

Информационная инфраструктура социально-гуманитарных наук в Евросоюзе

А.Б. Антопольский

г.н.с. ИНИОН РАН

Доклад на Ученом совете

12.05.2021 г. ale5695@yandex.ru

Управление наукой в Евросоюзе

https://ec.europa.eu/info/research-and-innovation_en

Стратегия в области исследований и инноваций

- Цели научно-инновационной политики
- Комиссар по науке и образованию
- Европейский инновационный совет (EIC)
- Европейский научный совет (ERC)
- Европейское исследовательское пространство (ERA)
- Европейская научная инфраструктура
- Европейское открытое научное облако (EOSC)

Европейская научная инфраструктура

https://ec.europa.eu/info/research-and-innovation_en

Научная инфраструктура – это организации, которые предоставляют исследовательским сообществам ресурсы и услуги для проведения исследований и стимулирования инноваций. Они включают в себя:

- научное оборудование или наборы инструментов
- коллекции, архивы или научные данные
- вычислительные системы и коммуникационные сети
- другие научные и инновационные средства открытые для внешних пользователей

Цели инфраструктуры

- уменьшение фрагментации исследовательской и инновационной экосистемы
- сокращение дублирования НИР
- координация разработки и использования научной инфраструктуры
- создание новых компонентов общеевропейской инфраструктуры
- объединение усилий для создания и управления крупными, сложными или дорогостоящими инфраструктурами,
- объединение навыков, и усилий лучших ученых мира
- результаты исследований - в открытый доступ
- повышения осведомленности промышленности о новых технологиях

Хартия доступа к инфраструктуре <https://op.europa.eu/en/publication-detail/-/publication/78e87306-48bc-11e6-9c64-01aa75ed71a1>

Инициативы, стратегии и сети научной инфраструктуры

- Европейский Стратегический форум по научной инфраструктуре (ESFRI)
- Группа старших должностных лиц
- EIROforum –соглашение о сотрудничестве
- Ассоциация европейской научной инфраструктуры (ERF-AISBL)
- Долгосрочная устойчивость научных инфраструктур
- Европейское открытое научное облако (EOSC)
- Европейские научные инфраструктурные консорциумы (ERIC) – всего 18, SSH -5, инфраструктуры не ERIC -3

Общие информационные системы ЕС

- Европейское открытое научное облако (EOSC)
 - OpenAIRE Европейская научная инфраструктура открытого доступа (The Open Access Infrastructure for Research in Europe)
- Издательское бюро ЕС
 - База данных исследовательских проектов (CORDIS)
 - Портал открытых данных ЕС (EU ODP)
 - Публикации структур ЕС
- Объединенный исследовательский центр ЕС (JRC)
 - Центры знаний и компетенций - 3 тыс. наборов данных - 156 коллекций

Научная инфраструктура по общественным и гуманитарным наукам

- Европейское социальное исследование (ESS ERIC);
- Европейская геронтологическая служба (SHARE-ERIC);
- Консорциум европейских архивов данных по социальным наукам (CESSDA ERIC);
- Европейская научная инфраструктура языковых ресурсов и технологий (CLARIN ERIC);
- Цифровая научная инфраструктура для искусств и гуманитарных наук (DARIAH ERIC);
- Открытая научная коммуникация в социальных и гуманитарных науках (OPERAS)
- Европейская исследовательская инфраструктура для наук о наследии (E-RIHIS)
- Сеть содействия языковым ресурсам и технологиям (FLaReNet)

Европейское социальное исследование (ESS ERIC)

- *Методология* : разработка показателей спецификация обследования, анкеты, порядок и режимы сбора, обработки и архивирования данных, оценка доступности и качество данных и документации
- *Данные и документация*: сведения по годам, странам и темам, инструменты онлайн-анализа, накопительные данные, методологические данные, многоуровневые данные, оповещения о данных ESS, условия использования данных

Европейская геронтологическая служба (SHARE ERIC)

- Общие данные - результаты обследований 140 тыс. чел. + 50 (380 тыс. интервью) из 28 стран
- Доступ к данным – бесплатно через 2 архива
- Специальные наборы данных
 - учебный набор данных
 - страновые проекты
 - опрос интервьюеров
 - биомаркеры
 - опрос Corona
 - голландские данные онлайн-опроса
- Публикации

Консорциум европейских архивов данных по социальным наукам (CESSDA ERIC)

Основные сервисы

- Европейский многоязычный социологический тезаурус (ELSST)
- Словарная служба (CVS)- 28 контролируемых словарей
- Каталог данных (CDC) – ИПС - 30 тыс. исследований
- Банк опросов (EQB)
- Руководство по управлению данными
- Депонирование данных для журналов
- Новые типы данных (социальные сети и проч.)

Европейская научная инфраструктура языковых ресурсов и технологий (CLARIN ERIC).

- Центры знаний – 20 в разных странах
- Архивы и репозитории
- Сервисы ч.1
 - Депозитные услуги.
 - VLO -поисковая система ЛИР (св. 1 млн)
 - Доступ к защищенным ресурсам
 - Коммутатор языковых ресурсов - поиск инструментов обработки данных.
 - Виртуальные коллекции размеченных данных
 - Реестр ЛИР. Идентификаторы
 - Инвентаризация ЛИР. Метаданные общедоступны.

CLARIN ERIC. Сервисы ч.2

- Долгосрочное архивирование ЛИР.
- Федеративный поиск
- CLARIN для исследователей (учебные и научные материалы)
- Обзоры доступных ЛИР по семействам
- Реестр курсов по ДН
- Обмен знаниями по взаимодействию между пользователями и операторами
- Мероприятия (12 в год)

Цифровая научная инфраструктура для искусства и гуманитарных наук (DARIAN ERIC)

Поддержка цифровой гуманитаристики (DH)

4 центра компетенции:

- Электронная инфраструктура
- Связи исследований и образования
- Управление научным контентом
- Пропаганда, влияние и охват

24 рабочих группы по разным направлениям

28 сервисов, ресурсов и проектов, не считая общих

DARIAN ERIC. Сервисы ч.1

- Календарь событий DH
- Публикации материалов SSH
- Библиография DARIAN
- Академические блоги SSH
- Открытый архив HAL
- Репозиторий DARIAN
- Vocabs - контролируемые словари
- SSK - стандарты и передовая практика

DARIAN ERIC. Сервисы ч.2

- Альтернатива Google Docs
- ISIDORE - платформа и поисковая система
- Реестр коллекций SSH
- Пространство Вики для DARIAN
- CodiMD - онлайн-инструмент для совместного редактирования текста.
- Общий метапоиск по ресурсам DARIAN
- TaDiRAN - таксономия цифровой гуманитаристики
- NeMO - онтология методов ДН

DARIAH ERIC. Сервисы ч.3

- AAI - сервис аутентификации и авторизации
- DARIAH-Campus - платформа для обучения и образования
- PARTHENOS –экспертная система по управлению данными в SSH
- Онлайн-реестр курсов и программ DH
- #dariahTeach - платформа учебных материалов по DH
- Topics Explorer - извлечение семантических кластеров “тем” из текстов

DARIAN ERIC. Сервисы ч.4

- MEISE- программа для редактирования нотной записи
- ConedaKOR - веб-система для администрирования и презентации материалов и коллекций DH.
- DH в немецкоязычных странах
- DKPro Wrapper - оболочка для извлечения лингвистической информации из книг
- TextGrid - среда для редактирования и публикации цифровых документов и данных.
- Geo-Browser - визуализация данных в привязке к географическим пространственным данным и к периодам времени..

Открытая научная коммуникация в социальных и гуманитарных науках OPERAS

Инфраструктура, не получившая статус ERIC

Текущие проекты

- OPERAS-P – развитие инфраструктуры открытого доступа
- COESO - сотрудничество между SSH и гражданским обществом
- OADJS - исследование журналов открытого доступа
- TRIPLE: Многоязычное и мультикультурное решение для предоставления ресурсов SSH в доступ

Законченный проект OPERAS-D – проектирование услуг и условий реализации, модель управления и бизнес-план развития открытого доступа

E-RINS - европейская исследовательская инфраструктура для науки о наследии

E-RINS ведет свою историю от проекта "Ариадна", интеграцию доступа к археологическим наборам данных.

Текущие проекты E-RINS:

- IPERION HS – новая интегрированная платформа, создание национальных узлов.
- ARCHLAB доступ к неопубликованным наборам данных из архивов европейских музеев, галерей.
- FIXLAB - доступ к стационарным исследовательским установкам для диагностики и археометрии.
- MOLAB - доступ к мобильному оборудованию для неразрушающих измерений произведений искусства, коллекций, памятников и объектов.

Сеть содействия языковым ресурсам и технологиям (FLaReNet)

Рабочие группы

- WG2 – План-график для области LRs
- WG3 – Методы и модели для LRs
- WG4 – Гармонизация форматов и стандартов
- WG5 – Протоколы и процедуры оценки
- WG6 – Автоматическое формирование LRs
- WG7 – Эволюционирующая дорожная карта

Общие проекты. Цифровая гуманитаристика и культурное наследие

- проект OpenAIRE "Единая точка входа к научным результатам в области цифровых гуманитарных наук и культурного наследия»
- Результаты исследований, данные, научные публикации и проекты, связанные с областью цифровых гуманитарных наук.
 - Публикации 3,913,851
 - Наборы данных 267,684
 - Программы 821
 - Другое 44,855

Общие проекты. CENDARI

- Совместная европейская цифровая архивная инфраструктура
- Доступ к историческим записям
- Платформа для интеграции библиотек и архивов
- Федеративный многоязычный поиск по разнородным ресурсам.
- Многоязычные авторитетные списки именованных сущностей
- Визуализация данных

Общие проекты. SSHOC.

- Создание Облака SSH
- Доступность для пользователей SSH данных, инструментов и методик
- Междисциплинарное и межстрановое сотрудничество
- Многоязычный поиск в гетерогенных ресурсах
- Безопасные среды для конфиденциальных данных.
- Принципы FAIR (Findable, Accessible, Interoperable, and Reusable)

Общие проекты. SERISS

Поиск синергии ERIC SSH: CLARIN, DARIAH, CESSDA, ESS, SHARE

- анализ различных архитектурных решений,
- оценка и улучшение качества данных и метаданных,
- создание форума инструментов и услуг,
- повышение качества данных обследований, локализация и улучшение услуг по сохранению и курированию данных,
- разработка совместной системы совместного доступа к данным и обогащения данных,
- совместная работа над правовыми и этическими аспектами,
- проведение учебной и просветительской работы,
- распространение результатов

Общие проекты BigDataEurope (BDE)

- Большие данные Европы
- Создание европейскими компаниями инновационных многоязычных продуктов и услуг на основе семантически совместимых, крупномасштабных многоязычных информационных ресурсов, доступных в соответствии с различными лицензиями и бизнес-моделями
- CESSDA в проекте BDE представляет потенциальных пользователей больших данных в SSH

Общие проекты. RISCARE.

- Европейские научные инфраструктуры в мировом контексте
- Цель проекта - представить аналитический отчет о положении и взаимодополняемости основных европейских научных инфраструктур в контексте мировой научной инфраструктуры. Отдельные разделы проекта посвящены социальным и гуманитарным наукам.

Общие проекты EOSC Enhance

- Развитие концепции европейского открытого научного облака (EOSC)
- Цель – обеспечить единую точку доступа к данным, охватывающим полный жизненный цикл данных. Это будет достигнуто путем объединения существующих научных инфраструктур, рассредоточенных по дисциплинам и государствам.

Выводы. Наиболее интересные инфраструктурные проекты ЕС в SSH

- Учет, мониторинг НИР, аналитика
- Ресурсы: депонирование, хранение, интеграция, открытый доступ
- Программные инструменты
- Лингвистика: словари, тезаурусы, метаданные, авторитетные файлы, многоязычие
- Идентификация и авторизация
- Стандарты, лучшая практика, методики, обучение
- Стратегическое планирование, устойчивость