**УДК 930.2:94(3)**

***Кузнецов Алексей Валерьевич***

*Кандидат исторических наук,*

*научный сотрудник*

*Института всеобщей истории РАН,*

*г. Москва, Российская Федерация*

*e-mail:*

**Компьютерный анализ текстов на латинском языке: Латентно-семантический анализ «Истории готов, вандалов и свевов» Исидора Севильского**

**The computer analysis of Latin texts: Latent Semantic Analysis of «Historia de regibus Gothorum, Wandalorum et Suevorum» by Isidore of Seville**

Аннотация

В статье предпринимается попытка с использованием современных методов интеллектуального анализа текстов исследовать латиноязычный текст хроники «История готов, вандалов и свевов» выдающегося богослова и ученого VII века Исидора Севильского. Ставится цель проверить выдвинутую в историографии гипотезу о наличии у автора представлений об определенной иерархии варварских народов. Основное внимание направлено на раскрытие неявных семантических взаимоотношений между различными частями произведения с целью уточнения отношения автора к трем варварским народам.

Анализ текста проводился на языке программирования R. В качестве конкретного метода бы выбран метод латентно-семантического анализа, позволяющий проводить сравнение и кластеризацию текстов на основе семантического пространства, построенного путем сингулярного разложения терм-документной матрицы.

Актуальность исследования заключается в том, что впервые реализован полный цикл латентно-семантического анализа средневекового латиноязычного текста. Проведена предварительная подготовка, построено семантическое пространство текста памятника, осуществлено сравнение семантической схожести текстов на основе меры косинусного сходства. Результаты анализа позволяют утверждать, что Исидор Севильский действительно выстраивает иерархию из трех варварских народов, придавая большее сходство описанию вестготов и свевов и особняком ставя вандалов.

The article makes an attempt to study the Latin text of the chronicle «Historia de regibus Gothorum, Wandalorum et Suevorum» by the theologian and scientist of the 7th century Isidore of Seville by means of the methods of text mining. The main goal is to verify the hypothesis that the author has ideas about a hierarchy of barbarian kingdoms. The main focus is on clarifying of implicit semantic relationships between the different parts of the chronicle in order to find out the author’s attitude to the three barbaric kingdoms.

The analysis of the text was performed with the R programming language. As a method for analysis was chosen the method of latent-semantic analysis. It allows the computation of semantic similarity and clustering of texts, which are represented as vectors in the semantic space constructed by the singular decomposition of the term-document matrix.

The relevance of the study is that for the first time provides a full cycle of latent semantic analysis of medieval Latin text: text preprocessing, creation of semantic space, calculating of the semantic similarity of the texts on the basis of the cosine similarity measure. The results of the analysis suggest that Isidore of Seville really builds a hierarchy of three barbarian kingdoms, giving greater similarity to the description of the Visigoths and Suevs and putting Vandals apart.

**Ключевые слова**: Исидор Севильский, раннесредневековая историография, компьютерный анализ текстов, латентно-семантический анализ, сингулярное разложение, кластерный анализ.

**Keywords**: Isidore of Seville, early Middle Age historiography, computational text analysis, Latent Semantic Analysis, singular value decomposition, cluster analysis.