

Т.С Маркарова, к. филол. наук
К вопросу о разработке информационной системы Центра лингвистических ресурсов и методических рекомендаций по оценке и отбору ресурсов¹

Препринт

Несмотря на обилие лингвистических ресурсов достаточно высокого содержательного и технологического качества в официальных источниках не существует академического определения понятия «лингвистические информационные ресурсы». Специалисты понимают под лингвистическими информационными ресурсами «множество определенным образом организованных речевых и языковых данных, находящихся на машинных носителях информации и используемых в различных сферах практической деятельности (образовании, промышленности, экономике, культуре, искусстве и т.д.)»².

Разрозненность, разножанровость и полифункциональность лингвистических информационных ресурсов, с одной стороны, и острая востребованность в них с другой - диктует необходимость подробного описания, систематизации и структуризации лингвистических информационных ресурсов. В Европе и США создан ряд организаций, занимающихся разработкой лингвистических ресурсов: LDC (Linguistic Data Consortium, USA), ELRA (European Language Resources Association), TELRI (TransEuropean Language Resources Infrastructure).

В то же время российское научное лингвистическое сообщество практически не имеет единого корпоративного профессионально организованного источника электронных лингвистических ресурсов всех возможных типов, отображающих/учитывающих максимально полный объем/реестр разрозненных отраслевых источников. Все еще далек от необходимого и уровень координации деятельности и научной коммуникации исследователей в области лингвистики/ языкознания. Совершенно очевидно, что назрела необходимость как в координации ученых-лингвистов, так и в интеграционном обновлении информационно-поисковых, справочно-энциклопедических, научно-образовательных лингвистических ресурсов. Такая координация и интеграция не только способствует эффективному решению теоретических, прикладных и академических задач, но и всячески улучшает корпоративную среду научных коммуникаций, оптимизирует и адаптирует поисковую систему научно-информационных ресурсов. Наиболее разумным решением, на наш взгляд, является создание корпоративной лингвистической интегрированной информационной системы (КЛИИС), которая, аккумулируя отраслевые научно-информационные

¹ Данное исследование поддержано грантом РФФИ № 18-00-002\18 " Интеграция научно-информационных ресурсов учреждений РАН (на примере языкознания) как части единого цифрового пространства РАН

² Лингвистические информационные ресурсы: Электронный ресурс [Режим доступа]: helpiks.org

ресурсы, может стать оптимальной средой и эффективной платформой для обмена информацией и коллаборативных коммуникаций.

Одной из наиболее острой проблемной задачей при разработке корпоративного электронного ресурса является экспертиза источников, в результате которой комплектование, ранжирование, классификация, идентификация источников должна приобрести верифицированный характер.

На наш взгляд, очень важным для научной деятельности/отечественной и международной научной коммуникации, для повышения конкурентоспособности исследовательской деятельности является грамотное, современное информационное сопровождение этой деятельности/коммуникации. В современной научной среде, объемной и все расширяющейся вне времени и вне границ, адекватное информационное сопровождение исследовательской деятельности предполагает не только наличие корректных информационных ресурсов, актуальную поддержку/обеспечение доступа к ним, полифункциональный сервис, но и формирование/функционирование высокопрофессионального экспертного пула во всех отраслях науки/образования. Экспертная деятельность, сочетающая в себе аналитическую и собственно экспертную, зачастую нуждается в грамотном профессиональном методическом сопровождении. Существует множество достаточно формализованных методических рекомендаций по сайтостроению, по рецензированию научных статей, по оценке деятельности учреждений/исследователей, по созданию наукометрических и библиометрических баз данных, по разработке всевозможных рейтингов и т.д. Совсем недавно появился очередной документ по Национальной электронной библиотеке (НЭБ). Это далеко не первая редакция, и, наверняка, не последняя рекомендаций по НЭБ, где в который раз рассуждают о методиках отбора, оценки и ранжирования информационных ресурсов в НЭБ³. О соотношении отраслевых информационных систем с НЭБ говорить пока рано, да и рекомендуемые критерии отличаются излишней формализованностью. Однако анализ современного объема научной информации/научных ресурсов с помощью лишь формальных критериев не может быть исчерпывающим: необходимо соблюсти корректный баланс Закона Ципфа и Принципа Парето⁴. Существенным, на наш взгляд, является еще и дифференцированный подход при экспертном анализе разноотраслевых, разножанровых информационных ресурсов с учетом специфики наук и дисциплин.

3«Положения о федеральной государственной информационной системе "Национальная электронная библиотека" и методики отбора объектов Национальной электронной библиотеки» Электронный ресурс. Режим доступа: <http://static.government.ru/media/files/yYAeqjBDCwITsAXr8LtXsdsdefAbLzi9.pdf>

4 Закон Ципфа – эмпирическая закономерность частоты слов естественного языка, количественное измерение информации; Принцип Парето заключается в следующем: 20% усилий дают 80% результата, а остальные 80% усилий — лишь 20% результата, один из способов оценки эффективности деятельности или смысловой насыщенности информации

При подготовке настоящих материалов⁵ мы постарались соблюсти разумную корреляцию формальных и содержательных критериев оценки лингвистических информационных ресурсов и учесть специфику отрасли. Как известно, лингвистика/языкознание помимо собственно фундаментальных (теоретических) проблем решает еще и множество прикладных задач в информационных системах практически всех научных дисциплин (компьютерная лингвистика, лингвистическое обеспечение информационных систем, программная лингвистика и т.д.). Это, в свою очередь, объясняет довольно широкую типовую палитру лингвистических ресурсов; документные, структурированные, фактографические. Методические рекомендации предназначены для экспертов в области лингвистики/языкознания, привлеченных к проектированию Центра лингвистических ресурсов. Цель методических рекомендаций состоит в описании/пропаганде/ наиболее рациональных способов организации лингвистической информационной системы/лингвистической базы данных, предназначенной для решения исследовательских, научных, профессиональных и образовательных задач. В системе НИРЯЗ, отражающей информационные ресурсы по лингвистике учреждений РАН, представлены следующие типы этих ресурсов (табл.1):

№	Типы лингвистических ресурсов	Кол-во
1.	Библиографии, библиотечные каталоги, архивные описи, каталоги ссылок	93
2.	Электронные коллекции и библиотеки полных текстов (книги, диссертации, отчеты, труды конференций и др.)	138
3.	Периодические, продолжающиеся издания и архивы периодики	77
4.	Корпуса	15
5.	Лексикографические ресурсы	83
6.	Этнолингвистические и социоллингвистические БД	10

⁵ Рекомендации разработаны на материале лингвистических ресурсов информационной системы НИРЯЗ <http://niryaz.alexo.beget.tech/>

7.	Лингвистические географические системы, атласы	5
8.	Электронные представления памятников письменности	5
9.	Активные лингвистические ресурсы (алгоритмы, процессоры, программы)	28
10.	Грамматики	11
11.	Описания языков и комплексные лингвистические сайты	14
12.	Информационные языки	12
13.	Энциклопедии, справочники, реестры языков	75
14.	Сведения об отдельных персонах (сайты и страницы ученых, личные фонды, биобиблиографии)	471
15.	Списки, перечни, указатели персон	31
16.	Медиаресурсы	8
17.	Сайты учреждений –владельцев ЛИР	116
18.	Сайты-сателлиты и ресурсы во внешних АИС	46
	Всего	1238

Таблица 1 Типы лингвистических ресурсов и их количество

В первую очередь необходимо:

- совершенствовать методы отбора и комплектования отраслевых ресурсов, дифференцированную технологию их представления и лингвистической обработки, т.е., строгий профессиональный отбор лингвистических ресурсов и их многоаспектную научную обработку;
- оптимизировать организацию и хранение данных ресурсов;
- обеспечить легитимный доступ к ним на основе современных компьютерных и интернет-технологий.

Корпоративная лингвистическая интегрированная информационная система (КЛИИС) должна быть многофункциональной. КЛИИС должна способствовать решению следующих задач:

- осуществление и совершенствование информационной поддержки фундаментальных и прикладных научных исследований в области лингвистики/языкознания на базе IT-технологий;
- реализация возможностей информационного взаимодействия и потенциала распределенных научных и образовательных лингвистических ресурсов локальных и глобальных сетей как основы функционирования единого информационно-лингвистического пространства;
- обеспечение легитимного доступа пользователям к ресурсам системы;
- обеспечение легитимного доступа к информации, существующей исключительно в электронной форме;
- целенаправленное актуальное информационное обеспечение специалистов;
- информационное сопровождение всех уровней и форм лингвистического образования и самообразования.

КЛИИС обеспечивает взаимодействие четырех необходимых функций:

- кумуляцию информационного потенциала лингвистики/языкознания;
- интеллектуальную, семантическую обработку информации;
- производство интегрированной и синтезированной информации;
- обеспечение коммуникационного процесса;
- возможность многоаспектного информационного поиска.

В ситуации лавинообразного роста информации, интеграционных процессов внутри и вне научных отраслей, совершенствования инструментария высокотехнологических систем становится возможной корректная и релевантная модельная визуализация вех этих процессов/явлений. Основным результатом кумуляции отраслевого информационного потенциала является формирование информационной

модели научной отрасли лингвистики/языкознания. Многоаспектный анализ информационной модели лингвистики/языкознания позволит установить соответствие между реальными ресурсами отрасли, информационными потребностями её пользователей; реконструировать в историко-географическом, предметно-содержательном и иных аспектах современное состояние науки и прогнозировать перспективы развития отечественной лингвистики/языкознания.

Интегрирование информации, различные каналы распространения информации, включение информации в сложную и разнотипную канву медиаформатов несомненно требует соблюдения некоторых условий существования информационного пространства:

- стандарты электронного взаимодействия
- взаимосвязь информационных систем различных учреждений
- координация в области информационных технологий
- централизованный подход к автоматизации
- стратегия интеграции единого информационно-образовательного пространства

В интегрированном электронном лингвистическом едином пространстве должны быть реализованы следующие сервисы (услуги):

- доступ к разнородным электронным документам в базах данных из одной точки (в среде одного экрана) благодаря единому интерфейсу, в единой поисковой среде, объединяющей разнородные электронные коллекции;
- реализация новых форм справочно-информационного обслуживания пользователей (исчерпывающий проблемно-тематический подбор информации, оперативное обеспечение доступа к полнотекстовым лингвистическим базам данных, расширение лингвистического информационного пространства, организация многофункциональных мультимедийных образовательных центров и т.д.)

Программные решения в сочетании с этими принципами помогут пользователям получать максимум информации, а создателям информации предоставить полный пакет услуг по построению и внедрению систем лингвистической информации любой сложности и назначения.

В структуру КЛИИС могут входить несколько подсистем, основанных на международных стандартах для разных типов информации, наиболее важных для каждой конкретной подсистемы. Одной из ключевых проблем интегрированной системы является управление корпоративными информационными сетями и разработка нестандартных программных и компьютерных решений для узковедомственных/узкотематических/проблемно-ориентированных систем. Это в свою очередь ставит перед разработчиками ряд программно-

технологических задач:

1. Интегрирование информации
2. Программное обеспечение по управлению информацией
3. Программное обеспечение для совместной работы, обмена информацией
4. Программное обеспечение рационализации
5. Инструмент автоматизации процессов
6. Программное обеспечение для инфраструктурного управления, включая управление безопасности и хранения данных
7. Программное обеспечение для Интернет-приложений
8. Программное обеспечение для управления жизненным циклом информации
9. Операционная система

Программно-технологическая составляющая КЛИИС, должна обеспечить необходимые и достаточные условия для развития и поддержки следующих ее свойств:

1. Наличие концептуальной модели
2. Наличие методического руководства для ввода и получения информации
3. Возможность совместной обработки ресурсов различных типов
4. Возможность экспорта/импорта информации или фрагментов информации
5. Наличие разветвленной системы метаданных
6. Возможность корректировки, декомпозиции системы
7. Возможность обработки фрагментарной информации
8. Возможность обработки динамических ресурсов
9. Возможность корректировки ошибочной информации
10. Возможность обработки нечеткой информации

Для эффективной работы корпоративной лингвистической информационной системы, для оптимальной работы в системе пользователей, наконец, для создания корректной информационной модели лингвистики/языкознания очень важны и значимы критерии (т.е, правила выбора) оценки ресурсов, составляющих контентное ядро КЛИИС. Несмотря на очевидность и мнимую простоту задача разработки критериев оценки и отбора отраслевых ресурсов довольно трудоемкая, эта задача, пожалуй, не имеет простых и однозначных решений. Безусловно, невозможно учесть все существующие и перспективные качества и нюансы как информационных ресурсов, направлений лингвистики/языкознания, так и возможных пользовательских запросов. При разработке критериев оценки лингвистических ресурсов мы использовали три аналитических принципа:

- **Синтаксический принцип** – учет типа носителя, способа представления информации, скорость передачи и обработки, размеры кодов представления информации;
- **Семантический принцип** – корреляция семантических свойств информационных блоков ресурса с лингвистикой/языкознанием и с пользовательскими потребностями;
- **Прагматический принцип** – полезность ресурса с точки зрения решения прикладных, практических задач лингвистики/языкознания.

Основываясь на информационных ресурсах, отраженных в системе НИРЯЗ и опираясь на перечисленные принципы анализа информационной системы, предлагаем ряд сгруппированных по разным основаниям/признакам критериев для оценки лингвистических ресурсов. Эти критерии целесообразно объединить в несколько типов, представленных на табл. 2:

Качественные критерии	Количественные критерии	Параметрические критерии
Актуальность	Видимость	Объем ресурса в логических единицах
Легитимность	Полнота	
Верифицированность	Цитируемость	Размер файлов/ страниц
Доступность	Скачиваемость	
Устойчивость	Посещаемость	Размер графических изображений
Уровни навигации	(статистические показатели)	
Структурированность	Количество	Длительность аудио и видео ресурсов
Интероперабельность	перекрестных (внутренних) ссылок	
Полифункциональность		
Универсальность?		
Уникальность		
Информационный дизайн (инфографика)		
Разнообразие форм представления		
Статичность		
Динамичность		
Практикоориентированность		
Когнитивность		
Проектность		
Аналитичность		
Интерпретационность		
Синхроничность		
Диахроничность		
Релевантность		
Технологичность		
Функциональность		

Лингводидактичность		
---------------------	--	--

Таблица 2 Критерии оценки ресурсов

Кроме того, критерии подразделяются на:

- **универсальные** (общие для всех типов/видов ресурсов любой тематики);
- **специфические** (дифференциальные для разных типов/видов ресурсов, коррелирующие с особенностями лингвистической науки/языкознания).

Универсальные критерии	Специфические критерии
Видимость	Когнитивность
Актуальность	Проектность
Полнота	Аналитичность
Легитимность	Интерпретационность
Верифицированность	Синхроничность
Цитируемость	Диахроничность
Скачиваемость	Релевантность
Доступность	Технологичность
Устойчивость	Функциональность
Уровни навигации	Лингводидактичность
Структурированность	
Интероперабельность	
Полифункциональность	
Посещаемость(статистические показатели)	
Универсальность?	
Уникальность	
Информационный дизайн (инфографика)	
Разнообразие форм представления	
Статичность	
Динамичность	
Практикоориентированность	

Таблица 3 Универсальные и специфические критерии

Разыскания в on-line словарях и энциклопедиях классических, академических дефиниций отобранных критериев отбора и оценки лингвистических ресурсов не дали положительных результатов. Поэтому предлагаем собственные определения/толкования, основанные на сведениях

из различных научных дисциплин с использованием материалов словарей/энциклопедий на Academic.ru и Википедии⁶.

1. **Когнитивность** – 1) степень обработки/переработки информации; 2) уровень «знаниевого» содержания
2. **Лингводидактичность** – корреляция с образовательными потребностями/проблемами/запросами в области лингвистики
3. **Проектность** – логическая цепочка алгоритмов, обуславливающая потенциальное развитие отдельного фрагмента ресурса
4. **Аналитичность** – 1) степень представления информации с учетом обобщенно-конкретных (аналитико-синтетических) признаков контента; 2) идентификационный уровень корреляции запроса/поиска/результата поиска внутри ресурса и вне его
5. **Технологичность** – 1) наличие/отсутствие лингвистического инструментария для исследовательских задач текстов/языков/произведений и поисковых задач ресурса
6. **Функциональность** - объем и качество набора операционных возможностей
7. **Релевантность** – степень соответствия заявленной тематике
8. **Диахроничность** – отражение исторического развития проблемных фрагментов контента
9. **Синхроничность** - исчерпывающее отражение актуального/современного состояния содержания ресурса
10. **Интерпретационность** – 1) степень раскрытия смысла контента; 2) перспективная возможность извлечения различных смыслов/ «знаниевых» фрагментов для построения отраслевой/тематической онтологии
11. **Видимость** – показатель позиции выдачи по соответствующим тематическим запросам в поисковых системах
12. **Актуальность** – значимость ресурса в настоящее время
13. **Полнота** – степень отражения существующих информационных потоков проблематики
14. **Легитимность** – степень законности/легальности контента
15. **Верифицированность** – соответствие ресурса/контента ресурса корректным истинным суждениям/ положениям лингвистической науки
16. **Цитируемость** – количество ссылок на ресурс в поисковых системах и на других ресурсах
17. **Скачиваемость** – цифровой показатель количества скачиваний фрагментов ресурса

⁶ Словари и энциклопедии на Академике. Электронный ресурс. Режим доступа: <https://dic.academic.ru/>; Википедия. Свободная энциклопедия. Русская версия. Электронный ресурс. Режим доступа: <https://ru.wikipedia.org/wiki>

18. **Доступность** – возможность обращения к ресурсу независимо от местонахождения и времени
19. **Устойчивость** – степень стабильности ресурса/информационной системы и способность к функциональному восстановлению
20. **Уровни навигации** – глубина и многообразие средств, способствующих ориентации в контенте и структуре ресурса
21. **Структурированность** – степень и уровень упорядоченности информационных массивов, релевантный основному целевому назначению ресурса
22. **Интероперабельность (функциональная совместимость)** – степень возможностей взаимодействия с другими ресурсами, системами различного технического оснащения и технологического сопровождения
23. **Полифункциональность** – исчерпывающий набор инструментария для работы
24. **Посещаемость** – 1) статистические показатели посещения ресурса; 2) основная метрическая характеристика эффективности
25. **Универсальность** – энциклопедичность, разносторонность контента
26. **Уникальность** – самобытность, эксклюзивность контента
27. **Информационный дизайн (инфографика)** – эргономическое графическое представление разноплановой, разножанровой, разной степени сложности информации
28. **Разнообразие форм представления** – наличие различных форм представления информации: текст, видео, графические символы, анимация, аудио
29. **Статичность** – наличие статичных страниц, снабженных мета-тегами, представляющими ресурс в поисковых системах
30. **Динамичность** – наличие динамичных страниц; возможность модификации и генерации контента по конкретному запросу
31. **Практикоориентированность** – потенциальная возможность использования ресурса при решении конкретных (исследовательских, научных, образовательных, просветительских задач)

Разумный баланс применения этих критериев при отборе и оценке лингвистических информационных ресурсов для КЛИИС позволит создать актуальную информационную модель лингвистики/языкознания с учетом как современных тенденций развития науки, так и классических направлений в преемственной их взаимосвязи.

Особый информационный массив в системе КЛИИС составляют научные журналы по лингвистике/языкознанию. Острая и болевая проблема с публикационной активностью ученых/исследователей, различными списками журналов (белыми, черными, статусными, «мусорными») в разных системах у разных агрегаторов значительно затрудняет исследовательскую

деятельность. Констатируя существующее реальное положение дел с научной периодикой, мы сочли необходимым выделить критерии оценки лингвистических журналов в специальный раздел. Несмотря на то, что некоторые критерии пересекаются с теми, которые мы предлагаем для оценки информационных лингвистических ресурсов вообще, в аспекте современного состояния периодики многие из критериев наделены другими смыслами. Для лингвистических журналов мы выделяем:

- Универсальные критерии
- Содержательные критерии
- Функционально-формальные критерии

Универсальные критерии:

- Номер свидетельства о регистрации
- Номер ISSN
- Тираж (для бумажных изданий)
- Периодичность
- Юридический адрес/электронный адрес
- Учредитель издания
- Наличие сайта
- Наличие англоязычной версии

Содержательные критерии:

- Соответствие целевому назначению
- Объективность (системность и комплексность отражения проблем, освещение разных точек зрения и т.д.)
- Достоверность информации (аргументированность, ссылки на документы, официальные публикации и т.д.)
- Оригинальность (проверка на антиплагиат)
- Качество подачи материала (аналитический стиль, корректное терминопотребление, элементарная грамотность)
- Практикоориентированность
- Тематико-отраслевая релевантность

Функционально- формальные критерии:

- Наличие в зарубежных и отечественных базах данных (Webofscience , Scopus, РИНЦ? и т.д.)
- Импакт-фактор журнала
- Наличие института экспертов/рецензентов
- Наличие в фондах крупных традиционных и электронных библиотек
- «Видимость» в Интернете, научных соц. сетях и т.д.
- Востребованность (по статическим данным из разных источников - библиотеки, сайты, поисковые системы – Rambler, Yandex, Google и т.д.)
- Экономическая транспарентность
- Экспертная транспарентность

- Наличие двуязычного сайта (дизайн, полифункциональность, поисковые и пользовательские возможности – работа с текстами статей, навигация по различным словарям, тезаурусам, классификациям; оперативная обратная связь и т.д.)
- Корректный научный аппарат (ссылки, библиография, аннотация/реферат, ключевые слова – на двух языках)

Будучи довольно объемным научно-информационным ресурсом, научная периодика, на наш взгляд, переживает в настоящее время кризис. Актуальность, информационная/информативная ценность научных фактов/сведений/информации, составляющих основное содержание отраслевых журналов теряет значимость в силу ряда причин. Это происходит по причине: несвоевременности доступа к информации, недоступности информации для исследователей, устаревшей в современную научную информационную эпоху журнальной идеологии, излишней формализации, коммерциализации, с одной стороны, и бессмысленной контрольно-надзорной функции над научной периодикой (отчетные показатели, финансовая зависимость и т.д.), с другой. По нашему мнению, для повышения эффективности научных достижений и повышения публикационной активности ученых необходимо создание распределенной отраслевой базы данных лингвистических знаний/лингвистической науки, своеобразного отраслевого репозитория (по примеру «Соционета»⁷). Перспектива развития КЛИИС от документных ресурсов к базе знаний требует более корректного использования некоторых терминов, прежде всего таких как: **сведения, информация, знания.**

«**Сведения** - 1. Известия, сообщения о чем-либо. 2. Факты, данные, характеризующие кого-либо, что-либо. 3. Познания в какой-либо области, осведомленность в чем-либо. 4. Отчет с цифровыми данными»⁸.

«**Информация** (от латинского *informatio* - разъяснение, изложение), первоначально - сведения, передаваемые людьми устным, письменным или другим способом (с помощью условных сигналов, технических средств и т.д.); с середины 20 века - общенаучное понятие, включающее обмен сведениями между людьми, человеком и автоматом, автоматом и автоматом, обмен сигналами в животном и растительном мире, передачу признаков от клетки к клетке, от организма к организму; одно из основных понятий кибернетики»⁹.

«**Знания** -1. Проверенный практикой результат процесса познания действительности, адекватное её отражение в сознании человека в виде представлений, понятий, суждений, теорий. 2. Исторически сложившаяся, непрерывно развивающаяся и постоянно сверяющая свою методологию и имеющийся багаж знаний с новым опытом система знаний о закономерностях развития природы, общества и мышления и о способах их

7 - <https://socionet.ru/>

8 Толковый словарь Ефремовой. Т. Ф. Ефремова. 2000. Электронный ресурс. Режим доступа: <https://dic.academic.ru/>

9 Современная энциклопедия. 2000. Электронный ресурс. Режим доступа: <https://dic.academic.ru/>

планомерного воздействия на окружающий мир; **III.** Обладание какими-либо сведениями, осведомленность относительно кого-либо или чего-либо; навыки, знания, полученные человеком; **2.** Владение какими-либо практическими навыками, умениями»¹⁰.

Из перечисленных дефиниций очевидно, что граница между этими понятиями довольно зыбкая, и во многих практических случаях одно составляет часть другого, синонимично третьему и т.д. Но хотя бы приблизительное понимание контекстного значения этих терминов, корректное и уместное их употребление значительно облегчит классификацию ресурсов, отражаемых в КЛИИС.

КЛИИС должна обладать широкой функциональностью современного репозитория со встроенной системой антиплагиата, с возможностью публикации и самопубликации результатов интеллектуальной деятельности, включая монографии, статьи, методические разработки, дипломные работы, диссертации, а также фрагменты всех этих и других видов научных работ.

Мы надеемся, что такой системой или отраслевым репозиторием может стать КЛИИС. Это позволит сформировать сетевую коллаборацию, оптимизирует научную коммуникацию, междисциплинарное, межотраслевое, межведомственное, международное сотрудничество. Кроме того, такое упорядочение информации будет иметь важное значение для повышения качества научной экспертизы, для прогнозирования и планирования проблем исследований, для разработок актуальных научных задач с учетом общей и частной ситуации в той или иной отрасли.

Существенным является разработка на основе такой базы данных рекомендательно-экспертного сервиса, обеспечивающего профессиональную и общественную экспертизу, консультирование, рецензирование, поддержку дискуссий и т.д.

Однако все эти функции могут получить развитие только в случае государственной заинтересованности и поддержки.

В мировой науке широко распространяется идеология Открытой науки (OpenScience)¹¹. Это конечно, относится и к лингвистическим ресурсам.

Отечественные базы данных лингвистической науки/лингвистического знания при определенных условиях (двухязычие: аннотации, ключевые слова, рефераты, корректные, с соблюдением формальных правил ссылки, библиографий и т.д.) будут подготовлены к интеграции в международные базы данных. Желательно предварительно сформировать алгоритм подсчета всяческих рейтингов, индексов цитирования внутри конкретной лингвистической науки. Безусловно, все это должно вырасти в мощную информационную систему с набором метаданных, поисковой навигацией, единым (для отрасли входом, гибким интерфейсом и т.д.

В настоящее же время весь наукометрический и библиометрический инструментарий, изначально предназначенный для создания комфортной

10 Толковый словарь Ефремовой. Т. Ф. Ефремова. 2000. Электронный ресурс. Режим доступа:

<https://dic.academic.ru/>

11 <http://iradche.ru/2014/04/04/open-science-3/>.

среды в помощь научным коммуникациям, используется весьма некорректно - только лишь для подсчета рейтингов/индексов цитирования/импакт-факторов и т.п. Большие человеко-затраты, финансовые вложения (на подписку, на издание многочисленных, зачастую дублирующих тематику, журналы) себя не оправдывают. Ученые в массе своей перестают думать об исследовании как таковом. И вместо того, чтобы дискутировать о результатах своей деятельности корректных/спорных/неверных/ - разных, исследователи озабочены публикациями и погоней за «нужными» журналами.

Мы надеемся, что существование репозитория лингвистической информации будет не только мощным информационным обеспечением и поддержкой исследовательской деятельности в области языкознания, но и послужит гибкой платформой/средой для научных коммуникаций.

Локальный, фрагментарный анализ лингвистических ресурсов показал, насколько важными при формировании отраслевой научно-информационной системы является отбор и оценка/ранжирование этих ресурсов. При всей закономерной субъективности методов отбора и оценки ресурсов отраслевая информационная модель, в нашем случае, модель лингвистики/языкознания все-таки объективно репрезентирует актуальное положение дел в конкретной науке.

Настоящие предложения по методическим рекомендациям по отбору и оценке лингвистических информационных ресурсов, носят экспериментальный или пилотный характер. Однако, по нашему мнению, они могут стать начальным этапом для формирования полифункциональной, многоаспектной системы экспертизы отраслевых информационно-научных ресурсов.