The ALINA project is devoted to the attempt to apply artificial intelligence to develop algorithms for post-processing the results of indirect examinations of historical manuscripts by means of advanced spectral methods. Spectral methods will be used to evaluate the possibility of distinguishing inks on the basis of their elemental composition encoded in special indicator papers.

Knowledge of source materials based on authentic historical texts is crucial for understanding our culture. Archives, libraries, and museums keep valuable manuscripts containing inks, almost all of which were created by mixing soluble iron compounds and plant tannins. A rich source of tannins was often gall nuts, the pathological growths formed on oak leaves after the gall wasp laid its eggs. Iron gall inks dominated writing materials for nearly 900 years for practical reasons, including simplicity of preparation. The diversity of the chemical composition was a result of the freedom of mixing, often accidental, components combined together in unobvious proportions. Such irreproducible randomness could result in chemical instability of inks and then lead to degradation of paper or parchment support. Acid hydrolysis of cellulose and reactions catalysed by free iron ions may cause ink corrosion, irreversibly destroying manuscripts and drawings of high historical value. Having developed an algorithm, we are planning to carry out a detailed study of the documents of the 3rd of May Constitution preserved in the collection of the Central Archives of Historical Records in Warsaw.

In spite of many years of research, many questions still remain concerning the factors essential for accurate diagnosis of the state of preservation and assessment of threats resulting from the presence of inks. Understanding the mechanisms involved in ink corrosion may be of key importance for developing effective conservation strategies for historical objects. Examination of the most valuable documents must be limited to use of methods that are completely safe for unique objects. A simple method for assessing the risk of ink corrosion to manuscripts has been developed by the Dutch scientists and conservators by introducing indicator papers soaked in a substance that takes on colour when it comes into contact with iron ions in unstable inks. This allows the conservator to partially assess whether a particular ink will cause parchment or paper to deteriorate. Partially, because only the iron ions react colourfully and the indicator does not detect other elements (copper, zinc, etc.), which are also part of the inks and can migrate with the iron ions to the indicator paper. Thus, the indicator paper can indirectly represent the source object during laboratory analysis, during which detailed data will be discovered by mass spectrometry measurements. These allow for determination of trace element/isotope contents, while direct examination of the surface of the sample can provide up to several thousand partial data about local composition. It is of high importance to correctly calculate the recorded numerical results and provide conservators with accurate information describing the composition of manuscripts and iron-gall inks.

The task would be trivial if it were not for the aforementioned unique composition of almost every historical ink. The variation in the proportions that occur between elements nonlinearly changes the strength of their binding by the indicator used. This variation is not accidental, it results from regularities governing chemical reactions occurring during contact of the indicator with inks. So far, the ways to decode the true information have been presented with huge approximations. The ALINA project aims to break the previous uncertainty limit resulting from the simplified post-processing of results recorded in the laboratory for indicator papers. We want to combine advanced methods of modern instrumental analysis with modern methods of data analysis used in artificial intelligence - the so-called Deep Learning methods. These methods have allowed in recent years to achieve spectacular results in areas as diverse as diagnostic medical imaging, synthesis of new antibiotics, or identification of whales on the basis of undersea recordings of their voices. The power of Deep Learning methods lies in their ability to detect complex, nonlinear and multivariate relationships in large Big Data databases, such as the large sets of results obtained in laboratories describing chemical processes in centuries-old inks.

The project will culminate with the testing of the refined methodology during examinations of the original manuscripts of the 3rd of May Constitution with the signatures of the officials and members of the Great Diet who signed the documents. We hope that ALINA will make it possible not only to assess the actual state of preservation of these important manuscripts, but also to distinguish the inks, which will then be able to be linked to the reconstruction of the history of changes made to the text of the Constitution just before its adoption.