертифікати з додаткових курсів та спеціалізацій. Більш детальна картина середньостатистичного експерта наведена нижче.

### **Кількісні результати експертного оцінювання.**

За статистикою найчастіша відповідь серед експертів на питання про кількість років досвіду роботи з проектами Big Data склала 2-5 років (37% відповідей), 5-7 років (27%) і менше 2 років (23%) (див. рис. 3.1).

Найбільш часта роль серед експертів - менеджер проекту (50% відповідей) (див. рис. 3.2).

Так, в середньому відповіді будуть представляти думку керівництва проекту з середнім досвідом роботи на 2-4 проектах.

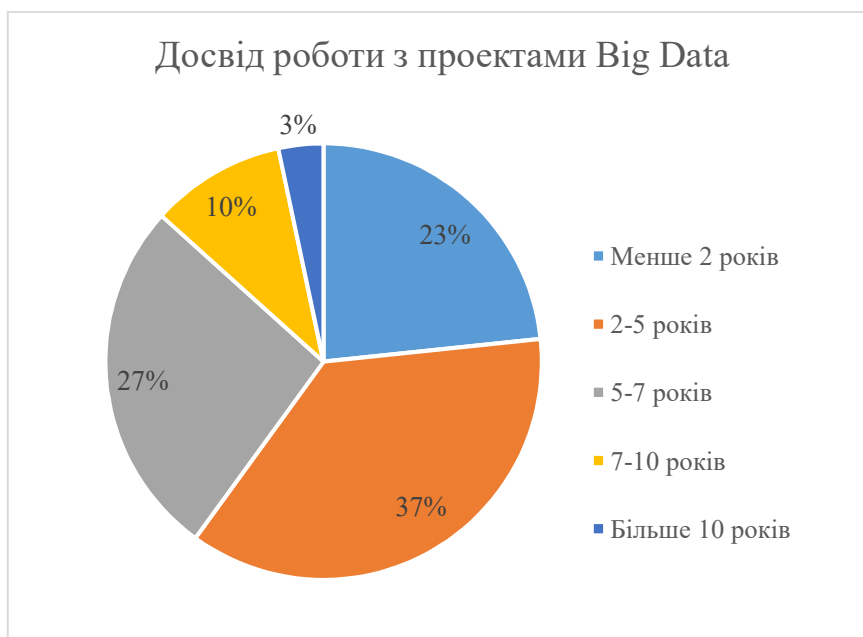


Рисунок 3.1 - Досвід роботи з проектами Big Data серед експертів, %



Рисунок 3.2 - Найчастіша роль на проєкті серед експертів, %

За статистикою, яка описує компанію, для якої/в якій здійснювався Big Data проєкт, можна побачити, що відповіді в середньому будуть характеризувати компанію

- 1) українського походження (див. рис. 3.3);
- 2) є великим підприємством (див. рис. 3.4);
- 3) віком понад 10 років (див. рис. 3.5).



Рисунок 3.3 - Походження компанії проєкту експерта, %

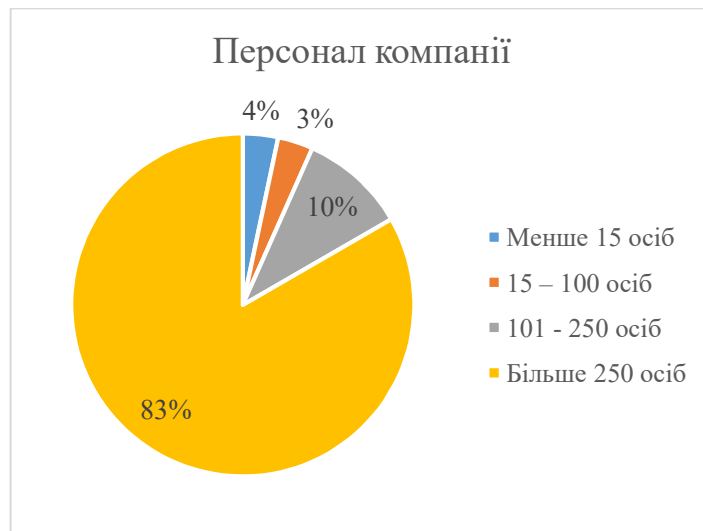


Рисунок 3.4 – Кількість співробітників компанії проекту експерта, %

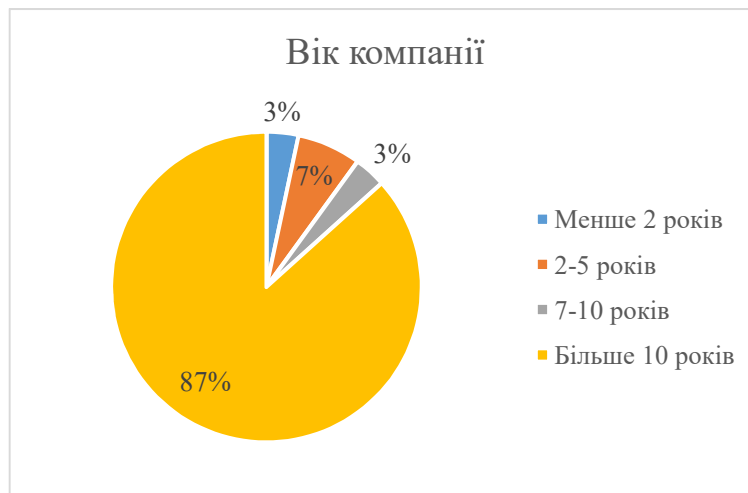


Рисунок 3.5 - Вік компанії, для якої / в якій здійснювався проект експерта, %

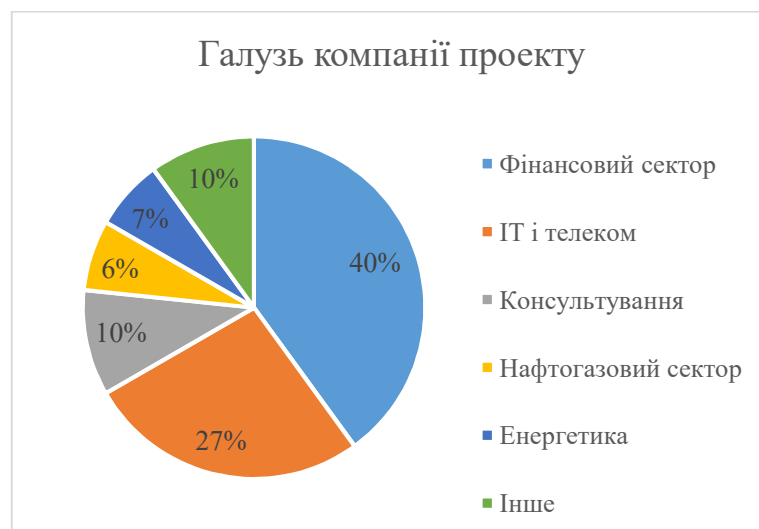


Рисунок 3.6 - Галузь компанії, для якої/в якій здійснювався проект експерта, %

Галузь компанії в більшості випадків представлена фінансовим сектором (40% відповідей) та ІТ-та телеком компаніями (27%), що очікувано, тому що ІТ-проекти з використанням гнучких методологій почалися саме в таких компаніях (див. рис. 3.6).

Щоб оцінити рівень відкритості до використання гнучких методологій і наявність експертів в області, що є одним з головних чинників успішного використання гнучких методів і методик, учасникам дослідження пропонувалося вибрати оцінку, що більшою мірою виражає згоду з наведеними твердженнями (1 - абсолютно не згоден, 2 - не згоден, 3 - важко відповісти, 4 - згоден, 5 - абсолютно згоден):

1) компанія (в тому числі керівництво) відкрито до використання гнучких методологій з управління проектами;

2) компанія має досвід у використанні гнучких методологій, в штаті є експерти в області гнучких методологій.

Середня відповідь склала 3,67 і 2,9 відповідно, що говорить про те, що в цілому компанії у вигляді керівництва починають усвідомлювати і приймати ефективність нових методологій, в тому числі гнучких, але в той же час в компанії не вистачає експертних знань для правильного використання гнучких методик. Дані чинники є важливими факторами впливу на вдале застосування методик гнучких методологій управління у проектах Big Data, тому їх необхідно також включити до регресійної моделі.

Далі представлена статистика по опису Big Data проекту, в якому використовувалися гнучкі методології.

На наступній діаграмі на рисунку 3.7 показані відповіді експертів про бюджет проекту Big Data, над яким вони працюють:

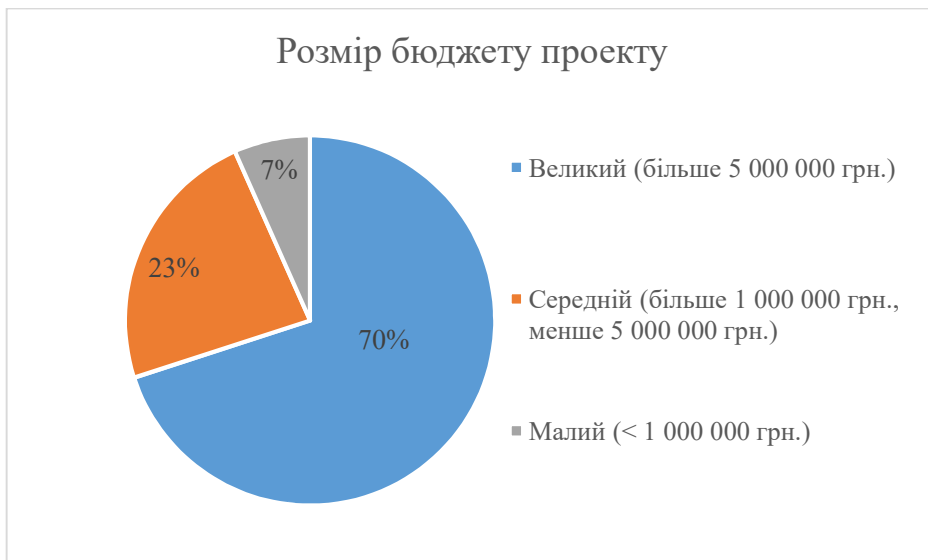


Рисунок 3.7 - Розмір бюджету проекту експерта, %

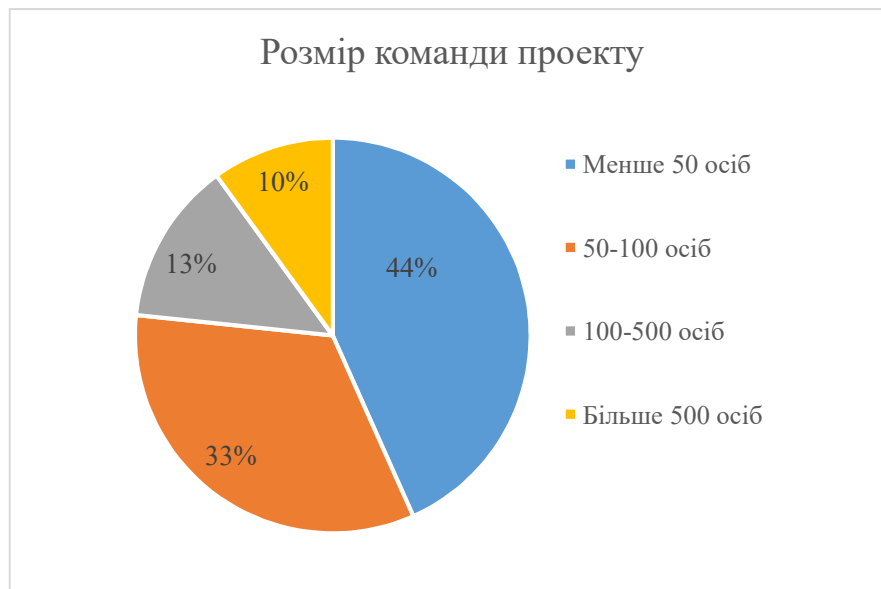


Рисунок 3.8 - Розмір команди проекту експерта, %

Проекти вибірки більшою мірою характеризуються великим бюджетом, більше 5 млн. грн. (див. рис. 3.7), невеликою командою в кількості менше 50 осіб (44% відповідей) (див. рис. 3.8), що пояснюється високою кваліфікацією і затребуваністю фахівців з одного боку і специфікою використання гнучких методологій, що складається в організації невеликих кросфункціональних команд, - з іншого боку.



Рисунок 3.9 - Місцезнаходження команди проекту експерта, %

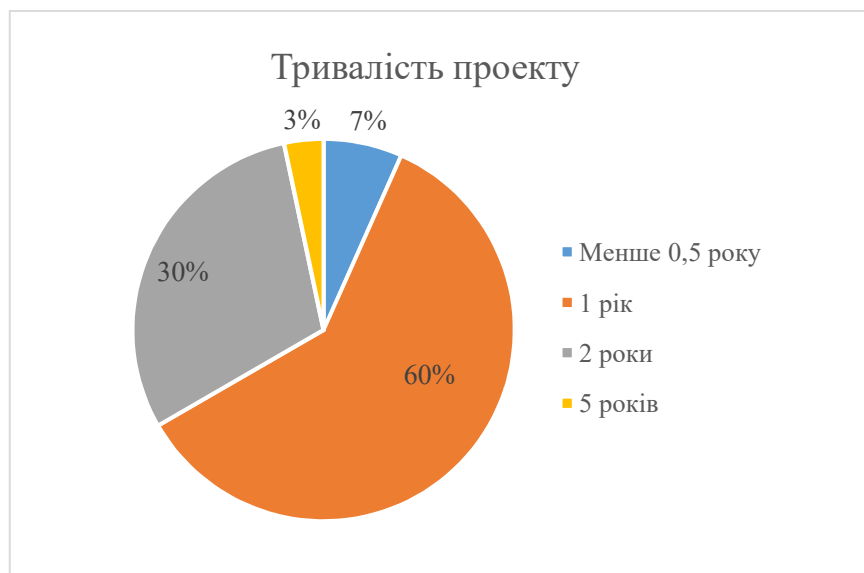


Рисунок 3.10 - Тривалість проекту експерта, %

За статистикою в 60% випадків команда проекту була розподіленою. Найчастіше віддалено працюють розробники і технічні фахівці (див. рис. 3.9). Сучасні способи зв'язку дозволяють підтримувати ефективну і постійну комунікацію і залучати експертів з різних регіонів. Тривалість проекту в 60% випадків склала 1 рік, що характерно для Big Data проекту середньої складності (див. рис.3.10) [32].

## Використання методик гнучких методологій.

Далі мова піде про використання окремих гнучких методик для управління проектами. Експертам пропонувалося вибрати гнучкі методики, які були використані на проекті (рис. 3.11).

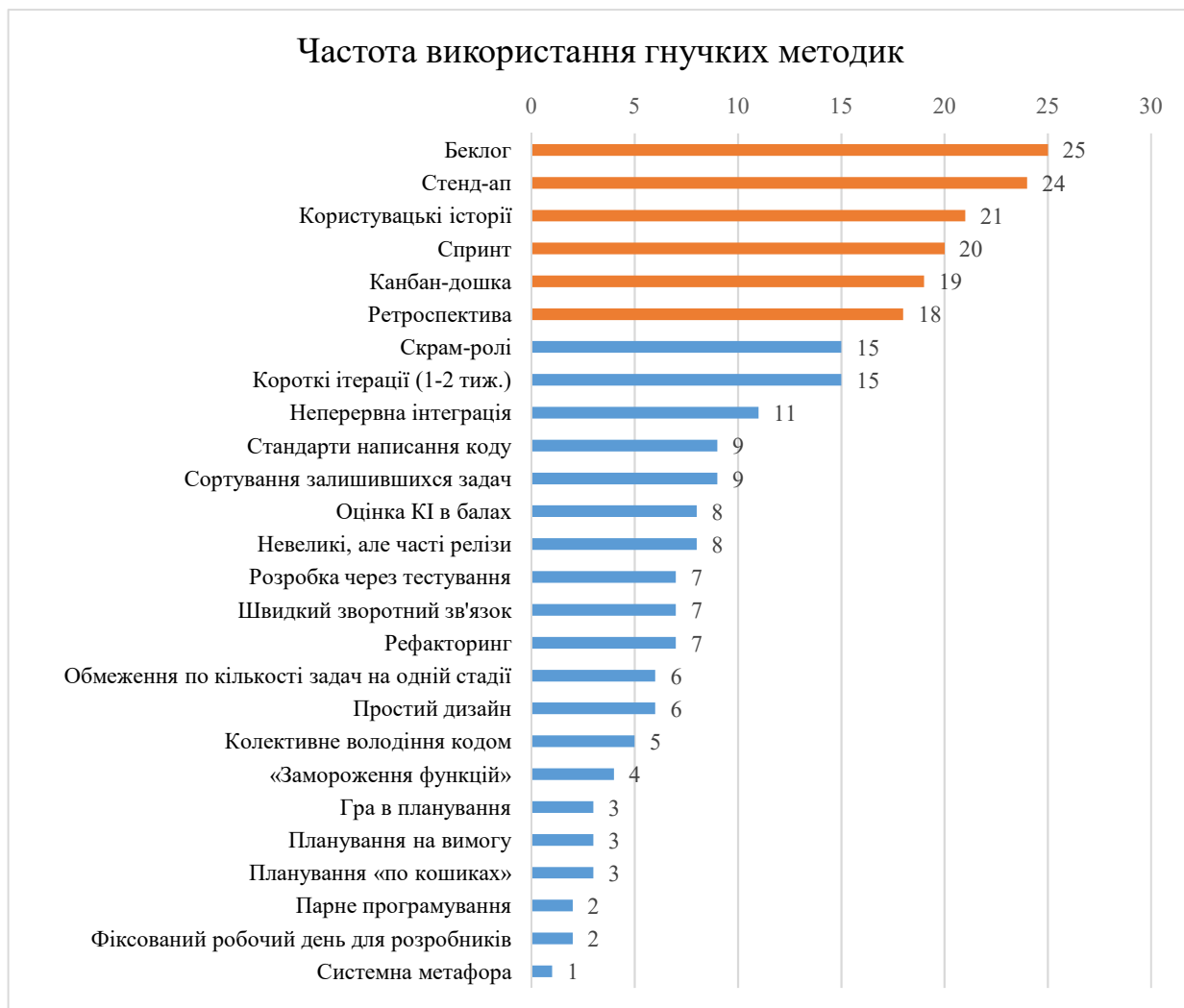


Рисунок 3.11 - Частота використання гнучких методик у проектах Big Data

Найбільш часто використовувані методики, виділені помаранчевим кольором на рисунку 3.11, а саме: Беклог, Стенд-ап, Користувацькі історії, Спринт, Канбан-дошка та Ретроспектива - будуть використані в подальшому для побудови регресійної моделі.

Фокус-групі, що складається з 7 експертів в області управління проектами, пропонувалося визначити пари гнучких методик, найбільш часто використовуваних разом в управління проектами Big Data. Далі загальній групі

експертів пропонувалося вибрати пари, які були використані на описуваному проекті. Була отримана наступна статистика (рис. 3.12).

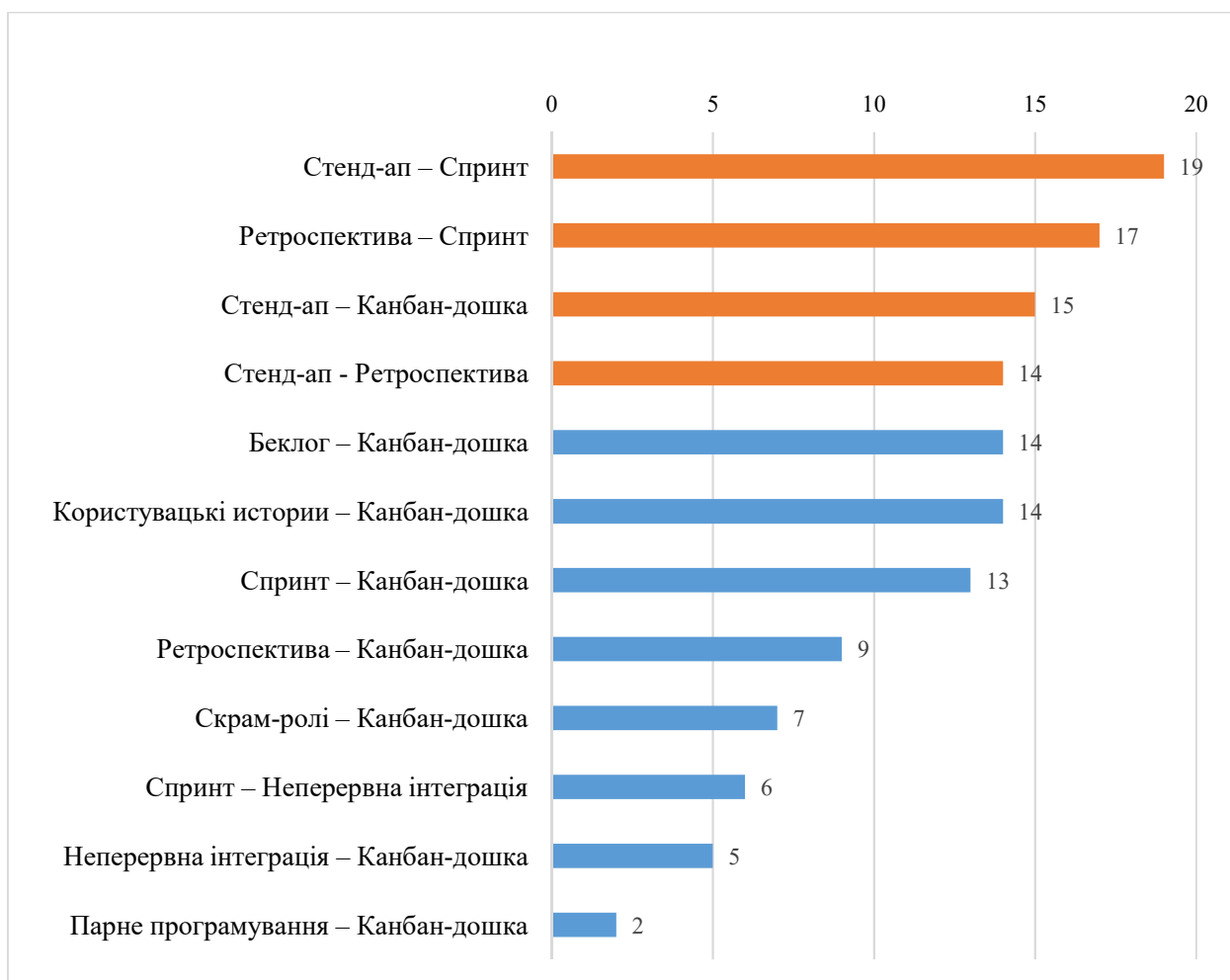


Рисунок 3.12 - Частота використання пар гнучких методик у проектах Big Data

Як видно на рисунку 3.12, найбільш часто використовувані пари, а саме ТОП-4 з 12, перетинаються з найбільш популярними окремо використовуваними гнучкими методиками. Можна очікувати, що при регресійному аналізі будуть виявлені пари і, можливо, трійки найбільш часто використовуваних гнучких методик.

Один з розділів опитування полягав в оцінці успіху проекту Big Data, в управлінні яких використовувалися гнучкі методики. Експертам пропонувалося спочатку вибрати характеристики успіху проекту, що входять в офіційні критерії успіху, а потім оцінити всі за 5-ти бальною шкалою Лайкерта з позначеннями: 1

- абсолютно не згоден, 2 - не згоден, 3 - важко відповісти, 4 - згоден, 5 - повністю згодний.

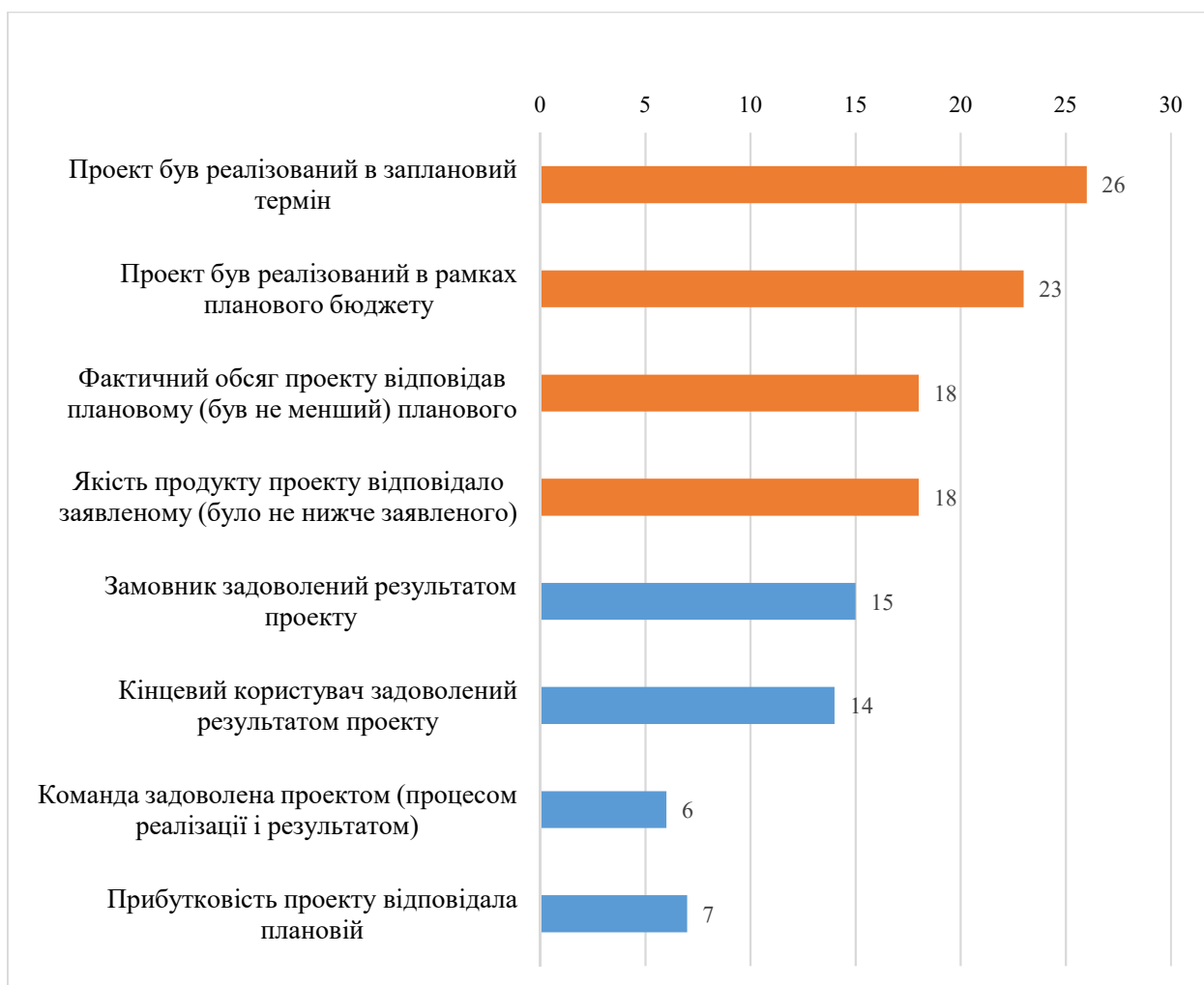


Рисунок 3.13 - Частота використання критерію успіху проекту Big Data

Як видно на рисунку 3.13, найбільш популярними критеріями виявилися параметри «проектного трикутника» - термін, бюджет і обсяг/якість. При побудові регресійної моделі кожному з критеріїв будуть присвоєні ваги відповідно до частоти їх використання при офіційній оцінці успіху проекту.

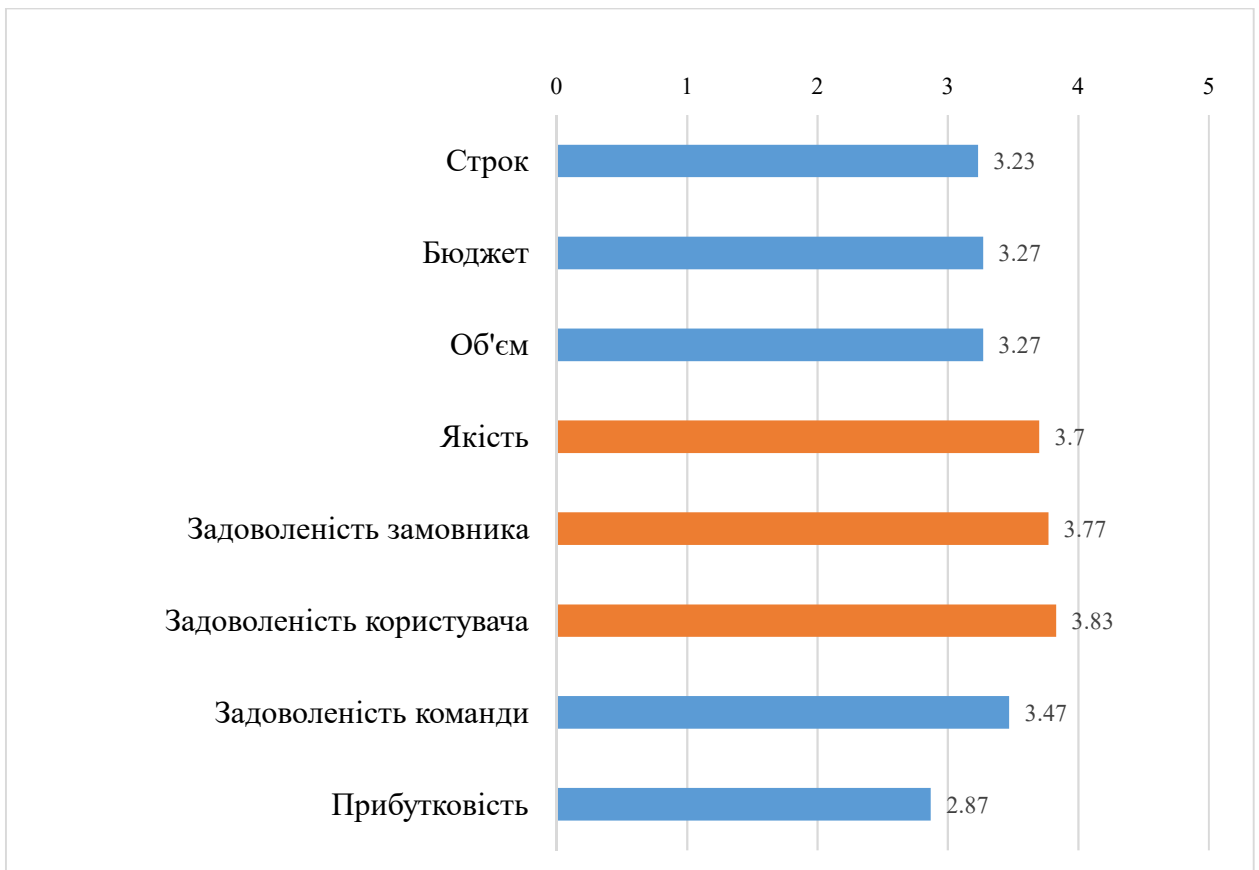


Рисунок 3.14 - Оцінка успіху проекту Big Data від експертів

Усереднена відповідь експерта представлена на рисунку 3.14. Однак, потрібно розуміти, що на відносно низьких оцінках (в середньому три бали) можна зробити висновок про неефективність використання гнучких методологій - причина може полягати в невмінні їх застосування або низькою часткою успішних проектів Big Data з конкретною специфікою. Важливо відзначити, що найбільший бал отримали критерії задоволеність замовника, користувача і якість - ключові «фокуси» гнучких методологій.

Далі експертам пропонувалося оцінити, на який з критеріїв успіху проекту Big Data впливає та чи інша гнучка техніка. Результати продемонстровані на рисунку 3.15.

Техніка	Термін	Бюджет	Якість	Об`єм	Замовник	Користувач	Команда	Прибуток	Не знаю / не використовую	Сума
Беклог	15	8	12	19	10	7	6	0	6	77
Користувацькі історії	7	6	14	13	9	13	9	4	6	75
Спринт	21	6	10	10	9	8	12	2	7	78
Стенд-ап	18	0	13	9	6	2	13	0	4	61
Оцінка КІ в балах	7	4	7	11	2	1	6	0	14	38
Scrum-ролі	3	2	10	6	4	3	11	0	12	39
Ретроспектива	4	0	14	4	9	5	17	1	5	54
Канбан-дошка	19	2	10	16	5	3	11	0	7	66
Обмеження за кількістю завдань на одній стадії	9	5	15	7	3	2	10	0	7	51
Гра в планування	6	3	3	3	7	6	9	1	18	38
Парне програмування	2	0	9	3	2	1	6	0	18	23
Розробка через тестування	4	1	10	5	1	2	6	0	17	29
Швидкий зворотний зв'язок	8	3	10	4	11	11	12	3	8	62
Неперервна інтеграція	10	1	7	5	7	5	5	2	15	42
Рефакторинг	2	3	10	2	1	0	2	1	20	21
Невеликі, але часті релізи	9	3	10	6	8	6	5	2	12	49
Стандарти написання коду	6	2	13	5	0	1	6	0	13	33
Колективне володіння кодом	5	1	12	5	0	0	6	0	16	29
Простий дизайн	4	0	8	0	4	5	3	0	16	24
Системна метафора	1	0	4	1	0	1	5	0	21	12
Фіксований робочий день для розробників	2	3	1	1	1	0	10	1	15	19
Короткі ітерації (1-2 тиж.)	11	3	6	6	4	6	6	1	11	43
Планування на вимогу	3	1	4	2	4	3	3	0	19	20
Планування «по кошиках»	5	2	3	2	4	1	3	0	21	20
«Замороження функцій»	4	3	2	5	2	2	3	1	22	22
Сортування залишившихся задач	7	2	6	11	3	1	5	0	16	35

Рисунок 3.15 - Вплив використання гнучких методик на критерії успіху проекту Big Data

З даної статистики можна зробити кілька важливих висновків:

1. Експерти, що здебільшого є менеджерами проекту, не інформовані про вплив тієї чи іншої методики на успіх проекту Big Data (див. стовпець «Не знаю / не використовую» на рисунку 3.15).

2. На думку експертів найбільший вплив надають методики методології Scrum - Беклог, Користувацькі історії, Спринт, Стенд-ап, Ретроспектива, методології канбан - Канбан-дошка та Обмеження за кількістю завдань на одній стадії, - і Швидкий зворотний зв'язок, що є елементом будь-якої гнучкої методології.

3. На думку експертів гнучкі методики мають найбільший вплив на такі показники успіху проекту, як якість, термін і задоволеність команди.

4. У передостанньому розділі форми експерти повинні були оцінити за шкалою Лайкерта, як той чи інший фактор може перешкодити ефективному використанню гнучких методологій. «Стоп-фактори», що були вказані у формі, були виявлені при аналізі літератури вище. Результати представлені на рисунку 3.16.

Найбільш важливими «стоп-факторами» були виділені:

- неготовність керівництва;
- консервативна культура компанії;
- відсутність необхідних експертів.

Дані результати підтверджують правильність включення в регресію факторів, що показують відкритість керівництва і наявність експертизи по тематиці гнучких методологій в компанії.

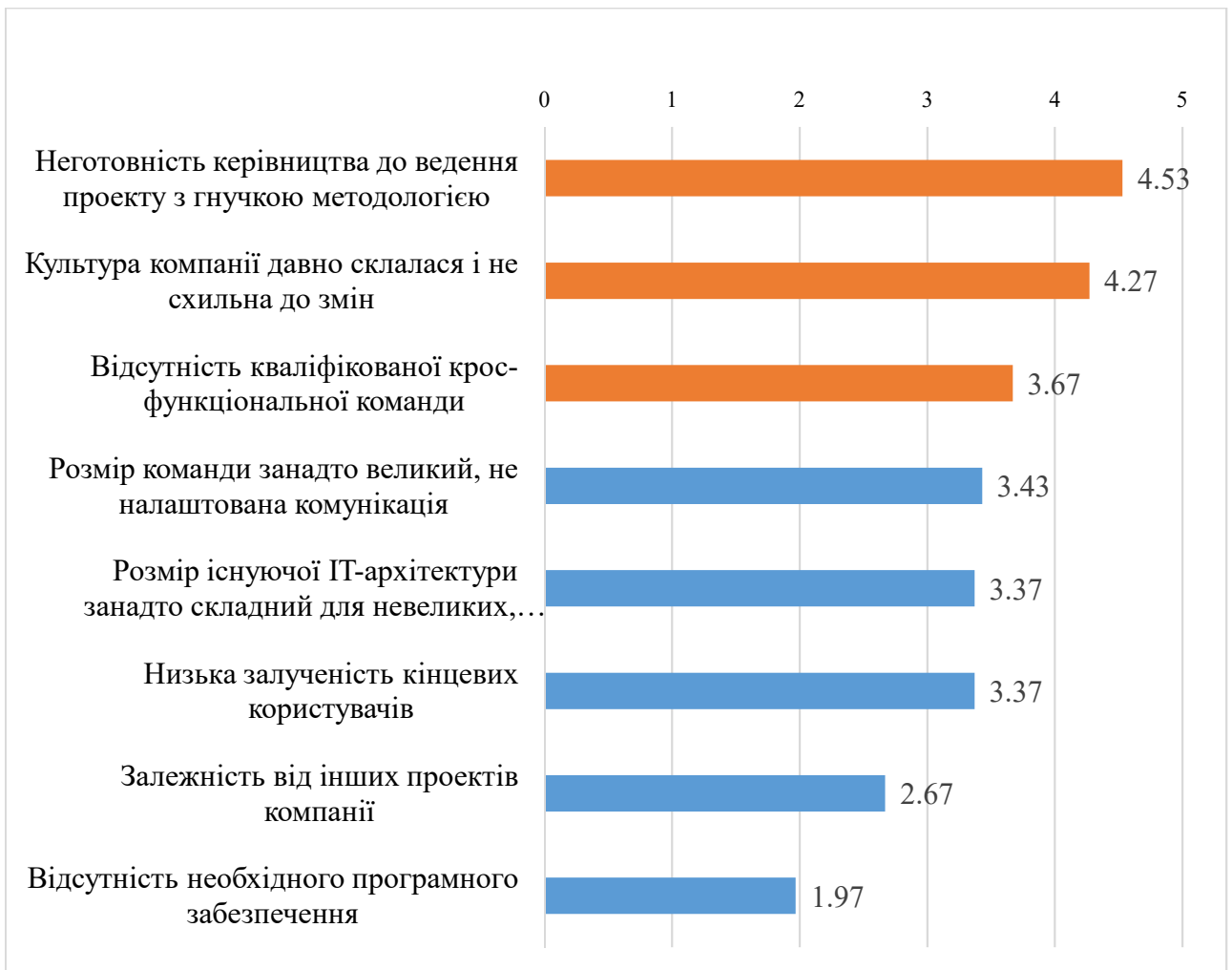


Рисунок 3.16 - Фактори, що перешкоджають ефективному використанню гнучких методологій у проектах Big Data

### **Якісні результати експертного оцінювання.**

У ході збору даних експертного оцінювання експертам пропонувалося залишити коментарі у відкритій формі на тему управління проектами Big Data за допомогою гнучких методик. Деякі думки щодо застосування методик гнучких методологій повторювалися в декількох відповідях, а саме:

1. У проектах Big Data важко точно спланувати, яку роботу потрібно буде закінчити. Потрібно постійно проводити експерименти та дослідження. Це робить роботу більш ітеративною. Ітеративний підхід також дозволяє зменшити витрати по внесенню змін до вимог. Ітеративність гнучких методологій ідеально підходить для проектів Big Data.

2. Гнучкі методології дозволяють швидко реагувати на зовнішні зміни ринку та ефективно і швидко змінювати бізнес-гіпотезу.

3. Гнучка методологія надає аналітикам даних можливість визначати пріоритети моделей та даних відповідно до цілей та вимог проекту. Це також допомагає аналітикам даних дати короткий огляд кожної мети нетехнічним зацікавленим сторонам.

4. Процеси проектів Big Data зазвичай включають високий ступінь невизначеності. Тут гнучкі методології узгоджуються з наукою про дані.

5. Гнучкість в плануванні дозволяє вчасно і повно враховувати зворотний зв'язок замовника і кінцевого користувача.

6. Для кожного проекту Big Data необхідно підбирати окремі методики різних методологій, а гнучку методологію легко адаптувати.

7. Гнучкі методології ефективні, якщо є орієнтація на замовника і користувача, а не на бюджет і термін.

Big Data забезпечує цінні результати моделей даних, які відповідають на найважливіші бізнес-питання. Використання гнучких методів може бути корисним під час управління робочими процесами Big Data. Такі методи, як Scrum, Kanban та XP, можуть допомогти керувати різними проектами з обробки великих даних. Ключовим є використання тих аспектів методологій, де обидві дисципліни перетинаються.

### **3.2 Вибір ефективних методик гнучких методологій для Big Data проекту**

Задача, що розглядається, полягає в наступному - потрібно вибрати параметри  $\theta$  \* моделі, які пояснюють вихідні дані якнайкраще.

Щоб формалізувати це, потрібно подумати про наступні проблеми:

- (Модель): Що таке хороший набір функцій  $\{f\}$ , який дає правдоподібні моделі?

- (Мета): У якому сенсі потрібно сформулювати оптимальність оцінки?

- (Оптимізація): Як вирішити проблему "схожості" (зокрема, це може бути мінімізація помилки)?

Дана постановка дозволяє реалізувати перехід від спостережуваної поведінки даних до дослідження параметрів, які безпосередньо не спостерігаються, але відображають суть процесу. І зробити це можна зробити по-різному. Ось деякі варіанти:

- LS: Традиційна задача оцінки найменших квадратів (Least Squares):

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N (f(t_i, \theta) - x_i)^2, \quad (3.1)$$

- WLS: оцінка зважених найменших квадратів (Weighted Least Squares):

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N w_i^2 (f(t_i, \theta) - x_i)^2, \quad (3.2)$$

- TA: Апроксимація Чебишева (Tchebychev Approximation):

$$\theta^* = \arg \min_{\theta} \max_{i=1, \dots, N} |f(t_i, \theta) - x_i|, \quad (3.3)$$

- L1-апроксимація:

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N |f(t_i, \theta) - x_i|, \quad (3.4)$$

- L<sub>p</sub> (p ∈ [0, ∞)): Загальний випадок, що включає в себе всі розглянуте вище (чебишовських випадок виходить при p → ∞)

$$\theta^* = \arg \min_{\theta} \sqrt[p]{\sum_{i=1}^N |f(t_i, \theta) - x_i|^p}, \quad (3.5)$$

- L0: Робастая апроксимація

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N |f(t_i, \theta) - x_i|_0, \quad (3.6)$$

де  $|z|_0$  дорівнює одиниці, якщо  $z \neq 0$  і  $|z|_0 = 0$  тоді і тільки тоді, коли  $z = 0$ .

Вибір інструмента рішення кожен раз залежить від розв'язуваної задачі. Найбільш підходящий метод – метод найменших квадратів. Це обумовлено тим, що ключовим моментом всіх перерахованих підходів є пошук мінімального значення деяких параметрів. Як пам'ятаємо з курсу математичного аналізу, ця задача розбивається на дві - знаходження необхідного і достатнього умов. Якщо необхідна умова знаходиться досить просто - шукаємо похідні функції мети за всіма параметрами, прирівнюємо їх нулю і отримуємо систему рівнянь. Значення параметрів, які реалізують нашу задачу знайдені. То з достатніми умовами все набагато складніше. У багатьох випадках знаходження достатніх умов екстремуму досить проблематичні. З іншого боку, якщо функція мети опукла, то необхідні умови екстремуму збігаються з достатніми. Прикладом опуклою функції мети є параболоїд, тобто, лінійна функція, зведена в квадрат.

Таким чином, якщо ми нічого не знаємо про суть завдання - беремо помилку, зводимо в квадрат і отримуємо опуклу функцію мети. Це дозволяє отримати наближення шуканого рішення.

Перейдемо до викладу методу найменших квадратів [69].

Поставимо у відповідність з вихідними даними, заданими у вигляді таблиці, функцію виду:

$$f(\{a_i\}_{i=0}^n, t) = \sum_{i=0}^n a_i \varphi_i(t) = a_0 \varphi_0(t) + a_1 \varphi_1(t) + \dots + a_n \varphi_n(t), \quad (3.7)$$

де  $\varphi_i(t), i=0, \dots, n$ -базисні функції, а  $a_i$  – невідомі коефіцієнти, що підлягають визначенню.

Зокрема, якщо в якості базисних функцій використовувати степеневі маномі  $\varphi_i(t)=t^i$ , задача зводиться до пошуку полінома

$$f(\{a_i\}_{i=0}^n, t) = \sum_{i=0}^n a_i t^i = a_0 + a_1 t + \dots + a_n t^n \quad (3.8)$$

ступеня  $n$ , що наближає вихідну таблицю.

Для визначення коефіцієнтів  $a_i$  будемо шукати функцію  $f(\{a_i\}_{i=0}^n, t)$ , відхилення значень якої від заданих таблицею значень  $x_i$  мінімально в деякому середньо інтегральному сенсі.

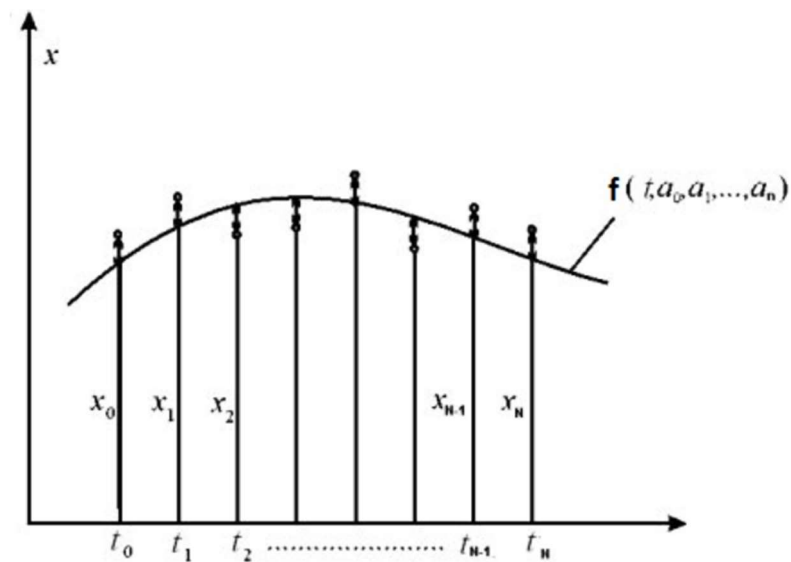


Рисунок 3.17 – Ілюстрація алгоритму методу найменших квадратів

Досить часто природною умовою є вимога опису вхідних даних прямою або параболою. У цьому випадку говорять, що використовується лінійна або квадратична регресія.

Розглянемо випадок опису апріорних даних прямою, тобто використаємо метод лінійної регресії. Наведемо вихідні дані  $(t_i, x_i)$ ,  $i = 0, 1, \dots, N$  прямий  $x = at + b$ . У цьому випадку функція мети набуде вигляду:

$$S(a, b) = \sum_{i=0}^N (at_i + b - x_i)^2 \rightarrow \min_{a, b} \quad (3.9)$$

Необхідна (і, в даному випадку, достатня) умова екстремуму має вигляд

$$\begin{cases} \frac{\partial}{\partial a} S(a, b) = 2 \sum_{i=0}^N t_i (at_i + b - x_i) = 0, \\ \frac{\partial}{\partial b} S(a, b) = 2 \sum_{i=0}^N (at_i + b - x_i) = 0, \end{cases} \quad (3.10)$$

або, що те ж саме

$$\begin{cases} a \sum_{i=0}^N t_i^2 + b \sum_{i=0}^N t_i = \sum_{i=0}^N x_i t_i, \\ a \sum_{i=0}^N t_i + b(N+1) = \sum_{i=0}^N x_i. \end{cases} \quad (3.11)$$

Звідси, застосовуючи метод Крамера рішення систем лінійних рівнянь, відразу отримуємо коефіцієнти прямої (лінійної регресії):

$$a = \frac{(N+1) \sum_{i=0}^N x_i t_i - \sum_{i=0}^N t_i \sum_{i=0}^N x_i}{(N+1) \sum_{i=0}^N t_i^2 - \left( \sum_{i=0}^N t_i \right)^2}, \quad b = \frac{\sum_{i=0}^N t_i^2 \sum_{i=0}^N x_i - \sum_{i=0}^N t_i \sum_{i=0}^N x_i t_i}{(N+1) \sum_{i=0}^N t_i^2 - \left( \sum_{i=0}^N t_i \right)^2}. \quad (3.12)$$

Таким чином, як регресійна модель буде використана лінійна МНК-модель, побудована за допомогою мови програмування Python (код вказаний в Додатку 2). Буде побудовано дві моделі - з присвоєнням різних ваг кожному показнику успіху проекту в залежності від частоти його згадки і з присвоєнням однакової ваги - 1.

Як регресант буде використаний консолідований показник успіху проекту:

- $Success1 = \sum c_i$ , де  $c$  - оцінка критерію від 1 до 5;
- $Success2 = \sum w_i c_i$ , де  $w$  - вага критерію,  $c$  - оцінка критерію від 1 до 5.

Як регресори будуть використані:

- даммі-змінні, що характеризують використання кожної з шести методик гнучких методологій, описаних вище: backlog, us, sprint, standup, retrosp, kanban;
- оцінка відкритості керівництва компанії до використання гнучких методологій: openness;
- оцінка наявності експертів в штаті в області управління проектами по гнучкими методологіями: expertise.

### Результати регресійного аналізу.

Тут і далі наведено результати моделі два (з регресанта success2). Результати консистентні з результатами моделі один, наведеної в Додатку 3. На першій ітерації регресійного аналізу були отримані наступні результати:

OLS Regression Results						
=====						
Dep. Variable:	success2	R-squared:	0.676			
Model:	OLS	Adj. R-squared:	0.553			
Method:	Least Squares	F-statistic:	5.478			
Date:	Sun, 22	Prob (F-statistic):	0.000821			
Time:	18:02:38	Log-Likelihood:	-70.525			
No. Observations:	30	AIC:	159.0			
Df Residuals:	21	BIC:	171.7			
Df Model:	8					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]
-----						
const	1.2353	2.314	0.534	0.599	-3.577	6.048
backlog	1.3899	2.312	0.601	0.554	-3.418	6.198
us	-2.8106	2.044	-1.375	0.184	-7.061	1.439
sprint	1.8974	1.930	0.983	0.337	-2.117	5.912
standup	3.6865	1.692	2.178	0.041	0.167	7.206
retrosp	-2.5564	1.522	-1.679	0.108	-5.722	0.609
kanban	3.4790	1.303	2.670	0.014	0.769	6.189
openness	2.0075	0.652	3.081	0.006	0.652	3.363
expertise	0.6382	0.509	1.254	0.224	-0.420	1.697
=====						

Рисунок 3.18 - Результати першої ітерації регресійного аналізу.

Гіпотеза про незначущості регресії відкидається на будь-якому адекватному рівні значущості. Іншим параметром для перевірки ефективності моделі є значення R-квадрата, яке представляє відсоток змін у залежній змінній

(success2), що пояснюється незалежними змінними. Чим вище значення, тим краще пояснювальність моделі, причому найвище значення - одне. У нашому випадку значення R-квадрат становить 68%, що говорить про нормальну пояснювальну силу регресії. Кількість спостережень - 30, що дозволяє зробити припущення про нормальний розподіл спостережень.

При лінійній регресії розмір коефіцієнта (coef) для кожної незалежної змінної дає розмір ефекту, який ця змінна надає на залежну змінну, а знак на коефіцієнті (позитивний чи негативний) - напрямок ефекту. У даному випадку з кількома незалежними змінними коефіцієнт повідомляє, наскільки очікувано збільшення залежної змінної (якщо коефіцієнт позитивний) або зменшення (якщо коефіцієнт від'ємний), коли ця незалежна змінна збільшиться на одиницю, утримуючи всі інші незалежні змінні постійними.

Проте як видно з рисунку 3.18, на 5%-му рівні значущості виявилися значущими тільки даммі-змінні standup, kanban і змінна openness, тому що – змінна  $P > |t| < 0,05$ . Вони будуть використані на наступній ітерації.

Результати другої ітерації представлені на рисунку 3.18:

OLS Regression Results						
Dep. Variable:	success2	R-squared:	0.576			
Model:	OLS	Adj. R-squared:	0.527			
Method:	Least Squares	F-statistic:	11.76			
Date:	Sun, 22	Prob (F-statistic):	4.70e-05			
Time:	18:59:48	Log-Likelihood:	-74.572			
No. Observations:	30	AIC:	157.1			
Df Residuals:	26	BIC:	162.7			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	0.9610	2.354	0.408	0.686	-3.877	5.799
standup	3.6091	1.439	2.508	0.019	0.651	6.567
kanban	3.6395	1.210	3.007	0.006	1.152	6.127
openness	2.2821	0.494	4.623	0.000	1.267	3.297

Рисунок 3.19 - Результати другої ітерації регресійного аналізу.

Регресія як і раніше залишається значущою, хоча R-квадрат трохи зменшився - до 58% (див. рис. 3.19). Всі змінні виявилися значущими. Перш ніж інтерпретувати результати, проведемо кілька статистичних тестів.

### **Перевірка на гетероскедастичність, мультиколінеарність і на ненормальність залишків.**

Перш за все був проведений тест на мультиколінеарність, що покаже, чи наявна лінійна залежність між пояснювальними змінними регресійної моделі. Така залежність може призвести до невизначеності значень параметрів чи до нестійкості їх оцінок. Нестійкість виражається в збільшенні статистичної невизначеності - дисперсії оцінок. Це означає, що конкретні результати оцінки можуть сильно відрізнятись для різних вибірок незважаючи на те, що вибірки однорідні. Як міра мультиколінеарності буде використовуватися VIF (Variance Inflation Factor, фактор інфляції дисперсії) – міра мультиколінеарності, що дозволяє оцінити збільшення дисперсії через лінійну залежність фактора  $x_i$  від інших

$$\text{Var}(\beta_j) = \frac{\sigma^2}{\sum_i (x_{ij} - \bar{x}_j)^2} VIF_j = \frac{\sigma^2}{\sum_i (x_{ij} - \bar{x}_j)^2} \frac{1}{1 - R_j^2} \quad (3.13)$$

де  $R_j^2$  – коефіцієнт детермінації лінійної залежності фактора  $x_i$  від інших факторів.

Можливі значення VIF:

- VIF > 10 – виражена мультиколінеарність;
- VIF = 5-10 – мультиколінеарність;
- VIF = 1-5 – немає мультиколінеарності.

Фактор інфляції дисперсії (VIF) для всіх змінних менше 5, що говорить про відсутність мультиколінеарності (рис. 3.20).

	VIF Factor	features
0	4.3	standup
1	2.2	kanban
2	4.5	openness

Рисунок 3.20 - Показник VIF для регресії.

Для перевірки гіпотези про те, що залишки цього ряду мають нормальний розподіл, була використана статистика Жака-Бера, яка склала 0,31. Нульова гіпотеза про нормальність розподілу залишків не відкидається на будь-якому адекватному рівні значущості ( $p\text{-value} = 0,86$ ).

За допомогою тесту Бройша-Пагана було досліджено наявність гетероскедастичності. P-value склало 0,16, тобто нульова гіпотеза про гомоскедастичність не відкидається на 15% рівні значимості.

Підсумкове рівняння регресії виглядає наступні чином:

$$Success2 = 0,961 + 3,609standup_i + 3,67kanban_i + 2,282openness_i + e_i \quad (3.14)$$

В результаті проведеного аналізу можемо проінтерпретувати результати та надати оцінку впливу методик гнучких методологій на успіх проекту Big Data:

1. Коефіцієнт в розмірі близько 3,6 при даммі-змінних свідчить про те, що при використанні Стенд-апу або Канбан-дошки консолідована змінна успіху проекту збільшується на 3,6 балів. Ці дві методики підвищують успішність проекту Big Data.

2. Коефіцієнт в розмірі 2,28 при змінній, що характеризує відкритість керівництва, говорить про те, що позиція керівників компанії позитивно впливає на успіх проекту: при збільшенні категоріальної змінної «відкритість» на 1 бал змінна «успіх» збільшується на 2,28 балів.

### **Висновки до Розділу 3**

У цьому розділі було проведено емпіричне дослідження на тему впливу методик гнучких методологій на успіх Big Data проекту. Серед найбільш популярних методик гнучких методологій найбільш ефективними виявилися Стенд-ап і Канбан-дошка. Важливо відзначити, що дані методики дійсно можливо використовувати окремо, при цьому вони не вимагають поглиблених знань в області гнучких методологій. Результатом відсутності експертизи виступає неефективність інших досліджуваних методик - Спринту, Беклога, парного програмування і інших.

Можна зробити висновок, що в Україні визнають важливість гнучких методологій, але в той же час фахівці поки не володіють достатньою кваліфікацією для їх успішного використання у Big Data проектах. У той же час адаптація методології та застосування різних методик вимагає ще більш глибоких знань за тематикою. Існуюча тенденція полягає в тому, що на проектах Big Data часто намагаються механічно ввести використання гнучких методологій без відповідної зміни в корпоративній культурі і штаті компанії. Такий метод в більшості випадків виявляється неефективним.

Інший тренд - використання класичних методів управління проектами з «гнучким» назвою, наприклад, Стенд-ап багато в чому нагадує класичні ранкові «планерки», Канбан-дошка - список поточних завдань і їх статус. При недостатньому зануренні в специфіку методики у менеджмента може виникнути відчуття, що гнучка методика вже освоєна, навіть якщо це не відповідає дійсності.

## РОЗДІЛ 4

### РЕКОМЕНДАЦІЇ І ПОДАЛЬШЕ ДОСЛІДЖЕННЯ

#### 4.1 Практичні рекомендації

У цьому розділі буде наведено практичні рекомендації за підсумками роботи і змальовано напрямок подальших досліджень по темі. Як було показано вище, застосування гнучких методологій на Big Data проекті дійсно дозволяє підвищити ймовірність успіху проекту, особливо в областях, яким раніше не приділялося належної уваги - задоволеність користувача, команди, замовника. Однак, застосування гнучких методологій вимагає кропіткої роботи як і на початку проекту, на етапі вибору методології, так і під час проекту.

Практичні рекомендації можна розділити на дві складові:

- «Що ми робимо» - які критерії успіху проекту закріплені як офіційні, яка методологія використовується, як реалізуються процеси управління проектом і виконання проекту;
- «Як ми це робимо» - рівень розвитку і прийняття гнучкої культури в компанії, правильність виконання того чи іншого процесу.

У таблиці нижче наведено, які дії виконує проектна команда при реалізації проекту і на що необхідно звернути увагу з точки зору того, як це виконується. При підготовці таблиці був використаний практичний досвід реалізації Big Data проекту за допомогою гнучких методологій (робота над Big Data проектом в ріелторській компанії, а також кейси компаній КіноПоиск, Work.ua, Sokolov, HeadHunter та ін., на основі даних компанії ScrumTrek):

Таблиця 4.1 – Дії проектної команди і специфічні характеристики їх виконання

№	Що ми робимо	Як ми це робимо
1	Вибір критеріїв успіху проекту	- залучення всіх стейкхолдерів у формування поняття «успіх даного проекту Big Data»

№	Що ми робимо	Як ми це робимо
		<ul style="list-style-type: none"> <li>- включення не тільки кількісних (як зазвичай у проектах Big Data), а й якісних показників - трансляція зворотного підходу всередині команди</li> </ul>
2	Формування методології проекту	<ul style="list-style-type: none"> <li>- навчання команди, включаючи керівництво, бізнес-користувачів</li> <li>- залучення експертів</li> <li>- використання Стенд-апу, Канбан-дошки</li> </ul>
3	Відкритість до змін	<ul style="list-style-type: none"> <li>- висновки на основі зібраних метрик і зібраного зворотного зв'язку</li> <li>- залучення керівництва / експертів для реалізації змін в процесі реалізації проекту</li> </ul>
4	Дотримання гнучкої культури	<ul style="list-style-type: none"> <li>- наявність комунікації між усіма учасниками проекту Big Data</li> <li>- постійне поліпшення процесів</li> <li>- прозорість виконання проекту</li> </ul>
5	Збір даних про виконання проекту	<ul style="list-style-type: none"> <li>- проведення «роботи над помилками» - фіксація і пошук рішення помилок, допущених в проекті</li> <li>- управління знаннями - збір корисних практик, успішних шаблонів зберігання даних</li> </ul>
6	Трансляція досвіду всередині компанії	<ul style="list-style-type: none"> <li>- комунікація всередині компанії</li> <li>- реалізація подібних проектів з досвідченою командою</li> </ul>

За підсумками Розділу 3 були виділені дві практики, гарантовано ефективні для реалізації Big Data проекту - Стенд-ап і Канбан-дошка. Дані методики часто виконуються неправильно, що призводить до їх неефективності. Далі буду описані основні помилки у використанні даних методик і запропоновані рекомендації щодо їх виправлення, а також можливі адаптації для покращення застосування практик у проектах з обробки великих даних.

### **Практичне застосування практики Стенд-ап.**

Основні помилки (як ми це робимо) при використанні практики Стенд-ап полягають в наступному:

#### **1. Занадто велика кількість учасників.**

Кожен учасник повинен бути активно залучений і відповідати на три питання - Що я робив вчора? Що я роблю сьогодні? Які перешкоди мені заважають? Якщо на зустрічі є «слухачі», то зустріч проходить неефективно, тому що проходить робочий час людей, не потрібних на Стенд-апі. З цієї причини в Стенд-апі беруть участь тільки керівники команд, скрам-майстер, якщо він є, і бізнес-експерти. Кількість учасників обмежується до 6.

#### **2. Ієрархічна, а не лінійна взаємодія учасників.**

Якщо зустріч побудована ієрархічними чином, де учасники звітують перед керівництвом, зустріч проходить неефективно, тому що губляться міжкомандні, кросфункціональні теми для обговорення, в тому числі потенційні проблеми. Якщо куратор зустрічі бачить, що поки один з учасників говорить, інші займаються своїми справами, це значить, що зустріч проходить неефективно. В такому випадку завданням куратора зустрічі є налагодити комунікацію «учасник, який виступає - інші».

#### **3. Гнучкий формат зустрічі.**

Важливо, щоб формат зустрічі був жорстко обмежений структурою трьох питань з дотримуваним таймінгом. В іншому випадку необхідний обсяг інформації не буде обговорено в заплановані 15 хвилин, зустріч пройде неефективно. Всі питання, що виникають під час Стенд-апу, в тому числі і нові завдання, повинні бути обговорені на окремій зустрічі.

## **Практичне застосування практики Канбан-дошка.**

Основні помилки (як ми це робимо) при використанні методики Канбан-дошка полягають в наступному:

1. Відсутність обмежень на кількість завдань на одній стадії.

Якщо немає обмеження на кількість завдань на одній стадії, швидкість реалізації проекту буде загальмована, тому що кілька завдань будуть в роботі одночасно. Для вирішення даної проблеми необхідно пріоритезувати завдання, щоб їх було менше, але при цьому висока швидкість реалізації зберігалася б.

2. Відсутність будь-яких завдань проекту на Канбан-дошці.

Якщо якийсь відділ не представлений на дошці, наприклад, представлена аналітика, але не представлено тестування, інструмент не буде правильно працювати, тому що реалізація проекту не буде достовірно проілюстрована. Перед запуском і використанням інструменту необхідно переконатися, що завдання всіх учасників проекту враховані.

3. Відсутність зворотного зв'язку - аналізу своєї роботи і роботи команди.

Кожен цикл через Канбан-дошку має на увазі наявність зворотного зв'язку як про свої роботи (як можна поліпшити свій шматок процесу), так і про роботи решти команди - взаємодія, швидкість комунікації, наявність проблем, що викликають затримки в робочому процесі. Система Канбан має на увазі «постійне поліпшення», що не представляється можливим без наявності озвучених областей для поліпшення всередині команди.

Спеціалізоване ПО дозволяє покрити частину проектного управління при використанні гнучких методологій, а саме:

- формування методології проекту - онлайн Канбан-дошка та візуалізація спринтів, занурення команди в методологію при навчанні системи;
- дотримання гнучкої культури - прозорість виконання процесу (всі учасники можуть побачити всю інформацію щодо процесу), наявність загального простору для комунікації, можливість швидкого зворотного зв'язку;

- збір даних про виконання проекту - налаштування метрик, використання кращих практик і шаблонів, організоване зберігання даних за проектом.

Приклади практичного застосування гнучких методик при управлінні ІТ-проектів за допомогою спеціалізованого програмного забезпечення (ПЗ) представлені нижче:

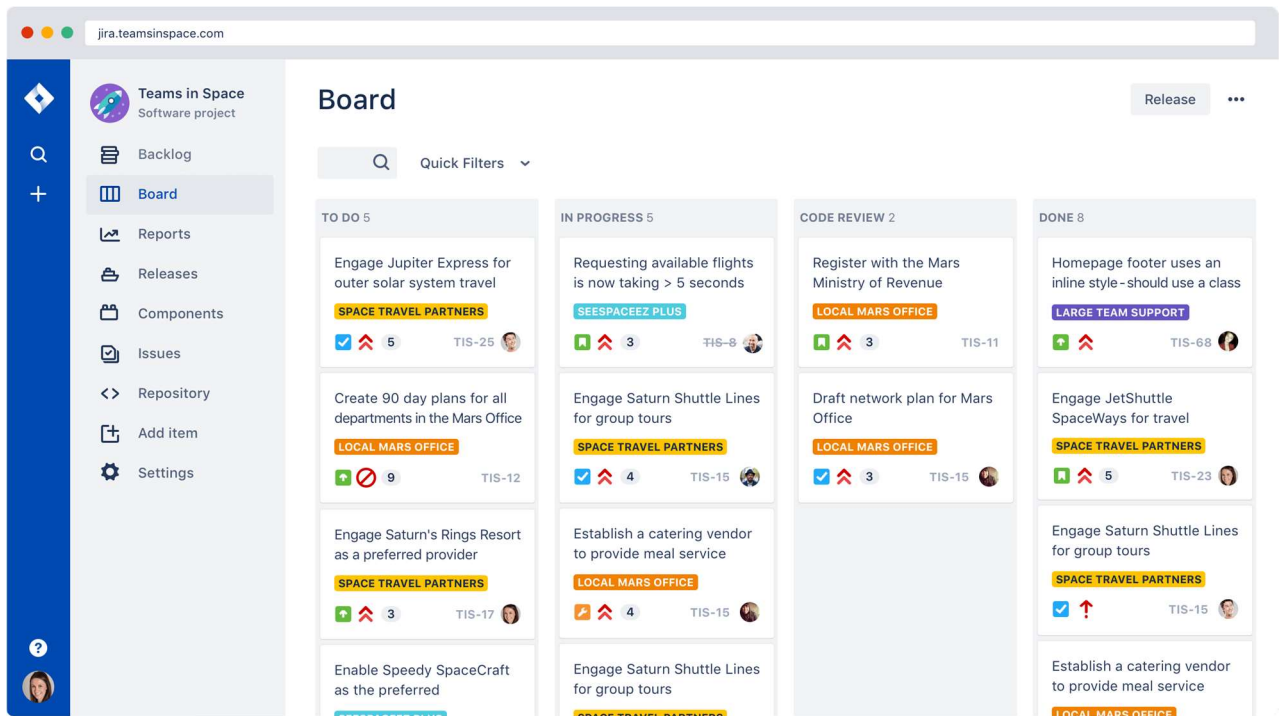


Рисунок 4.1 – Демонстрація інтерфейсу ПЗ Atlassian Jira [59]

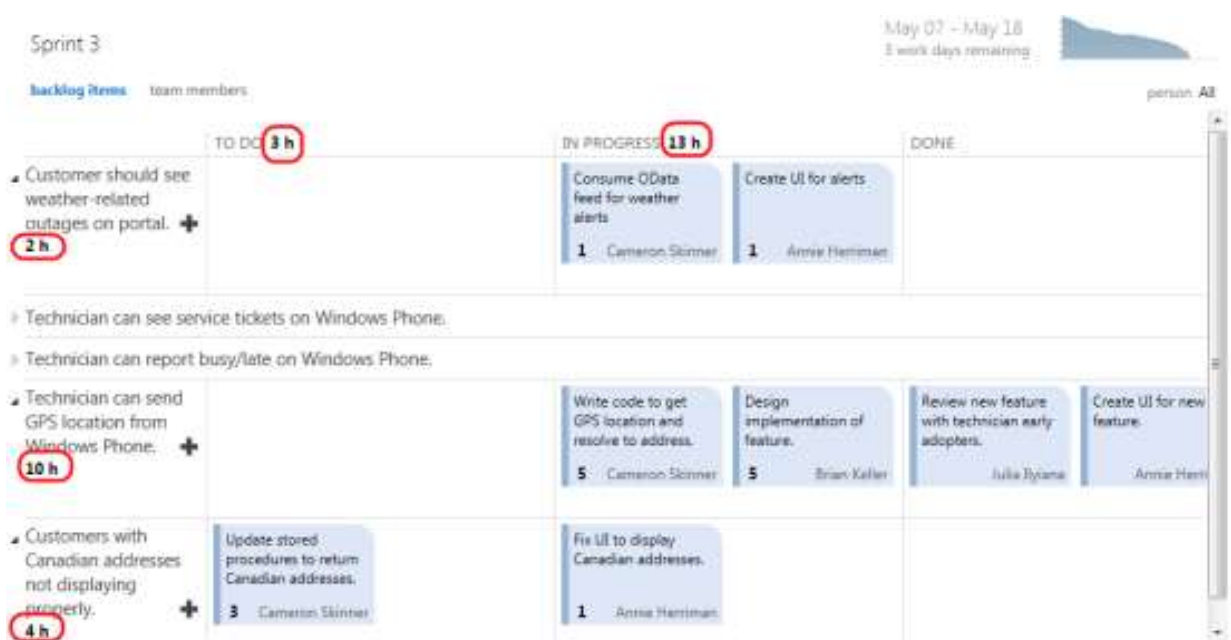


Рисунок 4.2 – Демонстрація інтерфейсу ПЗ Microsoft Azure DevOps Server [58]

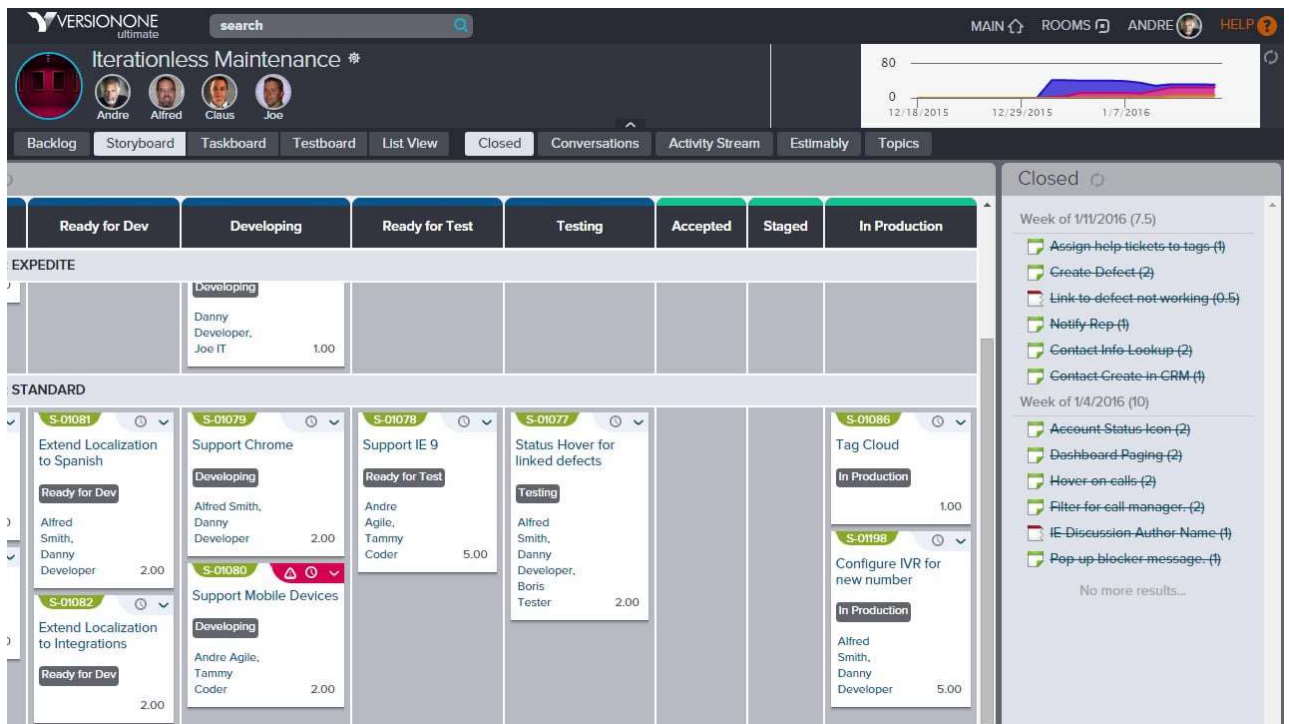


Рисунок 4.3 – Демонстрація інтерфейсу ПЗ VersionOne [60]

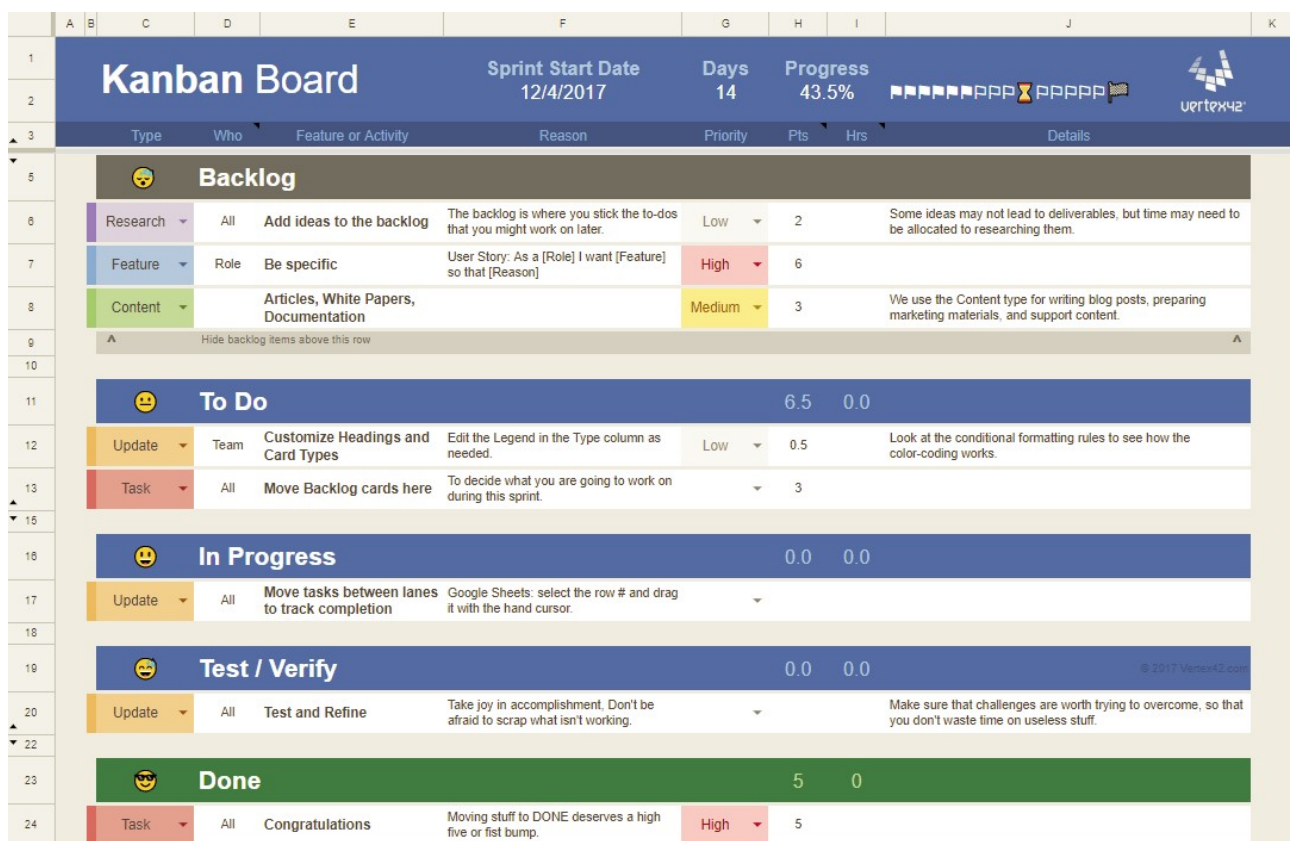


Рисунок 4.4 – Демонстрація інтерфейсу ПЗ Microsoft Excel, налаштованого у вигляді Канбан-дошки [57]

Використання канбан-дошки для проектів Big Data може бути дещо адаптовано та покращено.

Окрім традиційних трьох стовпців - Готово, Виконується та Зробити, нам треба додати додаткові стовпці: підготовка даних, в розробці, кодування, тестування тощо. Також, фази моделювання і підготовка даних необхідно об'єднати в одну фазу, тому що інакше дуже незручно рухати задачі по канбану: у цій варіації робота з даними та експерименти відбуваються в один і той же момент.

Нижче буде представлено схему роботи з канбан-дошкою для проекту з обробки великих даних:



Рисунок 4.5 – Використання канбан-дошки у проектах Big Data

Основні елементи схеми:

- Рядки - проекти або напрямки. Наприклад, частина робіт може бути пов'язана з автоматизацією Пайплайн, а не з самими експериментами.
- Елементи - гіпотези або групи питань. Якщо гіпотезу зі зрозумілою метрикою ще не вдається сформулювати, можна просто об'єднати кілька питань, щоб аналітик даних спробував знайти відповіді на них в даних.

- Черга з пріоритезованих гіпотез. Розставити пріоритет - завдання product owner.
- Experimenting: підготовка даних і моделювання.
- Discovery. Далеко не всі гіпотези доберуться до етапу реалізації, тому у останньому розділі виходить більш-менш пусто.

Ось так може виглядати приклад впровадження канбан-дошки у проект Big Data з використанням спеціалізованого програмного забезпечення:

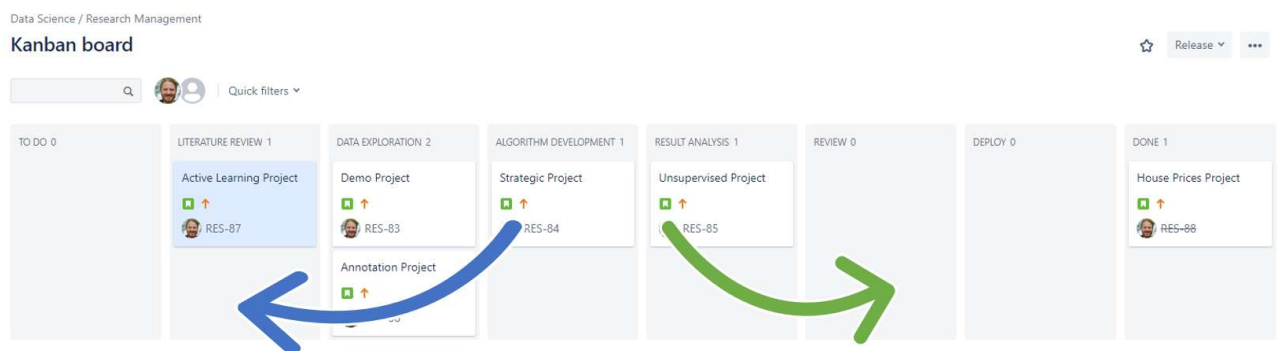


Рисунок 4.6 – Приклад використання допрацьованої канбан-дошки на проекті Big Data

## 4.2 Практичний досвід з впровадження гнучких методологій у Big Data команді

Під час проходження переддипломної практики у ТОВ «Проперті Маркет» були проведені спостереження та участь у роботі над проектом з великими даними у Agile команді. Виходячи з цього, можна підсумувати необхідні критерії для успішного застосування гнучких методик та методологій на проектах Big Data.

Процес розробки Agile фундаментально заснований на Agile Manifesto, який окреслює його відмінність від традиційних підходів до розробки програмного забезпечення. Agile дотримується ітеративної та поступової методології у своєму потоці.

Agile Manifesto робить акцент на взаємодії між інструментами, програмному забезпеченні над документацією, співпраці із замовником під час

переговорів про контракт та просто реагуванні на наступне. Agile процес нахилиє рівновагу до кращих та поточних аспектів розвитку. Під час проходження практики було взято участь у переговорах з замовниками, зустрічах команди, плануванні процесу розробки та складанні беклогу та багатьох інших Agile-подіях. Ці процеси якнайкраще впливають на самоорганізованість та розвиток команди, злагожену та швидку роботу та задоволеність замовника.

Великі дані формують цілу сферу даних, і їх важко визначити через свою складну структуру, розмір та різноманітність. Простіше кажучи, можна сказати, що вони включають всі дані з точки зору різноманітності, обсягу та швидкості їх створення.

Зараз, коли справа стосується аналізу великих даних, роль процесу Agile широко розглядається. Але основними проблемами прийняття підходу Agile у проектах з обробки великих даних є:

- конституція правильної суміші в команді, а саме: менеджери, розробники та аналітики;
- визначення обсягу проекту щодо кількості доступних джерел даних;
- межі безпеки проектів, заснованих на розповсюдженні даних.

Правильним способом інтеграції проектів великих даних буде чітке врахування вищезазначених факторів та дотримання п'ятиступеневого процесу:

1. Чітке визначення проблеми.
2. Виявлення розриву в знаннях.
3. Формулювання гіпотези.
4. Перевірка гіпотези.
5. Повторіть, тобто продовжуйте цикл.

Крокам цього процесу слідують у «Проперті Маркет», що дозволяє досягати значного успіху у проектах.

Проте під час практики було визначено і важливі своєрідні виклики інтеграції процесу Agile до аналізу великих даних.

Однією з найважливіших проблем в аналізі великих даних з використанням Agile-філософії є визначення сфер інтеграції. Параметри науки про дані повинні відповідати Agile процесам для розробки значущих результатів.

Методологія Agile потребує, щоб команди проєктів були чіткими, компактними та швидкими, що досить просто в проєктах з розробки програмного забезпечення чи мобільних додатків. Але для реалізації проєктів Big Data наука про дані з усією кількістю аналітиків, дизайнерів, розробників бізнесу, менеджерів, аналітиків даних тощо створює велику команду.

Ще однією проблемою є перекриття функцій різних команд.



Рисунок 4.7– Структура команди Agile Big Data

Наприклад, рисунок 4.7 вище демонструє, що присвоєння функції зберігання необроблених даних може входити в рамки інженера платформи, а

також прикладного дослідника. Чим більше людей виконує одне завдання, тим ускладнюється спілкування між ними. Все це призводить до проблеми збільшення терміну доставки.

Проте досвід приймаючої компанії говорить про деякі існуючі методи інтеграції Agile Framework та науки про дані.

Найкращий спосіб інтегрувати Agile Framework у науку даних, щоб зробити процес аналізу великих даних гнучким, - сприймати зміни з певною проактивністю.

Першим кроком до того, щоб зробити взаємодію гладкою, є зміщення переваг на користь загальних спеціалістів.

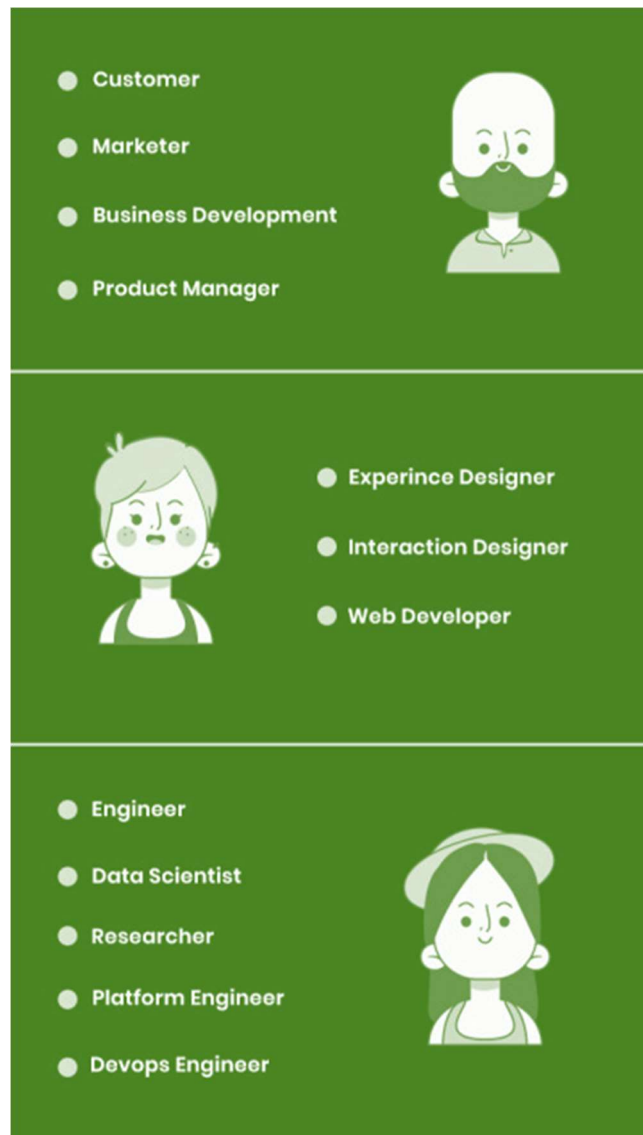


Рисунок 4.8 - Спеціалісти є головними в командах Agile Big Data

Це означає, що при створенні гнучких команд для аналізу великих даних буде зосереджено увагу як на глибині, так і на ширині знань серед членів команди, ніж на спеціалізації. Це не означає, що фахівці залишаться поза увагою, але як міст для заповнення місць у робочому підрозділі, як очікується, спеціалісти відіграватимуть більш значну роль.

Отже, досвідчена Agile команда Big Data матиме:

1. Аналітиків даних, які добре працюють із створенням додатків та послуг в Інтернеті, а також з функціями дослідження.
2. Веб-розробники, які розуміють UI / UX, а також можливість створювати додатки.
3. Дослідники, які можуть перевірити вихідний код, пояснити та обробляти дані.
4. Менеджери продуктів, які можуть обробляти всі пов'язані домени.

Окрім цього, можна підсумувати деякі основні принципи інтеграції команди Big Data Analytics:

1. Розділіть команду проекту на менші групи.
2. Слідкуйте за станом найсучасніших технологій для прискорення процесів, високорівневими інструментами та платформами, такими як PaaS (Платформа як послуга, хмарні обчислення, розподілені системи)
3. Діліться посередницькою роботою, тобто даними постійно, навіть якщо вони є неповними, оскільки характер аналізу великих даних вимагає регулярного періодичного обміну.

Після подолання основних перешкод Agile команда почне функціонувати, як описано нижче.

Таким чином, практична команда Agile Big Data складатиметься з невеликої групи спеціалістів, що виступають мостом між іншими членами. Загальні спеціалісти використовуватимуть масштабовані висококласні інструменти та хмарні обчислення для виконання ітерацій, які додають значення вихідним даним і, таким чином, уточнюють їх для вищого значення.

### 4.3 Подальше дослідження

З огляду теоретичного досвіду по темі подальше дослідження слід направити на вимір ефекту впровадження тієї чи іншої методики гнучких методологій на успіх Big Data проекту.

Незважаючи на досить великий досвід вивчення теми в академічному світі, проектний менеджер зазнає труднощів в:

- 1) ознайомленні з методологією;
- 2) з'ясуванням того, чи найбільш ефективною буде саме ця методологія на проекті Big Data;
- 3) з впровадженням методології на проекті і в компанії.

Для практичних цілей необхідно розробити мануал, який включити в себе, з одного боку, теоретичні знання, з іншого боку, досвід помилок і їх виправлень в бізнесі. Для цього необхідно зібрати базу даних успішних і неуспішних кейсів використання гнучких методологій на проектах Big Data як в чистому вигляді, так і окремих методик і їх комбінацій. Досить велика база знань дозволить виділити патерни і особливості використання гнучких методів і практик в залежності від:

- регіону проекту/компанії;
- індустрії проекту;
- віку компанії;
- співвідношення приватного і державного капіталу в компанії;
- характеристик замовника
- та ін.

Така база знань допоможе відповісти на три не до кінця вивчених дослідних питання:

1. Чи існують «універсальні» методики гнучких методологій, гарантовано збільшують ймовірність успіху проекту Big Data незалежно від його характеристик?
2. Від чого залежить ефективність застосування гнучких методологій?

3. Чи є вичерпний список характеристик проекту Big Data, які не можуть співіснувати з гнучкими методологіями?

#### **Висновки до Розділу 4**

У цьому розділі були описані практичні рекомендації щодо використання методик гнучких методологій на проектах Big Data і по збільшенню їх ефективності зокрема. Незважаючи на доведену ефективність гнучких методик, часто вони застосовуються неправильно і виявляються неспроможними. Були описані основні помилки при їх застосуванні і те, як їх можна уникнути. Також була допрацьована методика Канбан-дошка та запропонований варіант її застосування в проектах з обробки великих даних.

У розділі було окреслено також основні здобутки проходження переддипломної практики та описано досвід впровадження гнучких методологій у проекти Big Data на бази практики. Основними критеріями успішності такої команди є кросфункціональність, відкритість до змін та використання сучасних технологій.

Також було окреслено коло тем, найбільш актуальних для дослідження з даної проблематики: фактори, що впливають на успішність застосування методик гнучких методологій у проектах Big Data; пошук і доказ наявності «універсальних» ефективних методик гнучких методологій; існування вичерпного переліку «стоп-характеристик» проекту Big Data, при яких застосування гнучких методологій недоцільно.

## ВИСНОВКИ

У сучасній ері інформатизації прискорення зростання даних є об'єктивною реальністю. Соціальні медіа, мобільні пристрої, дані з вимірювальних пристроїв, ділова інформація - лише кілька джерел, які можуть генерувати гігантські обсяги інформації. Тому проекти Big Data є окремим та одним з ключових напрямків в ІТ-галузі. Проектна діяльність у Big Data значно зростає, проте кількість невдалих проектів все ще залишається високою. Одним з важливих критеріїв успіху є вибір правильної методології управління проектами.

У даній роботі було розглянуто традиційні та гнучкі методології управління проектами як підходи до управління проектами Big Data. За статистичними даними гнучкі підходи мають на 18% більше успішних проектів і на 20% менше неуспішних ніж традиційні послідовні підходи. Враховуючи основну проблему проектів Big Data - нерозуміння між ними та цілями бізнесу, а також такі специфічні проблеми як величезний об'єм даних та неперервні зміни, у роботі було вирішено дослідити процес управління проектами Big Data саме з допомогою Agile методологій. А оскільки це відносно нова галузь і для великих даних може знадобитися щось нове або, принаймні, комбінація стандартних підходів, було вирішено оцінити вплив саме конкретних методик гнучких методологій на успіх проектів Big Data.

У дипломній роботі було проведено аналіз академічної літератури, за допомогою якого описано принцип використання гнучких методологій у проектній діяльності, успішність та особливості цього підходу. Були визначені особливості управління проектами Big Data та механізми, що звичайно застосовуються до таких проектів, а також описано як саме та чому Agile методології можуть бути застосовані для них. Можна зробити висновок, що у проектах з такою кількістю невизначеності, як у проектах Big Data, застосування гнучких методологій чи принаймні їх методик буде дуже вдалим через їх підхід до внесення змін у проект та «гнучкість» практики загалом.

Було розглянуто й описано практичне визначення поняття та критеріїв успіху проекту. Проаналізовано роботи більше 25 авторів і сформовано 33 основних критерія успішності проектів, що вони виділяють. Для гнучких методологій кількість критеріїв становила 38. У роботі виявлено основні методики гнучких методологій та проведено експертне оцінювання впливу методик на успіх проектів Big Data. Всього були визначено 30 основних методик з таких методологій як Scrum, Kanban, XP, ScrumBan та XP Hybrid Scrum.

Для проведення експертного оцінювання відібраних методик було залучено 30 експертів. За статистикою 37% експертів мають 2-5 років досвіду роботи з проектами Big Data, 27% мають 5-7 років досвіду і 23% експертів - менше 2 років. Найбільш популярна роль серед експертів - менеджер проекту (50% відповідей). Таким чином, в середньому відповіді представляли думку керівництва проекту з середнім досвідом роботи на 2-4 проектах. За статистикою, що описує компанію, для якої/в якій здійснювався Big Data проект, можна побачити, що відповіді в середньому будуть характеризувати компанію українського походження (70% відповідей), що є великим підприємством (83%), віком понад 10 років (87%). Галузь компанії в більшості випадків представлена фінансовим сектором (40% відповідей) та ІТ-та телеком компаніями (27%).

За 5-ти бальною шкалою Лайкерта було також оцінено рівень відкритості до використання гнучких методологій і наявність експертів в області, що є одними з головних чинників успішного використання гнучких методів і методик. Середня відповідь склала 3,67 і 2,9 відповідно, що говорить про те, що в цілому компанії у вигляді керівництва починають усвідомлювати і приймати ефективність Agile методологій, але в той же час в компанії не вистачає експертних знань для правильного їх використання. Ці фактори були включені у подальший аналіз.

Під час проведення експертного оцінювання з відібраних методик експертами було визначено 6 найефективніших, які були також включені у регресійну модель. Результатами оцінювання став відбір таких методик як

беклог - 25 голосів, стенд-ап - 24, користувацькі історії - 21, спринт - 20, канбан-дошка - 19 і ретроспектива - 18.

Експертами була надана оцінка успіху проекту Big Data, в управлінні якого використовувалися гнучкі методики. Спочатку обиралися характеристики успіху проекту, що входять в офіційні критерії успіху, а потім вони оцінювалися всі за 5-ти бальною шкалою Лайкерта. Найбільш популярними критеріями виявилися параметри «проектного трикутника» - термін – 26 голосів, бюджет – 23, обсяг – 18, якість - 18. При побудові регресійної моделі кожному з критеріїв були присвоєні ваги відповідно до частоти їх використання при офіційній оцінці успіху проекту. Усереднена відповідь експерта становила: якість – 3.7, задоволеність замовника – 3.77, задоволеність користувача – 3.83.

Найбільшими стоп-факторами, що перешкоджають ефективному веденню проектів Big Data гнучкими методологіями були виділені: неготовність керівництва – 4.53, консервативна культура компанії 4.27, відсутність необхідних експертів – 3.67.

Також було зібрано цінні коментарі у відкритій формі від експертів щодо використання гнучких методик у проектах Big Data. Згідно до них можна зробити висновок, що для Big Data проектів є корисними ітеративний підхід і швидкість реагування на зміни Agile методик. Гнучка методологія надає аналітикам даних можливість визначати пріоритети моделей та даних відповідно до цілей та вимог проекту і, що важливо, бізнесу. Також важливими аспектами є гнучкість в плануванні, адаптивність методології та орієнтація на замовника. Зібрані дані переконують нас у можливості та ефективності використання Agile у Big Data.

Для оцінки впливу методик гнучких методологій на успіх проектів Big Data побудовано регресійну модель з використанням методу найменших квадратів для аналізу отриманих з експертного опитування даних та визначення ефективних методик гнучких методологій для Big Data проекту. Була розглянута і описана теоретична база, на яку спирається метод найменших квадратів. Результатами регресійного аналізу стало визначення двох найбільш ефективних методик – Стенд-ап і Канбан-дошка. При чому значення R-квадрат становило

68%, що говорить про нормальну пояснювальну силу регресії. На другій ітерації було визначено, що при використанні Стенд-апу або Канбан-дошки консолідована змінна успіху проекту збільшується на 3,6 балів. Важливо, що дані методики дійсно можливо використовувати окремо, вони не вимагають поглиблених знань в області Agile. Було виявлено й інший значимий фактор, що впливає на успіх проекту, а саме відкритість компанії до використання гнучких методологій: при збільшенні категоріальної змінної «відкритість» на 1 бал змінна «успіх» збільшується на 2,28 балів. Даний результат підтверджує теоретична основа гнучких методологій: Agile-культура, перш за все. Регресійна модель не показала наявності статистичної значущої залежності між показником успіху проекту Big Data і використанням інших гнучких методик (для них P-value > 0,05). Однак, не можна однозначно зробити висновок, що інші методики не мають ефекту, питання вимагає вивчення в рамках подальших досліджень.

Хотілося б звернути увагу на те, що ціллю даної роботи не стояв доказ використання гнучких методологій у проектах Big Data в цілому. Ціль дослідження - показати, чи існують сформовані комплекси методик гнучких методологій, що гарантовано впливають на успіх проекту Big Data, і оцінити їх.

Можна зробити висновок, що в Україні визнають можливість використання гнучких методологій для проектів Big Data, але фахівці поки не володіють достатньою кваліфікацією для їх успішного використання. До того ж адаптація методології та застосування різних методик вимагає ще більш глибоких знань за тематикою. Існуюча тенденція полягає в тому, що на проектах Big Data часто намагаються механічно ввести використання гнучких методологій без адаптації самої методології та відповідної зміни в корпоративній культурі і штаті компанії. Такий метод в більшості випадків виявляється неефективним.

Практичними рекомендаціями за підсумками роботи були наведені дії, що виконує проектна команда при реалізації Big Data проекту і на що необхідно звернути увагу з точки зору того, як це виконується. Наприклад, при виборі критеріїв успіху проекту - залучення всіх стейкхолдерів у формування поняття «успіх даного проекту Big Data» і включення не тільки кількісних, а й якісних

показників. При підготовці рекомендацій був використаний практичний досвід реалізації Big Data проекту за допомогою гнучких методологій.

У роботі було визначено, що серед особливостей успішної Agile-команди у проектах Big Data є необхідність слідувати таким принципам як невеликий обсяг спеціалістів, кросфункціональність команди та використання сучасних технологій. Кросфункціональні спеціалісти мають використовувати масштабовані висококласні інструменти та хмарні обчислення для задач. Таким чином команда Big Data проекту буде достатньо компактною та професійною для вдалого використання Agile. Було адаптовано і допрацьовано класичну канбан-дошку для роботи з Big Data проектами та запропоновано схему роботи з нею. Такий адаптивний підхід можна застосувати до будь-якої гнучкої методології чи методики та використовувати тільки кращі для проекту їх аспекти.

Коло тем, найбільш актуальних для подальшого дослідження з даної проблематики включає в себе фактори, що впливають на успішність застосування методик гнучких методологій у проектах Big Data; пошук і доказ наявності «універсальних» ефективних методик гнучких методологій; існування вичерпного переліку «стоп-характеристик» проекту Big Data, при яких застосування гнучких методологій недоцільно.

Таким чином, цілі і завдання даної кваліфікаційної роботи магістра були виконані, було знайдено і оцінено в кількісному вигляді вплив використання методик Стенд-ап і Канбан-дошка на показник успіху Big Data проекту.

Важко знайти одну методологію, яка б чудово працювала протягом життєвого циклу проекту в галузі Big Data. Проекти з обробки великих даних можуть поєднувати в управлінні ними декілька інструментів різних методологій. Дослідження показало, які саме інструменти впливають на успіх Big Data проектів. Їх синтез у єдиний підхід зможе позитивно відобразитися на результатах проекту. На закінчення хотілося б відзначити, що для успішної адаптації і складання своєї методики з різних гнучких елементів, необхідна накопичена експертиза і відповідний технічний і соціальний ландшафт і клімат в компанії.

## ПЕРЕЛІК ВИКОРИСТАНИХ ІНФОРМАЦІЙНИХ ДЖЕРЕЛ

1. The Kübler-Ross factor in managing the performance of technical and socio-economic systems / D. Lukianov, K. Kolesnikova, O. Mezentseva, V. Rudenko // Scientific Journal of Astana IT University, Vol.2, 2020, Pp. 32–43.
2. Influence Analysis of Different Management Methodologies on the Result of Big Data Projects / V. Rudenko, O. Mezentseva // 7th International conference "Information Technology and Interactions (Satellite: Conference Proceedings, December 04, 2020, Kyiv, Ukraine / Taras Shevchenko National University of Kyiv and [etc]; Vitaliy Snytyuk (Editor). Kyiv: Stylos, 2020.– Pp. 165-167.
3. Kolesnikova, K., Rudenko V., Mezentseva, O., Project management methodologies techniques in Big Data analysis. Herald of Advanced Information Technology. Vol.2 No. 3.
4. Mike Barlow, Real-Time Big Data Analytics: Emerging Architecture - O'Reilly Media, 2013, P. 55.
5. Trevor Hastie, Robert Tibshirani, Jerome Friedman. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. - Second Edition. February 2009, P. 72.
6. DJ Patil, Building Data Science Teams. - O'Reilly Media, 2011, P. 81.
7. CHAOS Report. The Standish Group International, Inc. 2015. – Режим доступу:  
[https://www.standishgroup.com/sample\\_research\\_files/CHAOSReport2015-Final.pdf](https://www.standishgroup.com/sample_research_files/CHAOSReport2015-Final.pdf)  
– (Дата звертання: 01.12.2020).
8. The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East. Доступно за адресою:<https://www.emc.com/leadership/digital-universe/2012iview/index.html>.  
(Дата звернення: 14.12.2020)
9. The 12th Annual State of Agile Report. CollabNet VersionOne. 2018. Available at: <https://explore.versionone.com/state-of-agile/13th-annual-state-of-agile-report>. (Дата звертання: 04.12.2020)

10. Highsmith, J. Agile Project Management: Creating Innovative Products - Addison-Wesley, Upper Saddle River (NJ), 2010.
11. Hällgren, M., Maaninen-Olsson, E. Deviations, ambiguity and uncertainty in a project-intensive organization // Project Management Journal. 2005. Vol. 36 (3). c. 17–26.
12. Collyer, S., Warren, C. M. J., Hemsley, B., Stevens, C. Aim fire aim – project planning styles in dynamic environments // Project Management Journal. 2010. Vol. 41 (4). c. 108–121.
13. Project Management Institute. A Guide to the Project Management Body of Knowledge – Fourth Edition, Project Management Institute Inc., 2004.
14. Cusumano, M.A., Selby, R.W. How Microsoft Builds Software // Communications of the ACM. 1997. Vol. 40 (6).
15. Project Management Institute. A Guide to the Project Management Body of Knowledge – Fourth Edition, Project Management Institute Inc., 2008.
16. Project Management Institute. A Guide to the Project Management Body of Knowledge – Fifth Edition, Project Management Institute Inc., 2013.
17. Smite, D., Moe, N.B., Agerfalk, P.J. Fundamentals of agile distributed software development. Agility Across Time and Space. - Springer-Verlag, Berlin, Heidelberg. 2010
18. Moe, T. L., Khang, D. B. Success Criteria and Factors for International Development Projects: A Life-Cycle-Based Framework // Project Management Journal. 2008. Vol. 39 (1). P. 72–84.
19. Vallona, R., José da Silva Estácioc, B., Prikladnickic, R., Grechenigb, T. Systematic literature review on agile practices in global Data Science development // Information and Software Technology. 2018. Vol. 96. P. 161–180.
20. Chow, T., Cao, D.B. A survey study of critical success factors in agile Big Data projects // The Journal of Systems and Software. 2008. Vol. 81. P. 961–971.
21. Sheffield J., Lemétayer, J. Factors associated with the software development agility of successful projects // International Journal of Project Management. 2013. 31. 459–472.

22. Takeuchi, H.; Nonaka, I. The New New Product Development Game. [Електронний ресурс]: Harvard Business Review. 1986. URL: <http://damiantgordon.com/Methodologies/Papers/The%20New%20Product%20Development%20Game.pdf> (Дата звернення: 01.02.2021).
23. Schwaber, K.; Beedle, M. Agile software development with Scrum - Prentice Hall, 2002.
24. Kniberg, H., Skarin, M. Kanban and Scrum – Making the Most of Both. [Електронний ресурс]: Lulu, 2010. URL: <https://lulu.com> (Дата звернення: 01.02.2021).
25. Akif, R., Majeed, H. Issues and Challenges in Scrum Implementation // International Journal of Scientific & Engineering Research. 2012. Vol. 3 (8).
26. Turk, D., France, R., Rumpe, B. Limitations of Agile Software Processes // Proceedings of the Third International Conference on Extreme Programming and Flexible Processes in Software Engineering. .2002.
27. Anderson, D. Kanban-Successful Evolutionary Change for Your Technology Business - WA: David J. Anderson & Associates Inc, Seattle, 2010.
28. Sipper, D., Bulfin, R.L. Variations of the kanban system: Literature review and classification // International Journal of Production Economics. 2010. Vol. 125(1). P. 13-21.
29. Кент Бек: Екстремальне програмування - Пітер, 2002. 224 с.
30. Boehm, B., Turner, R. Balancing Agility and Discipline: A Guide for the Perplexed // Boston, MA: Addison-Wesley, 2004.
31. Ladas, C. Scrumban: Essays on Kanban Systems for Lean Software Development - Modus Cooperandi Press. 2009
32. Reddy, A. Scrumban [R] Evolution- Get the most out of Agile, Scrum and Lean Kanban - Pearson, 2016
33. Mushtaq Z., Rizwan Jameel Qureshi Novel, M. Hybrid Model: Integrating Scrum and XP [Електронний ресурс]: International Journal of Information Technology and Computer Science. 2012. Vol. 6. P. 39-44. URL: <http://www.mecspress.org/> (Дата звернення: 01.04.21).

34. Albert, M., Balve, P., Spang, K. Evaluation of project success: a structured literature review // *International Journal of Managing Projects in Business*. 2017. Vol. 10 (4). P. 796–821.
35. Alami, A. Why do Information Technology Projects Fail? // *Procedia Computer Science*. 2016. Vol. 100, P. 62–71.
36. Chang, A., Chih, Y. Y., Chew, E, Pisarski, A. Reconceptualising mega project success in Australian Defence: Recognising the importance of value co-creation // *International Journal of Project Management*. 2013. Vol. 31 (8). P. 1139–1153.
37. Baccarini, D., Collins, A. The concept of project success - what 150 Australian project managers think // *Australian Institute of Project Management (AIPM) Conference: Perth 2004, 10-12th October, 2004*.
38. Müller, R., Jugdev, K. Critical success factors in projects: Pinto, Slevin, and Prescott - the elucidation of project success // *International Journal of Managing Projects in Business*. 2012. Vol. 5 (4). P. 757-775.
39. Müller, R., Turner, R. The Influence of Project Managers on Project Success Criteria and Project Success by Project Type // *European Management Journal*. 2007. Vol. 25 (4). P. 298-309.
40. Davis, K. An empirical investigation into different stakeholder groups perception of project success // *International Journal of Project Management*. 2017. Vol. 35 (4). P. 604-617.
41. Atkinson, R. Project management Cost time and quality, two best guesses and a phenomenon, it's time to accept other success criteria // *International Journal of Project Management*. 1999. Vol. 17 (6). P. 337-342.
42. Kerzner, H. *Project Management: A Systems Approach To Planning, Scheduling, And Controlling* [Электронный ресурс]: John Wiley & Sons, Inc., 2003 URL: <https://books.mec.biz/tmp/books/55F1OL4WQC7HL2OBCGHS.pdf> (Дата звернення: 01.03.2021).
43. Baccarini, D. The logical framework method for defining project success // *Project Management Journal*. 1999. Vol. 30 (4). P. 25-32.

44. Shenhar, AJ, Dvir, D., Levy, O., Maltz, AC Project success: A multidimensional strategic concept // Long Range Planning, 2001. Vol. 34 (6). P. 699-725.
45. Ika, LA Project Success as a Topic in Project Management Journals // Project Management Journal. 2009. Vol. 40 (4). P. 6-19.
46. Diallo, A., Thuillier, D. The success dimensions of international development projects: the perceptions of African project coordinators // International Journal of Project Management. 2004. Vol. 22. P. 19-31.
47. Castro, MS, Bahli, B .; Farias Filho, JR et al. A contemporary vision of project success criteria. [Электронный ресурс]: Brazilian Journal of Operations & Production Management. 2019. Vol. 16 (1). P. 66-77. URL:<https://bjopm.emnuvens.com.br/bjopm/article/view/723> (Дата звращения: 01.04.2021).
48. Pankratz, O., Basten, D. Ladder to success - eliciting project managers 'perceptions of IS project success criteria // International Journal of Information Systems and Project Management. 2014. Vol. 2 (2). P. 5-24.
49. Thomas, G. Fernández, W. Success in IT projects: A matter of definition? // International Journal of Project Management. 2008. Vol. 26 (7). P. 733-742.
50. Todorović, ML, Petrović, D. Č., Mihić, MM, Obradović, VL, Bushuyev, SD Project success analysis framework: A knowledge-based approach in project management // International Journal of Project Management. 2015. Vol. 33 (4). P. 772-783.
51. Hassan, MM, Bashir, S., Abbas, SM The Impact of Project Managers 'Personality on Project Success in NGOs: The Mediating Role of Transformational Leadership // Project Management Journal. 2017. Vol. 48 (2). P. 74-87.
52. Mazur, AK, Pisarski, A. Major project managers 'internal and external stakeholder relationships: The development and validation of measurement scales // International Journal of Project Management. 2015. Vol. 33 (8). P. 1680-1691.
53. Hussein, BA, Ahmad, SBS, Zidane, YJT Problems Associated with Defining Project Success // Procedia Computer Science. 2015. Vol. 64. P. 940-947.

54. Agarwal, N., Rathod, U. Defining "success" for software projects: An exploratory revelation // International Journal of Project Management. 2006. Vol. 24 (4). P. 58-70.
55. Stephens, M., Rosenberg, D. The irony of extreme programming. - MA: Dr Dobbs journal. Apress LP 2004.
56. Ohno, T. Toyota Production System - beyond large-scale production // Productivity Press. 1988. Vol. 1, P. 29.
57. Приклади адаптації та використання Microsoft Excel для прикладних бізнес-задач [Електронний ресурс]: Vertex42 Blog URL: <https://www.vertex42.com/blog/> (Дата звернення: 17.04.2021).
58. Офіційний сайт Microsoft [Електронний ресурс]: URL: <https://azure.microsoft.com/en-us/services/devops/server/> (Дата звернення: 17.04.2021).
59. Офіційний сайт Atlassian [Електронний ресурс]: URL: <https://www.atlassian.com/company> (Дата звернення: 17.04.2021).
60. Офіційний сайт Collab.Net [Електронний ресурс]: URL: <https://www.collab.net/> (Дата звернення: 17.04.2021).
61. Wells, D. Extreme Programming. [Електронний ресурс]: Extreme Programming, 2000. URL:<http://www.extremeprogramming.org/> (Дата звернення: 01.04.2021).
62. Mushtaq Z., Rizwan Jameel Qureshi Novel, M. Hybrid Model: Integrating Scrum and XP [Електронний ресурс]: International Journal of Information Technology and Computer Science. 2012. Vol. 6. P. 39-44. URL:<http://www.mecspress.org/> (Дата звернення: 01.04.2021).
63. Scrumtrek - кейси виконання проектів з використанням гнучких методологій [Електронний ресурс]: ScrumTrek. URL:<https://scrumtrek.ru/> (Дата звернення: 01.04.2021).
64. Takeuchi, H .; Nonaka, I. The New New Product Development Game. [Електронний ресурс]: Harvard Business Review. 1986.

URL:<http://damiantgordon.com/Methodologies/Papers/The%20New%20Product%20Development%20Game.pdf> (Дата звернення: 01.04.2021).

65. Beck, K., Grenning, J., Martin RC., Beedle, M., Highsmith, J., Mellor, S., Bennekum, A., Hunt, A., Schwaber, K., Cockburn, A., Jeffries, R., Sutherland, J., Cunningham, W., Kern, J., Thomas, D., Fowler, M., Marick B. Principles behind the Agile Manifesto. [Електронний ресурс]: Agile Manifesto. 2001. URL:<https://agilemanifesto.org/principles.html> (Дата звернення: 01.04.2021).

66. Dvir, D., Raz, T., Shenhar, A. An empirical analysis of the relationship between project planning and project success. International Journal of Project Management. 2003. Vol. 21 (2). P. 89-95.

67. Wysocki, RK Effective Project Management: Traditional, Agile, Extreme [Електронний ресурс]; John Wiley & Sons, Inc., 2011 URL:<http://index-of.co.uk/Project%20Management/Effective%20Project%20Management%20Traditional,%20Agile,%20Extreme%20by%20Robert%20K.%20Wysocki%207th%20Edition.pdf> (Дата звернення: 01.03.2021).

68. Manifesto for Agile Software Development [Електронний ресурс]: Agile Manifesto. 2001. URL:<http://agilemanifesto.org/iso/en/manifesto.html> (Дата звернення: 01.04.2021).

69. Шашков В.Б. Ш 12 Прикладной регрессионный анализ. Многофакторная регрессия: Учебное пособие. - Оренбург: ГОУ ВПО ОГУ, 2003. - 63 с.

70. Світлана Гуцикова. Метод експертних оцінок. Теорія і практика // Інститут психології РАН, 2011р. с. 56.

71. Anna Vichugova – CRISP-DM. [Електронний ресурс]: Big Data School. URL: <https://www.bigdataschool.ru/wiki/crisp-dm> (Дата звернення: 01.04.21).

## ДОДАТКИ

### Додаток А

Екземпляр для заповнення експертних оцінок «Використання гнучких методик в управлінні проектами Big Data»

#### **Розділ 1. Загальна інформація.**

##### ***Відомості про експерта***

1. Вкажіть, будь ласка, Ваш досвід роботи в галузі управління проектами:
  - Менше 2 років
  - 2-5 років
  - 5-7 років
  - 7-10 років
  - Більше 10 років
  
2. Вкажіть, будь ласка, в якій ролі Вам доводиться найчастіше брати участь в проектній діяльності у Вашій компанії (виберіть найбільш підходящий варіант):
  - Представник вищого керівництва організації
  - Керівник функціонального підрозділу
  - Менеджер проекту
  - Спонсор проекту
  - Адміністратор проекту
  - Замовник проекту
  - Виконавець (член команди проекту)
  - Інша роль (вкажіть яка саме)

---

##### ***Відомості про компанію (для / в якій здійснювався проект)***

3. Вкажіть назву компанії (якщо можливо)
- 

4. Компанія є ...
  - Українською
  - Зарубіжною
  - Компанією зі змішаним капіталом
  
5. Яка чисельність персоналу компанії?
  - Менше 15 чоловік (мікро-підприємство)
  - 15 – 100 чоловік (мале підприємство)
  - 101 - 250 чоловік (середнє підприємство)
  - 250+ чоловік (велике підприємство)

6. Скільки років компанія існує на ринку?

- Менше 2 років
- 2-5 років
- 5-7 років
- 7-10 років
- Більше 10 років

7. В якій галузі переважно працює компанія, в якій Ви працюєте:

- Будівництво
- Енергетика
- ІТ-та телеком
- Освіта
- Консультування
- Фінансовий сектор
- Нафтогазовий сектор
- Інше

(вкажіть)

8. Вибір оцінки, що більшою мірою виражає Вашу згоду з наведеними твердженнями (1 - абсолютно не згоден, 2 - не згоден, 3 - важко відповісти, 4 - згоден, 5 - абсолютно згоден)

Компанія (в т.ч. керівництво) відкрито до використання гнучких методологій по УП	1	2	3	4	5
Компанія має досвід у використанні гнучких методологій, в штаті є експерти в області гнучких методологій	1	2	3	4	5

### ***Відомості про проект***

9. Бюджет проекту:

- Малий (<1 000 000 грн.)
- Середній (більше 1 000 000 грн., менше 5 000 000 грн.)
- Великий (більше 5 000 000 грн.)

10. Чисельність команди проекту:

- Менше 50 чоловік
- 50-100 чоловік
- 100-500 чоловік
- Понад 500 чоловік

11. Локація команди:

- одна локація
- розподілена команда

12. Тривалість проекту:

- Менше 0,5 року
- 0,5 -1 рік
- 1-2 роки
- 2-5 років
- Більше 5 років

## Розділ 2. Використання гнучких методологій в управлінні проектами Big Data

13. Відзначте «X» методики гнучких методологій, які були використані на проекті:

1	SCRUM	2	KANBAN	3	XP*
	Беклог		Канбан-дошка		Гра в планування
	Користувацькі історії		Обмеження за кількістю завдань на одній стадії		Парне програмування
	Спринт				Розробка через тестування
	Стенд-ап				Швидкий зворотний зв'язок
	Оцінка КІ в балах				Неперервна інтеграція
	Scrum-ролі				Рефакторинг
	Ретроспектива				Невеликі, але часті релізи
					Стандарти написання коду
					Колективне володіння кодом
					Простий дизайн
					Системна метафора
					Фіксований робочий день для розробників
					Інше:

\* XP - Extreme Programming

4	SCRUMBAN	5	Інша гнучка методологія
	Короткі ітерації (1-2 тиж.)		
	Планування на вимогу		
	Планування «по кошиках»		
	«Замороження функцій»		
	Сортування залишившихся задач		

14. Відзначте «X» пари методик гнучких методологій, використаних в проекті:
- Стенд-ап - Канбан-дошка
  - Стенд-ап - Спринт
  - Стенд-ап - Ретроспектива
  - Беклог (розрахований по ПІ) - Канбан-дошка
  - Користувацькі історії - Канбан-дошка
  - Спринт - Канбан-дошка
  - Ретроспектива - Спринт
  - Ретроспектива - Канбан-дошка
  - Скрам-ролі - Канбан-дошка
  - Парне програмування - Канбан-дошка
  - Спринт - Неперервна інтеграція
  - Неперервна інтеграція - Канбан-дошка
  - Інше \_\_\_\_\_

### Розділ 3. Успіх проекту

15. Вибір оцінки, що більшою мірою виражає Вашу згоду з наведеними твердженнями (1 - абсолютно не згоден, 2 - не згоден, 3 - важко відповісти, 4 - згоден, 5 - абсолютно згоден).

Відзначте «X» пункти, що входять до офіційних критеріїв успіху проекту:

Проект був реалізований в запланований термін	1	2	3	4	5
Проект був реалізований в рамках планового бюджету	1	2	3	4	5
Фактичний обсяг проекту відповідав плановому (був не менший) планового	1	2	3	4	5
Якість продукту проекту відповідало заявленому (було не нижче заявленого)	1	2	3	4	5
Замовник задоволений результатом проекту	1	2	3	4	5
Кінцевий користувач задоволений результатом проекту	1	2	3	4	5
Команда задоволена проектом (процесом реалізації і результатом)	1	2	3	4	5
Прибутковість проекту відповідала плановій	1	2	3	4	5
Інше:	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5

16. Поставте «X» в перетині практики і критерію, якщо, на Ваш погляд, використання даної практики впливає на досягнення критерію.

	Термін	Бюджет	Якість	Об`єм	Задоволеність замовника	Задоволеність користувача	Задоволеність команди	Прибутковість	Не знаю / не впливає
<b>SCRUM</b>									
Беклог									
Користувацькі історії									
Спринт									
Стенд-ап									
Оцінка КІ в балах									
Scrum-ролі									
Ретроспектива									
<b>KANBAN</b>									
Канбан-дошка									
Обмеження за кількістю завдань на одній стадії									
<b>XP</b>									
Гра в планування									
Парне програмування									
Розробка через тестування									
Швидкий зворотний зв'язок									
Неперервна інтеграція									
Рефакторинг									
Невеликі, але часті релізи									
Стандарти написання коду									
Колективне володіння кодом									
Простий дизайн									
Системна метафора									
Фіксований робочий день для розробників									
<b>SCRUMBAN</b>									
Короткі ітерації (1-2 тиж.)									
Планування на вимогу									
Планування «по кошиках»									
«Замороження функцій»									
Сортування залишившихся задач									

17. Як ви вважаєте, чи дає використання гнучких методів управління проектами Big Data перевагу в порівнянні з традиційними методами (наприклад, з каскадної моделлю)? Дайте, будь ласка, пояснення.

Так, \_\_\_\_\_

---

Ні, \_\_\_\_\_

---



---

18. Який вплив, на Ваш погляд, той чи інший фактор надає на ефективне використання гнучких методик в проекті Big Data: 1 - незначний вплив, 2 - слабкий вплив, 3 - середній вплив, 4 - сильний вплив, 5 - дуже сильний вплив (виберіть один варіант для кожного показника):

Неготовність керівництва до ведення проекту з гнучкою методологією (особливе управління плануванням, змінами)	1	2	3	4	5
Культура компанії давно склалася і не схильна до змін (занадто консервативна)	1	2	3	4	5
Розмір команди занадто великий, не налаштована комунікація	1	2	3	4	5
Розмір існуючої ІТ-архітектури занадто складний для невеликих, але частих релізів	1	2	3	4	5
Залежність від інших проектів компанії	1	2	3	4	5
Відсутність кваліфікованої крос-функціональної команди	1	2	3	4	5
Низька залученість кінцевих користувачів	1	2	3	4	5
Відсутність необхідного програмного забезпечення для управління проектом за допомогою гнучкої методології	1	2	3	4	5
Інше:	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5
	1	2	3	4	5

19. Якою ще інформацією з тематики управління проектами Big Data за допомогою методик гнучких методологій Ви б хотіли поділитися з автором дослідження?

---



---



---



---

## Результати першої регресійної моделі

## 1. Результати першої ітерації

OLS Regression Results						
Dep. Variable:	success1	R-squared:	0.680			
Model:	OLS	Adj. R-squared:	0.558			
Method:	Least Squares	F-statistic:	5.574			
Date:	Sat, 14	Prob (F-statistic):	0.000737			
Time:	14:47:46	Log-Likelihood:	-88.154			
No. Observations:	30	AIC:	194.3			
Df Residuals:	21	BIC:	206.9			
Df Model:	8					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	3.2149	4.165	0.772	0.449	-5.446	11.876
backlog	2.2647	4.161	0.544	0.592	-6.388	10.917
us	-4.7914	3.678	-1.303	0.207	-12.440	2.858
sprint	3.2370	3.474	0.932	0.362	-3.988	10.462
standup	6.8871	3.046	2.261	0.034	0.553	13.221
retrosp	-4.8146	2.740	-1.757	0.093	-10.512	0.883
kanban	6.1079	2.345	2.605	0.017	1.231	10.985
openness	3.7564	1.173	3.203	0.004	1.318	6.195
expertise	1.1142	0.916	1.216	0.237	-0.791	3.019

## 2. Результати другої ітерації

OLS Regression Results						
Dep. Variable:	success1	R-squared:	0.579			
Model:	OLS	Adj. R-squared:	0.530			
Method:	Least Squares	F-statistic:	11.92			
Date:	Sat, 14	Prob (F-statistic):	4.25e-05			
Time:	14:50:12	Log-Likelihood:	-92.261			
No. Observations:	30	AIC:	192.5			
Df Residuals:	26	BIC:	198.1			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	2.7811	4.245	0.655	0.518	-5.944	11.506
standup	6.5139	2.595	2.510	0.019	1.180	11.848
kanban	6.3303	2.183	2.900	0.007	1.844	10.817
openness	4.1996	0.890	4.718	0.000	2.370	6.029

### 3. Результати статистичних тестів

- VIF

VIF Factor	features
0	4.3 standup
1	2.2 kanban
2	4.5 openness

- Жака-Бера

```
[('Jarque-Bera', 0.5337516239450696),
 ('Chi^2 two-tail prob.', 0.7657681647984069),
```

- Бройша-Пагана

```
[('Lagrange multiplier statistic', 4.240416307878593),
 ('p-value', 0.2366464691188857),
```

## Код програмування на мові Python

```
import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as seabornInstance

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

from sklearn import metrics

% Matplotlib inline

from sklearn import datasets, linear_model

from sklearn.linear_model import LinearRegression

import statsmodels.api as sm

from scipy import stats

from statsmodels.stats.outliers_influence import variance_inflation_factor

df = pd.read_excel ('disser.xlsx')

X = df [['backlog', 'us', 'sprint', 'standup', 'retrosp', 'kanban', 'openness', 'expertise']]
```

```

Y = df [ "success1"]

X3 = df [[ 'backlog', 'us', 'sprint', 'standup', 'retrosp', 'kanban', 'openness', 'expertise']]

X2 = sm.add_constant (X3)

est = sm.OLS (Y, X2)

est2 = est.fit ()

print (est2.summary ())

X3 = df [[ 'standup', 'kanban', 'openness']]

X2 = sm.add_constant (X3)

est = sm.OLS (Y, X2)

est2 = est.fit ()

print (est2.summary ())

vif = pd.DataFrame ()

vif [ "VIF Factor"] = [variance_inflation_factor (X3.values, i) for i in range (X3.shape [1])]

vif [ "features"] = X3.columns

vif.round (1)

import statsmodels.stats.api as sms

from statsmodels.compat import lzip

name = [ 'Jarque-Bera', 'Chi ^ 2 two-tail prob.', 'Skew', 'Kurtosis']

test = sms.jarque_bera (est2.resid)

lzip (name, test)

```

```

name = [ 'Lagrange multiplier statistic', 'p-value',
        'F-value', 'f p-value']

test = sms.het_breuschpagan (est2.resid, est2.model.exog)

lzip (name, test)

«Для показника проекту з вагами»

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as seabornInstance

from sklearn.model_selection import train_test_split

from sklearn.linear_model import LinearRegression

from sklearn import metrics

% Matplotlib inline

from sklearn import datasets, linear_model

from sklearn.linear_model import LinearRegression

import statsmodels.api as sm

from scipy import stats

from statsmodels.stats.outliers_influence import variance_inflation_factor

```

```

df = pd.read_excel ('disser.xlsx')
X = df [['backlog', 'us', 'sprint', 'standup', 'retrosp', 'kanban', 'openness', 'expertise']]
Y = df ["success2"]
X3 = df [['backlog', 'us', 'sprint', 'standup', 'retrosp', 'kanban', 'openness', 'expertise']]
X2 = sm.add_constant (X3)
est = sm.OLS (Y, X2)
est2 = est.fit ()
print (est2.summary ())
X3 = df [['standup', 'kanban', 'openness']]
X2 = sm.add_constant (X3)
est = sm.OLS (Y, X2)
est2 = est.fit ()
print (est2.summary ())
vif = pd.DataFrame ()
vif ["VIF Factor"] = [variance_inflation_factor (X3.values, i) for i in range (X3.shape [1])]
vif ["features"] = X3.columns
vif.round (1)
import statsmodels.stats.api as sms
from statsmodels.compat import lzip

name = ['Jarque-Bera', 'Chi ^ 2 two-tail prob.', 'Skew', 'Kurtosis']

```

```
test = sms.jarque_bera (est2.resid)
```

```
lzip (name, test)
```

```
name = [ 'Lagrange multiplier statistic', 'p-value',
```

```
'F-value', 'f p-value']
```

```
test = sms.het_breuschpagan (est2.resid, est2.model.exog)
```

```
lzip (name, test)
```