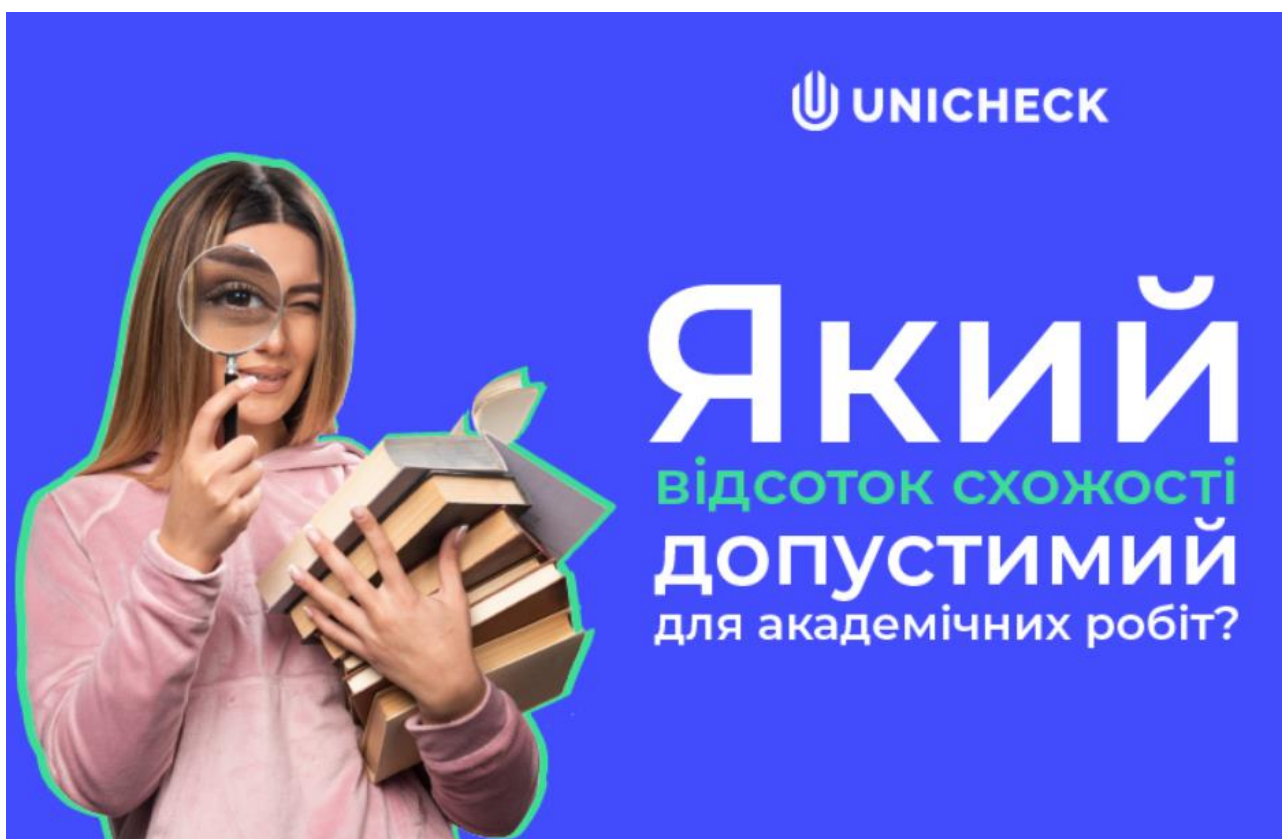


Який відсоток схожості можна вважати допустимим для академічних робіт?

<https://unichек.com/ua/blog/yakij-vidsotok-shozhosti-mozhna-vvazhaty-dopustymym-dlya-akademichnyh-robot>



Що насправді означає відсоток схожості та скільки текстових збігів дозволяється для академічних робіт?

Ясність. Це те, що дає відчуття розуміння. Це те, що вказує напрямок руху. Якщо ви користуєтесь інструментами для запобігання плагіату, скоріш за все, ви також цього прагнете. Знаючи, який відсоток подібності можна

вважати плагіатом, було б набагато простіше визначати долю наукової роботи: швидше знайти її слабкі місця та обговорити зі студентом допущені помилки.

Чи існує якась певна кількість відсотків схожості, яка може вважатися занадто великою для академічної роботи? Чим плагіат відрізняється від подібності? І який прийнятний відсоток схожості може бути у роботі? Ось деякі з наших думок і висновків.

Магія обчислення відсотка схожості у звіті Unicheck



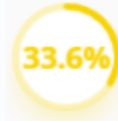
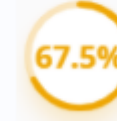

Що люди найбільше люблять в перевірці на схожість, так це те, що вони отримують відсоток подібності після кожної перевірки. Джерела, з якими порівнюється робота, можуть включати вебсторінки, книги, журнали, бази даних з відкритим доступом, приватні сховища, бібліотеки закладів освіти та багато іншого.

Щодо Unicheck, то його алгоритм проводить перевірки по веб індексу в реальному часі та по базі робіт навчального закладу. Кожне речення поділяється на менші складові. Потім проходять через пошукову систему. Якщо подібність встановлена, Unicheck перевірить джерело, щоб переконатися, що його варто додати до звіту.

Потім розраховується **загальний відсоток схожості** (ЗВС) за цією формулою:

$$\text{ЗВС} = \frac{\text{Загальна кількість слів у всіх текстових збігах}}{\text{Загальна кількість слів} - \text{Загальна кількість слів у посиланнях}} * 100$$

Колір відсотку буде залежати від “щільності” знайдених збігів і загальної кількості слів у документі:

0%	1-24%	25-49%	50-74%	75-100%
				

На перший погляд, червоний відсоток здається найсерйознішим ... Однак, у багатьох випадках, це може вводити в оману.

Червоний провокує миттєву реакцію

Часто викладачі та студенти сприймають показники схожості, як явні ознаки потенційного плагіату.

Аналітика Unichesk підтверджує той факт, що звіти з високим відсотком схожості перевіряються набагато ретельніше ніж інші.

Це й не дивно. Історично, червоний символізував тривогу. Тож, люди підсвідомо звертають більше уваги на нього. Інші кольори відсотків схожості не мають такого ефекту.

Однак, менші відсотки – не гарантія того, що текст повністю оригінальний.

На що саме вказує відсоток подібності у тексті?

Відсоток схожості вказує на кількість текстових збігів, знайдених після перевірки роботи..

Схожість може бути різною. Unichesk розрізняє:

- **Збіги** – текст, знайдений у базі академічних робіт установи або інтернеті
- **Цитати** – правильно відформатовані прямі/непрямі цитати
- **Список використаних джерел** – процитований перелік робіт в кінці документа

Таким чином, збіги найчастіше свідчать про запозичення в тексті та впливають на відсоток схожості. Цитати та список використаних джерел, при правильному форматуванні, не впливатимуть на відсоток, оскільки Unichesk автоматично виключає їх з результатів пошуку.

Проте, кінцеві відсотки у звіті не повинні одразу ж конвертуватися в оцінки. Ось чому.

Що може вплинути на відсоток подібності

Завжди є дві сторони однієї монети. Нижче ми перерахували декілька прикладів високих і низьких відсотків:

Чи вказує високий відсоток схожості на те, що в тексті є плагіат?

Високий відсоток інформує педагогів про проблеми, з якими стикаються студенти. Проте, відсоток схожості може збільшитися через відключені фільтри пошуку, неправильне форматування цитат і посилань або дублювання шаблонів завдань.

Розглянемо кожен випадок окремо:

- Фільтр, що виключає цитати та список джерел з результатів пошуку, вимкнений

Якщо цитати та список використаних джерел не виключені з результатів пошуку, вони будуть збільшувати загальний показник подібності.

Переконайтеся, що ця функція включена в налаштуваннях сервісу. Це позбавить вас від проблем, пов'язаних з видаленням їх вручну.

- Нова робота дублює попередню, написану одним і тим же студентом. У такому випадку, ви отримаєте найвищий відсоток схожості. Загалом, результати попередніх перевірок студентів повинні автоматично видалятися з результатів пошуку. Однак, якщо ваша інтеграція не підтримує такої функції, ви можете видалити попередню роботу вручну. У такий спосіб ви отримаєте той відсоток схожості, який є тільки в цій роботі. Вам залишиться перевірити цитати та список джерел.

- Цитати виділені як запозичений текст

Якщо студенти не слідують правилам цитування, Unicheck не зможе ідентифікувати цитату і буде вважати її збігом. У такому випадку слід виключити вручну, щоб знизити відсоток схожості.

- Типові шаблони/інструкції до завдань збільшують індекс подібності. Шаблони або інструкції до завдань часто згадуються в роботах студентів. Це неминуче призвело б до масового дублювання і збільшення відсотка схожості. Функція "Ігнорувати текст" дозволяє пропускати такі абзаци. Як тільки ви її застосуєте, Unicheck оновить кількість слів і не буде вважати шаблони текстовими збігами.

- Збіги, які виявилися загальними фразами/загальновизнаними фактами. Часто, в список текстових збігів включаються термінологія, описи, загальні знання. Багато студентів використовують загальну термінологію у текстах, тому ці частини тексту будуть збігатися, в результаті чого збільшується ЗВС. Очевидно, що ці збіги не мають нічого спільного з запозиченим текстом і повинні бути виключені для всіх документів.

Чи може 0-20% відсотків схожості бути прийнятним для всіх академічних робіт?

В Unicheck ми рекомендуємо проводити подвійну перевірку звітів навіть у тому випадку, якщо схожість низька. Ситуації бувають різними. Ось чому сервіси для перевірки на плагіат не дають ніяких інструкцій з інтерпретації відсотка схожості. Ось ще два випадки, які підтверджують нашу точку зору:

- Крихітний відсоток може свідчити про серйозну проблему

Іноді, перевірка мікро збігів теж варта зусиль. Unicheck може виявити невеликий текстовий збіг з джерелом, яке було перефразовано. Включіть режим порівняння у звіті Unicheck для того, щоб порівняти текст роботи студента та джерела. Це полегшить процес порівняння роботи та її аналіз.

- Робота може мати низький відсоток подібності через текстові зміни

Уявіть, що дві роботи мають однаковий відсоток схожості, нижче 10%. Чи варто перевіряти кожен? Безперечно. Одна може мати декілька скопійованих фраз або неправильно процитованих слів. Інша може містити текстові маніпуляції, які мінімізують реальний показник схожості. Якщо такі маніпуляції будуть виявлені, сервіс повідомить про підозрілі зміни в тексті та запропонує відкрити вкладку “Модифікація” для більш детального аналізу.

Відсоток плагіату, дозволений в університетах

Простіше кажучи, сервіс з перевірки текстів робить в основному дві речі для педагога. Він обчислює кількість збігів і показує, звідки вони беруться. Unichesk – це лише інструмент. Він допомагає викладачам більш детально вивчати документ і приймати обґрунтовані рішення, а студентам – вказує на те, що потребує цитування.

Часто виникає ще одне питання: “Який відсоток схожості вважається плагіатом?”

Сергій Єльцов, директор навчального центру менеджменту якості та моніторингу освітнього процесу Харківського національного університету імені В. Н. Каразіна:

У нас є рекомендовані показники оригінальності, але вчені ради факультетів можуть в певних межах їх змінювати.

Рекомендовані показники оригінальності за величиною коефіцієнта подібності становлять:

- не більше 20% – оригінальна робота,
- від 21% до 50% – задовільно оригінальна робота,
- від 51% до 90% – умовно оригінальна робота,
- понад 90% – неоригінальна робота.

Подальший розгляд роботи залежить від її класифікації за цим критерієм, наприклад, умовно оригінальна робота може бути за рішенням кафедри направлена на доопрацювання, а неоригінальна – відхилена від захисту. Якщо робота була на доопрацюванні, її знов перевіряють, якщо рівень запозичень не зменшився, то робота також відхиляється від захисту.”

Прийнятний відсоток подібності може відрізнитися від одного університету до іншого. Для технічних наук високі відсотки схожості можуть бути цілком нормальними через безліч використовуваних формул/шаблонів. На відміну від цього, високі процентні показники можуть свідчити про неоригінальність для філології та історії. Рішення про те, що дозволено і наскільки гнучкими можуть бути допустимі відсотки схожості для курсу або завдання, приймає установа.

І ще декілька корисних порад:

Для того, щоб викладачі й студенти розуміли один одного й отримували максимальну користь зі звіту подібності Unicheck, ми склали цей перелік порад:

- Роз'ясніть студентам, що збіги з іншими роботами допустимі і якими вони можуть бути
- Перевірте налаштування сервісу та включіть фільтри пошуку
- Переглядайте всі збіги в роботі та виключайте ті, що неправильно цитуються або повторюють формули/загальні фрази
- Порівняйте кількість цитат з об'ємом оригінального тексту студента роботи
- Прийміть остаточне рішення щодо оцінки роботи