



**Ігор КРАСОВСЬКИЙ**,  
менеджер  
відділу  
корпоративних  
продуктів  
компанії АBBYY  
Україна

## FlexiCapture: як з ним працювати

Складна всередині програма FlexiCapture, про яку ми розповідали у попередній статті, після налаштування доступна для роботи навіть для «чайників». У цій статті ми навчимо, як вводити стоти паперових документів у систему, перевіряти їх та зберігати у потрібному форматі у потрібному місці. Програма впроваджується із зображенням тексту та навіть написаним від руки

Процес автоматизованої обробки стосів документів з допомогою програми FlexiCapture нескладний і складається з чотирьох етапів: власне сканування, розпізнавання відсканованого тексту, перевірки та експорту даних. Розгляньмо їх детальніше.

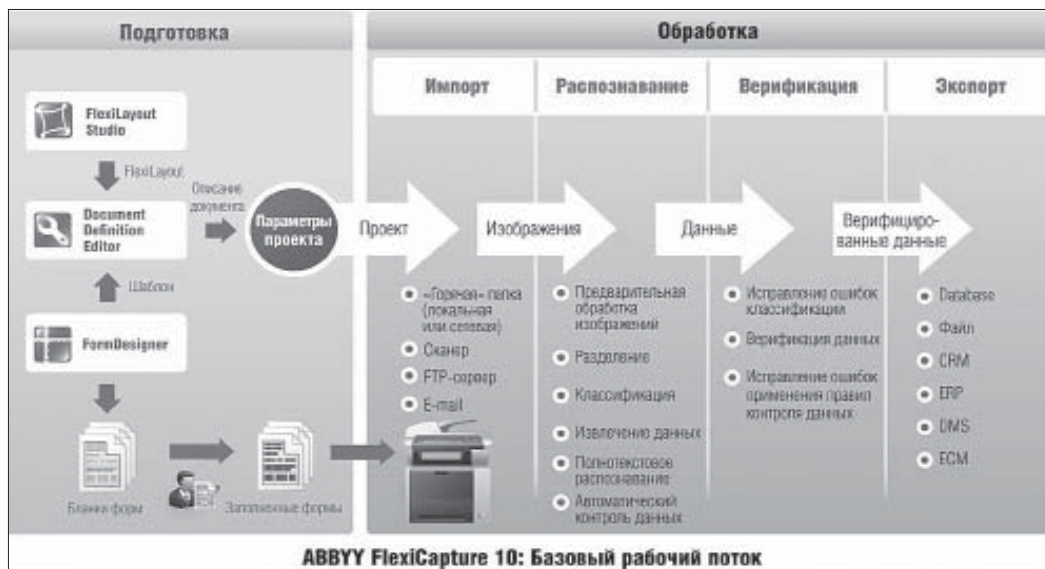


Рис. 1. Етапи обробки документів у програмі FlexiCapture

### Перший етап — сканування

Щоб розпочати роботу з АBBYY FlexiCapture, закладіть у сканер документи, які необхідно опрацювати. Краще використовувати сканери з автоподавачем, адже в таку машину можна закласти як один аркуш, так і цілу пачку документів (зазвичай близько 50 документів).

Після того як оператор запустить процес сканування, система АBBYY FlexiCapture автоматично відсканує документи. Якщо використовувати сучасні потужні сканери, то на ці 50 аркушів піде 1–3 хвилини. Кожна пачка сформує окремий пакет, якому можна присвоїти унікальні параметри, щоб полегшити подальший пошук даних.

**Ключові слова:**  
автоматизація  
обліку, програмне  
забезпечення,  
HMA, 1С

## Другий етап — розпізнавання й автоматична перевірка

Після сканування документів програма розпізнає текст, автоматично знайде заповнені поля у формі та перевірить правильність даних за допомогою зовнішніх та внутрішніх словників.

Автоматичне визначення типу тексту дозволить обробляти форми, роздруковані на різних типах принтерів, друкованій машинці, типографічним способом та вручну друкованими літерами. Також обробляються зображення тексту з графічних файлів (формати TIFF, JPG, PCX, DCX, PNG, BMP).

Примітно, що можна задати певні обмеження для документів. Приміром, якщо про дату відомо, що вона має потрапляти в певний часовий інтервал, то після розпізнавання дані перевіряються на виконання цієї умови, і, якщо вона не виконується, оператор отримує повідомлення про помилку.

Більше того, система може самостійно перевіряти інформацію за словником, базою даних, сумою кількох полів, спеціальними правилами. Налаштування цих правил досить гнучке, тож адміністратор системи може легко їх виправляти та доповнювати.

## Третій етап — перевірка даних оператором

Оскільки точність розпізнавання завжди трохи нижча за 100 відсотків, після програми інформацію має перевірити оператор. Для цього в системі є три способи перевірки:

- групова;
- контекстна;
- перевірка за формою.

**Групова перевірка** ідеально підходить для перевірки даних, які належать до відносно невеликої множини (цифр чи міток). Під час групової перевірки всі невпевнено розпізнані символи одного виду (скажімо, нулі) виводяться на екран перед оператором (рис. 2).



Рис. 2. Перевірка цифр під час групової верифікації

Завдяки особливостям людського сприйняття оператор виділяє один нетиповий символ з великої кількості однотипних і таким чином підтверджує правильність сотні однотипних символів простим натисканням кнопки Enter.

У ході **контекстної перевірки** на екран одночасно виводяться фрагмент вихідного зображення та результат розпізнавання (рис. 3). Тож оператор може порівняти результати розпізнавання з оригіналом та підтвердити правильність натисканням однієї клавіші або виправити неправильно розпізнані символи.

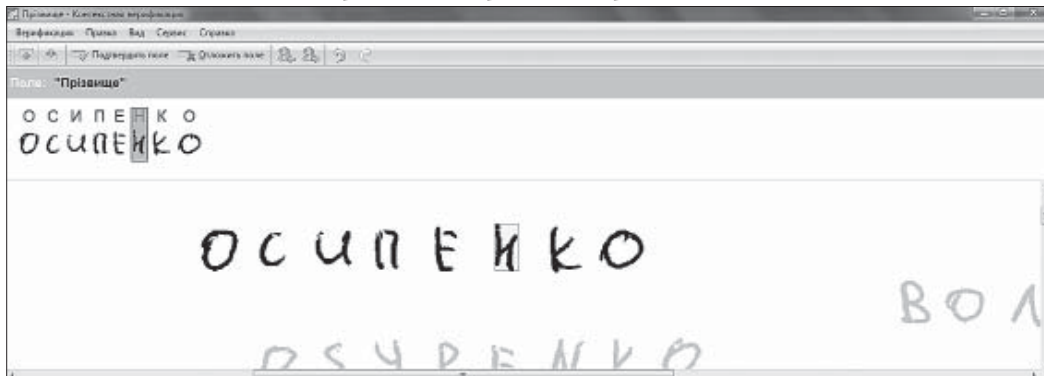


Рис. 3. Контекстна перевірка невпевнено розпізнаних символів

**Перевірка за формою** проводиться, якщо контрольні перевірки свідчать про наявність серйозних помилок у документах. Документ помічається «прапорцем», і потім зображення документа виводиться на екран, щоб оператор передивився помилкові поля та вніс зміни (рис. 4).

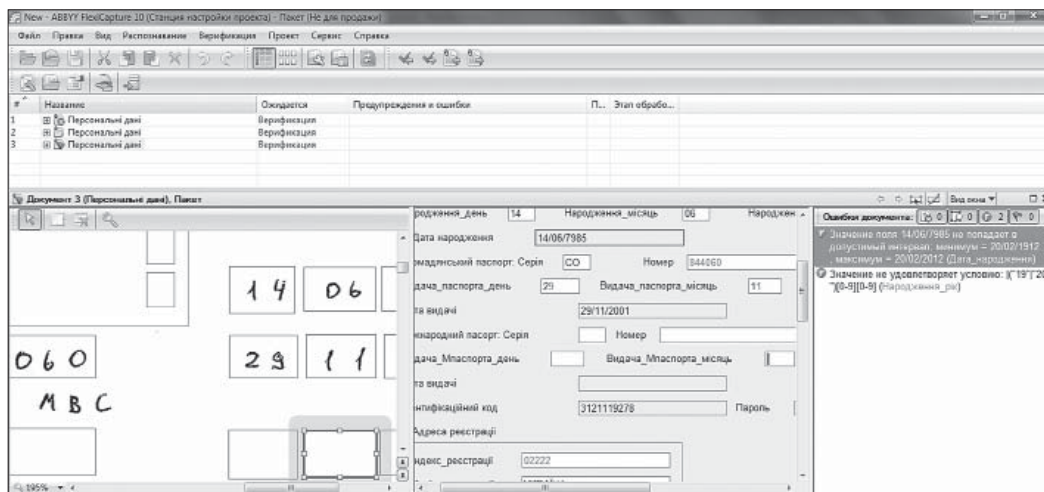


Рис. 4. Перевірка даних у формі

#### Четвертий етап — експорт

Перевіривши документ, можна експортувати його у файл потрібного формату (Excel, TXT, CSV, DBF) або одразу у бази даних чи будь-яку інформаційну систему для подальшого використання.

Тож тепер немає необхідності витрачати час на передрукування потрібних даних (друкованих, сфотографованих чи написаних від руки) зі стосів документів, а потім перевіряти правильність введення. Достатньо довірити механічну роботу розумній програмі, долучаючись до процесу лише для перевірки невпевнено розпізнаних слів та символів.

