

## **МЕТОД ИЗВЛЕЧЕНИЯ ДАННЫХ ЦИФРОВОЙ ЛИЧНОСТИ ИЗ СОЦИАЛЬНЫХ СЕТЕЙ**

**Автор – С. А. Михайлов**

(Университет ИТМО, Санкт-Петербург)

**Научный руководитель – А. М. Кашевник**

(Университет ИТМО, Санкт-Петербург)

Цифровая личность (Digital Identity) — это информация об объекте, используемом компьютерными системами для представления внешнего агента, который может быть, как и человеком, так и организацией. Согласно стандарту ISO/IEC 24760-1 цифровую личность можно определить, как «набор атрибутов, которые относятся к сущности». Информация, содержащаяся в цифровой личности, позволяет оценивать и аутентифицировать пользователя, взаимодействующего с системами в Интернете, без участия операторов-людей. Использование концепции цифровой личности позволяет облегчить работу среди разнообразных сервисов-федераций, позволяя обмениваться информацией в рамках заранее разработанной модели.

Для описания цифровой личности можно применить онтологический подход. Онтологией является точной спецификацией предметной области в виде концептуальной схемы и состоит из классов, экземпляров, атрибутов и различного вида отношений. Данный подход позволяет создать семантическую модель данных, которой можно пользоваться среди разнообразных сервисов с целью обеспечения интероперабельности.

Целью данной работы является создания метода извлечения данных цифровой личности из различных открытых источников, в том числе и социальных сетей. Данный метод позволит получить большой набор данных, на основе которых можно проводить анализ и построение математических моделей. В основе данного метода предлагается использовать автоматическое сопоставление структуры данных из открытых источников и предложенной онтологии. Для извлечения сущностей среди представленных сырых данных используются так называемые графы знаний (knowledge graphs) – большие семантические сети, в которых описываются сущности и которым можно строить семантические запросы.