

**Аннотация научного доклада ведущего научного сотрудника
ИПМИ КарНЦ РАН к.т.н. А.А. Крижановского
«Раскрытие данных скрытого Веба на примере системы топонимии Карелии»**

Разрабатывается информационная онлайн-система «Топонимия Карелии» (ТопКар). При этом решается задача интегрирования ТопКар с поисковыми системами (Google, Яндекс), а именно, чтобы пользователи поисковых систем при вводе названия топонима могли получить в списке результатов страницы сайта ТопКар, содержащие искомую информацию. Трудность для данных поисковых систем заключается в том, что страницы сайта ТопКар являются динамическими, то есть после того, как пользователь вводит запрос система ТопКар обращается к базе данных и генерирует веб-страницу автоматически. Без применения специальных методов такие страницы не могут быть проиндексированы поисковыми системами, они являются для них «невидимыми». Таким образом, задача состоит в то, чтобы включить динамические страницы скрытого Веба в массив доступных и индексируемых поисковых ресурсов на примере геоинформационного ресурса «Топонимия Карелии» (ТопКар). Парадигма распределённых связанных данных (Linked data) изменилась при возникновении базы Викиданных, ключевым становится понятие единого уникального идентификатора, связывающего различные источники информации. Предложен подход включения системы ТопКар в общую сеть связанных данных за счёт нескольких факторов: (1) предоставление открытого онлайн-доступа к большому объёму топонимической информации, (2) связывание топонимических статей с идентификаторами Викиданных, (3) использование информации топонимических карточек (перевод, ударение, этимология, географические координаты, связь с другими топонимами) в статьях энциклопедии.

Сотрудники Института языка, литературы и истории КарНЦ РАН и сотрудники отдела ГИС ПетрГУ разработали в 2003-2010 гг. географическую информационно-аналитическую систему «Топонимия Карелии». Это была локальная система без доступа к сети Интернет. На основе данных этой системы начата разработка онлайн-системы ТопКар. Поскольку в системе ТопКар используются карты открытой географической системы OpenStreetMap, необходимо преобразовать локальные карты в электронные карты открытой географической системы OpenStreetMap. Для решения этой задачи был разработан автоматизированный метод для преобразования координат топонимов из старой локальной системы ТопКар в новую онлайн-версию.

Сейчас база данных ТопКар включает 1 377 топонимов с координатами, около 14 тыс. топонимов без координат, но с привязкой к поселениям на карте. Из 3310 поселений 1926 г. в ТопКар имеют координаты 75 поселений. Из 1287 современных поселений около половины (586 поселений) имеют координаты. Участники проекта продолжают работать над пополнением базы данных ТопКар.