The recent advancement in both the computational technology and the analytical equipment has lead to the development of the hyphenated (coupled) analytical techniques that mostly combine chromatographic and spectroscopic methods coupled on-line to exploit the advantages of both. Hyphenated chromatographic systems, such as gas chromatography coupled with mass spectrometer (GC-MS) can provide rich analytical information and are some of the most reliable and suitable methods for analyzing complex samples that firstly require efficient separation of the components in a mixture using gas chromatography and detailed information about their structure by breaking them down into ionized fragments and detecting via their mass-to-charge ratio afterwards. The improvements in hyphenated methods over the last decades have remarkably broadened their applications in the forensic studies, medicine and environmental and food analysis, which arise from the virtue of providing possibly the most complete information about the examined samples. In forensics they are used for comparing samples e.g. to establish whether the analysed small fragments of fiber, may be two pieces of the same object (hypothesis H_1) or not (H_2), to find any links between person and the crime place in further investigation. Another application involves food or other valuable goods authenticity verification, which attempts to indicate whether the product is original (H_1) or has been falsified (H_2), mainly by intentional mislabelling.

Hyphenated techniques generate massive amount of data as a consequence of coupling two techniques. They are often visualised in the form of a 2D image or a map, with each pixel coloured according to the signal intensity. Their similarity may be easily investigated using e.g. hypotheses testing coupled with pattern recognition tools. However, it should be borne in mind that the key to inspect the similarity of the samples in forensics or product authentication studies is the possible coincidental similarity of the samples, which risk grows with increasing frequency of the data. Then the likelihood ratio (LR) is the method of choice, which provides a way for expressing the similarity of the compared data by accounting their similarities and discrepancies in view of two contrasting hypotheses (H_1 and H_2) and includes the frequency of observing the characteristics, their average variability within the samples and between them and correlations. Moreover, unlike typical statistical tests, LR values directly show the magnitude of the samples similarity and thus the strength of the support for H_1 or H_2.

Classical LR models are suitable for comparing samples using the database containing more samples (m) than the number of variables (p). Also, the variance of the data within each sample should be much lower than the variance of samples averages to facilitate their differentiation. Since the time when these models were sufficient, the rapid development of analytical tools has also forced the evolution of the LR models. Their primary refinement attempted to deal with the *curse of dimensionality* of highly multidimensional data such as spectra of chromatograms for which $m << p$, with the application of the hybrid LR models. They use the chemometric data dimensionality reduction techniques to keep $m > p$ and incorporate their outcomes in LR framework afterwards. We hypothesize that the hybrid LR models are also applicable on more complex data, such as 2D GC-MS images but after severe customisation to succeed in commenting on their (un)common origins, which is the motivation of the research. We presume that due to growing complexity of these data their interpretation entails some new problems, that were not for less complex data, such as spectra, including the impediments in feature selection and visualisation of data that is valuable for their proper inspection. Our objective is to use more demanding algorithms for data mining, pattern recognition and machine learning that will help uncover hidden patterns/trends in the data, reduce their huge dimensionality by maintaining the most unique chemical information to make samples easily distinguishable, and compose a well performing hybrid LR model to assess the similarity of e.g. 2D GC-MS images based on the outcomes of this method.

We believe that the studies become a milestone in comparing the 2D images in forensics, food or other valuable goods authentication, medicine and other fields, which do not only concentrate on the similarity of the data but their specificity is also vital and reinforces the strength of the drawn conclusions. This picture is still incomplete and the rapid development of hyphenated analytical tools stimulates also the evolution of the LR models in this direction. We believe that the idea of developing the LR models for interpreting the similarity of 2D images in the context of two hypotheses is innovative and pioneering in that it would significantly enhance the credibility of conclusions, which become more persuasive. By utilising advanced analytical techniques and sophisticated chemometric, statistical resources the project is believed to improve knowledge of the factors that mostly influence the similarity of 2D images. It will remarkably contribute to the LR framework evolution, specifically in the aspects of compressing dimensionality and maintaining the most unique features to facilitate differentiation of the samples and value of the conclusions. The significance of this research relates to going far beyond the problems found for LR models on two-way data, for which many methods are available to feature selection, studies on variance components and correlations.