Osiągnięcia technologiczne prowadzące do udoskonalenia technik analitycznych pozwoliły na rozwój analitycznych technik łączonych/sprzężonych. Łączą one głównie metody chromatograficzne oraz spektroskopowe wykorzystując jednocześnie potencjał obu technik. Techniki łączone, takie jak chromatografia gazowa sprzężona ze spektrometrem mas (GC-MS), dostarczają ogromnej ilości informacji, stanowiąc jedne z najbardziej przydatnych narzędzi do analizy próbek o złożonych matrycach wymagających wstępnego rozdzielenia chromatograficznego, a następnie identyfikacji składników na podstawie stosunku masy do ładunku jonów powstałych w wyniku fragmentacji związków. Nowe rozwiązania udoskonalające techniki łączone na przestrzeni ostatnich lat znacząco zwiększyły pole ich zastosowań w badaniach dla celów sądowych, medycynie, analizie środowiskowej i żywności ze względu na dostarczanie jak najbardziej kompletnej informacji o analizowanych próbkach. W naukach sądowych techniki łączone wykorzystywane są w celu porównania dwóch próbek, np. włókien, by ustalić, czy mogą one pochodzić z tego samego źródła (hipoteza H_1), czy dwóch różnych źródeł (H_2) i w ten sposób powiązać ewentualnego podejrzanego z miejscem zdarzenia w toku dalszego postępowania. Innym przykładem ich aplikacji jest analiza autentyczności produktów żywnościowych lub innych cennych dóbr, której istotą jest ustalenie, czy analizowany produkt jest oryginalny (H_1) czy podrobiony (H_2).

W wyniku łączenia dwóch technik analitycznych, techniki sprzężone generują ogromną ilość danych. Często wizualizuje się je w formie dwuwymiarowej mapy/obrazu, na której kolor każdego piksela odpowiada intensywności zmierzonego sygnału. Ich podobieństwo jest często ocenione z wykorzystaniem testów statystycznych i metod rozpoznawania obrazów. Należy mieć jednak na uwadze, że kluczem w ocenie podobieństwa próbek w naukach sądowych, czy podczas weryfikacji autentyczności produktów jest fakt, że gdy cechy obu próbek są bardzo popularne, ich podobieństwo może być przypadkowe. Niebezpieczeństwo to rośnie wraz ze wzrostem częstotliwości danej cechy. Z tego powodu teoria testowania hipotez oparta na ilorazie wiarygodności (ang. likelihood ratio, LR) jest rekomendowanym sposobem pozwalającym na interpretację podobieństwa danych fizykochemicznych w świetle dwóch przeciwstawnych hipotez (H_1 i H_2), uwzględniając unikatowość i zmienność cech w obrębie próbek i pomiędzy nimi oraz korelacje między cechami. Ponadto, w przeciwieństwie do klasycznych testów statystycznych, wartości LR bezpośrednio wskazują, jak duże jest zaobserwowane podobieństwo, a więc siłę wsparcia dla H_1 i H_2.

Klasyczne modele LR pozwalają na porównanie cech próbek na podstawie baz danych, w których liczba próbek (m) jest większa niż liczba cech je opisujących (p). Ponadto, rozrzut danych w obrębie każdej próbki powinien być znacznie niższy od rozrzutu średnich dla tych próbek, aby umożliwić ich efektywne odróżnienie. Modele te przestały być wystarczające ze względu na szybki rozwój metod analitycznych, który wymusił ich ewolucję. Pierwsze udoskonalenia dotyczyły analizy podobieństwa widm lub chromatogramów, dla których $m<<p$, z zastosowaniem hybrydowych modeli LR. Wykorzystują one chemometryczne metody redukcji wymiarowości przestrzeni cech, by zapewnić $m>p$, i następnie wkomponowują ich rezultaty w modele LR. Pojawia się więc pytanie, czy hybrydowe modele LR mogą zostać również zastosowane do analizy podobieństwa danych, jak np. dwuwymiarowe sygnały GC-MS, po ich uprzednim zmodyfikowaniu i dostosowaniu. Uzyskanie odpowiedzi na to pytanie jest więc motywacją podjęcia niniejszych badań. W związku z rosnącą kompleksowością tego typu danych ich interpretacja wiąże się z pojawieniem się nowych problemów, których nie obserwowano dla danych o mniejszej wymiarowości, np. widm. Utrudnienia mogą więc pojawić sie podczas wyboru zmiennych, czy ich wizualizacji, która niejednokrotnie dostarcza cennych wskazówek. Celem badań jest więc wykorzystanie potencjału algorytmów do wnikliwej eksploracji danych (data mining), rozpoznawania obrazów (pattern recognition) oraz uczenia maszynowego (machine learning), które wspomogą proces poszukiwania ukrytych trendów w danych, redukcji wymiarowości przestrzeni cech poprzez wydobycie najbardziej unikatowych cech opisujących próbki i utworzenie efektywnego modelu LR do oceny podobieństwa np. danych GC-MS.

Badania te będą krokiem milowym w dziedzinie porównywania danych uzyskanych z technik łączonych między innymi w naukach sądowych, medycynie, analizie autentyczności produktów żywnościowych i innych cennych dóbr, gdzie o podobieństwie decyduje nie tylko zgodność cech, ale także ich unikatowość, wzmacniająca wyciągane wnioski. Idea utworzenia modeli LR do oceny podobieństwa danych z technik łączonych w kontekście dwóch hipotez z uwzględnieniem ich unikatowości jest innowacyjna i pionierska, gdyż prowadzi do wyciągania bardziej trafnych i przekonujących wniosków o podobieństwie próbek.

Dzięki zastosowaniu zaawansowanych technik analitycznych, chemometrycznych i statystycznych badania mogą pozwolić na poszerzenie wiedzy o czynnikach mających kluczowe znaczenie dla analizy porównawczej dwuwymiarowych danych. Będą one stanowiły istotny wkład w ich ewolucję, szczególnie w odniesieniu do redukcji wymiarowości przestrzeni cech oraz poszukiwania cech unikatowych w celu jak najlepszego rozróżnienia próbek i podniesienia wartości wyciąganych wniosków. Ogromne znaczenie projektu wiąże się z faktem podejmowania nowych wyzwań, innych niż w przypadku danych o mniejszej wymiarowości, które można analizować z wykorzystaniem zdecydowanie większej plejady metod, szczególnie do selekcji zmiennych, czy analizy źródeł zmienności i korelacji.