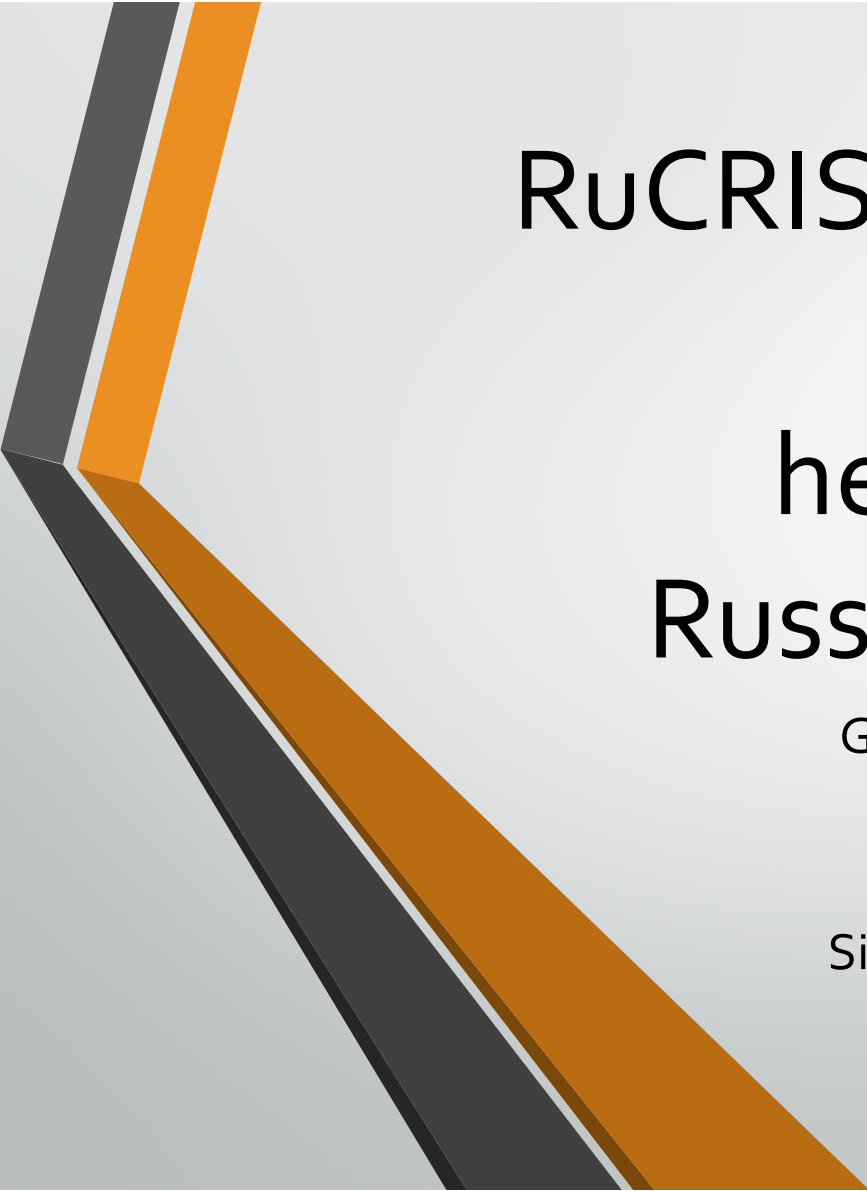


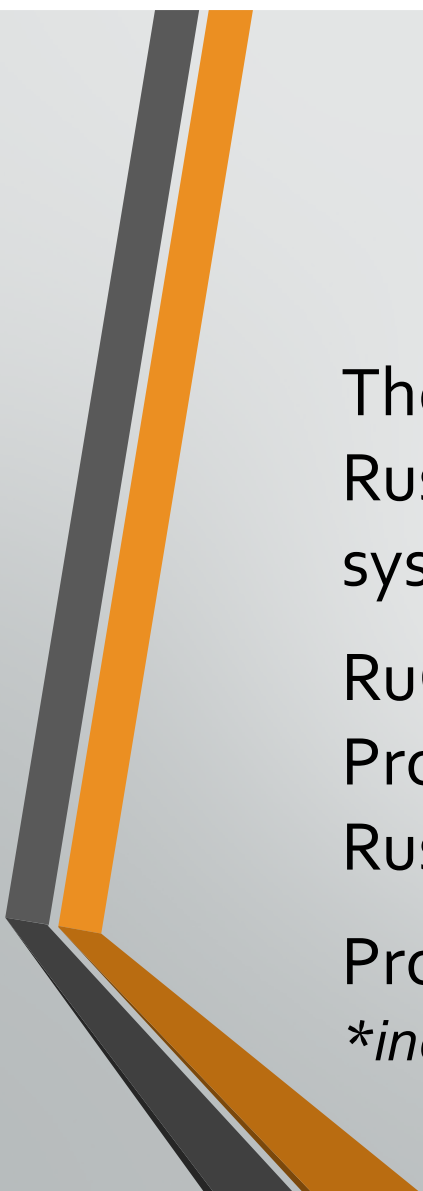


RuCRIS: a pilot CERIF-based system to aggregate heterogeneous data of Russian research projects

Guskov A.E.¹, Zhizhimov O.L.¹, Kikhtenko V.A.²,
Skachkov D.M.¹, Kosyakov D.V.¹

¹Institute of computational technologies of the
Siberian branch of Russian Academy of Sciences,

²Novosibirsk State University
(Novosibirsk, Russia)



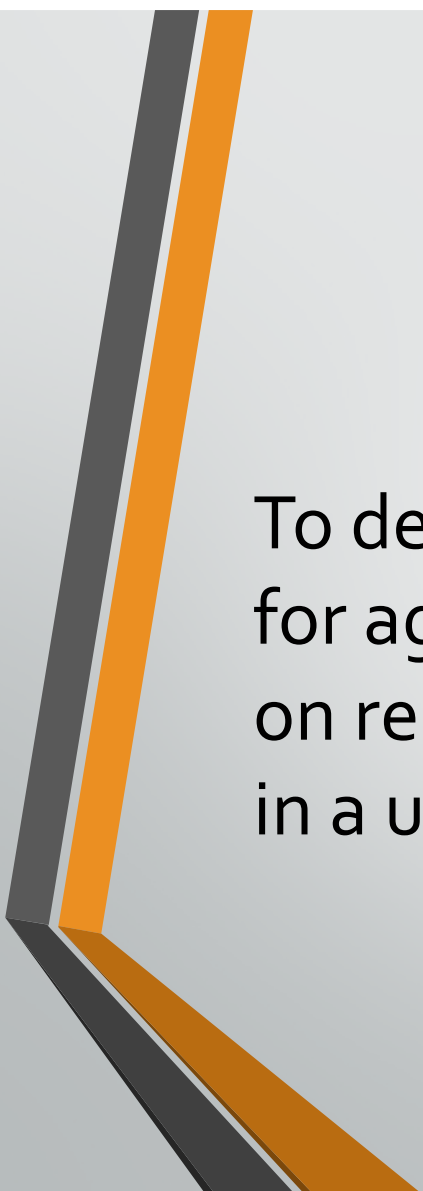
Introduction

There is a strong demand among scientific society and Russian government for information and analytics system for Russian research projects.

RuCRIS project, that was supported by Federal Program of Ministry of Education and Sciences of Russia is the effort to build such kind of system.

Project duration: 3 months*

**including 1 month for reporting stage*



Goal

To develop technologies
for aggregating and managing information
on research projects and result outputs
in a unified repository

Project tasks

- 1. Analysis of the existing methods for collecting, processing and distributing research deliverables**
- 2. Design data gathering methods and procedures**
- 3. Pilot implementation, which provides:**
 - Research projects and results data harvesting.
 - Data classification, mapping and linking.
 - Storing the data in the central repository.
 - Visualizing data from the repository.

Potential consumers

Consumers	Purposes
Government	Research strategic management
Funders	Grants distribution
Research institutes and universities	Search for collaboration
Researchers	Search for related projects
Business, Industry	Search for innovations

Data sources

Repository	Access type	Project	Person	Organization	Result product
Russian Foundation for Basic Research	HTML	+			
Russian Foundation for Humanities Research	HTML	+	+		
Republican Research Scientific and Consulting Centre for Expertise	HTML	+	+	+	
Repository of Federal Programs of Russian Ministry of Education and Science	XML	+		+	
Center for Information Technologies and Systems for Executive Authorities	XML	+		+	+
Projects repository of the Siberian Branch of the Russian Academy of Science	DB	+	+	+	

Problems

Problem	Description
Low collaboration level	Data owners aren't engaged into integration efforts

Low collaboration problem

Data owners aren't engaged into integration efforts



Project catalogues on the website
(name, annotation, executors)

Some analytics in annual reports



No definite results for almost all projects

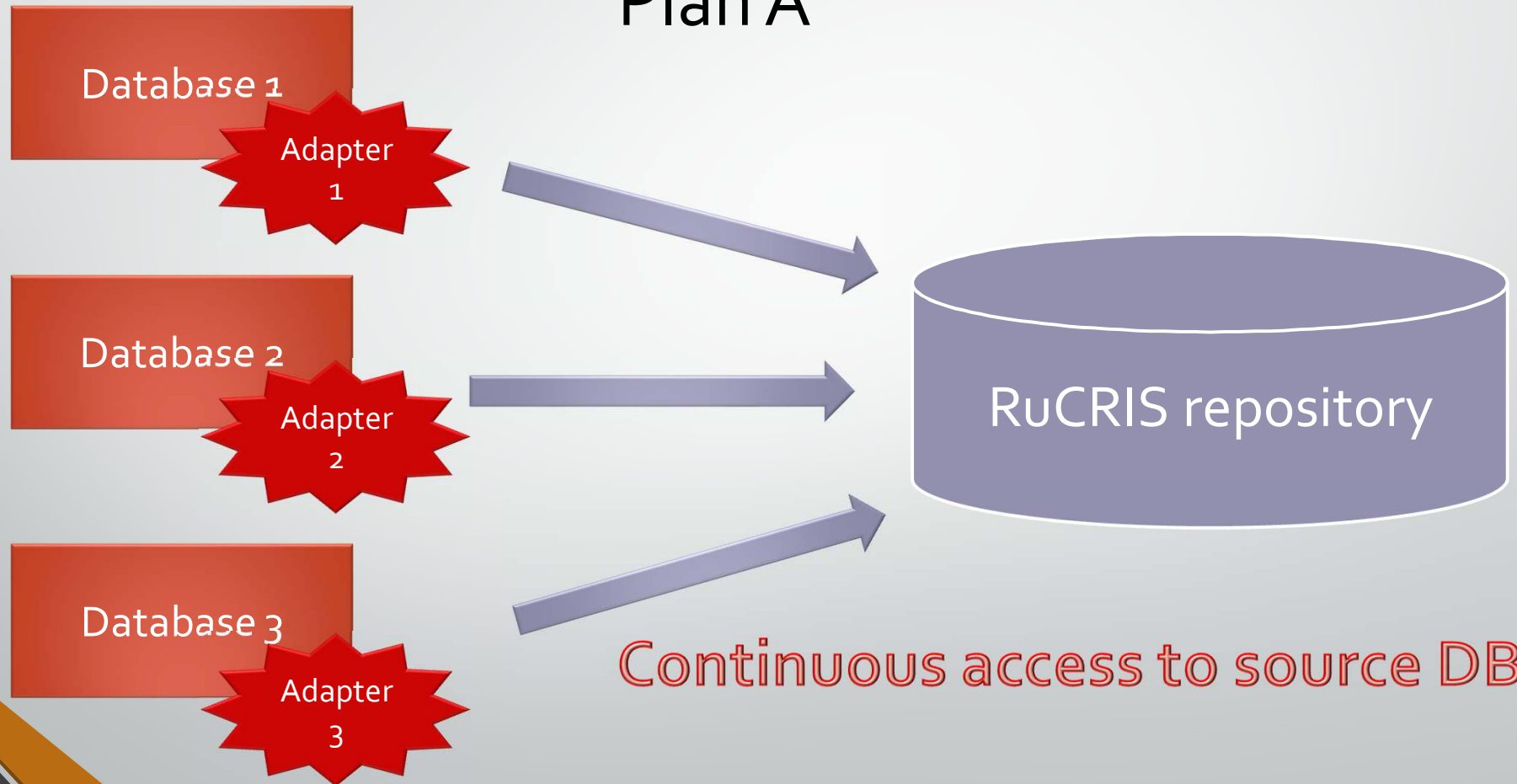
No API to access the data

No custom analytics

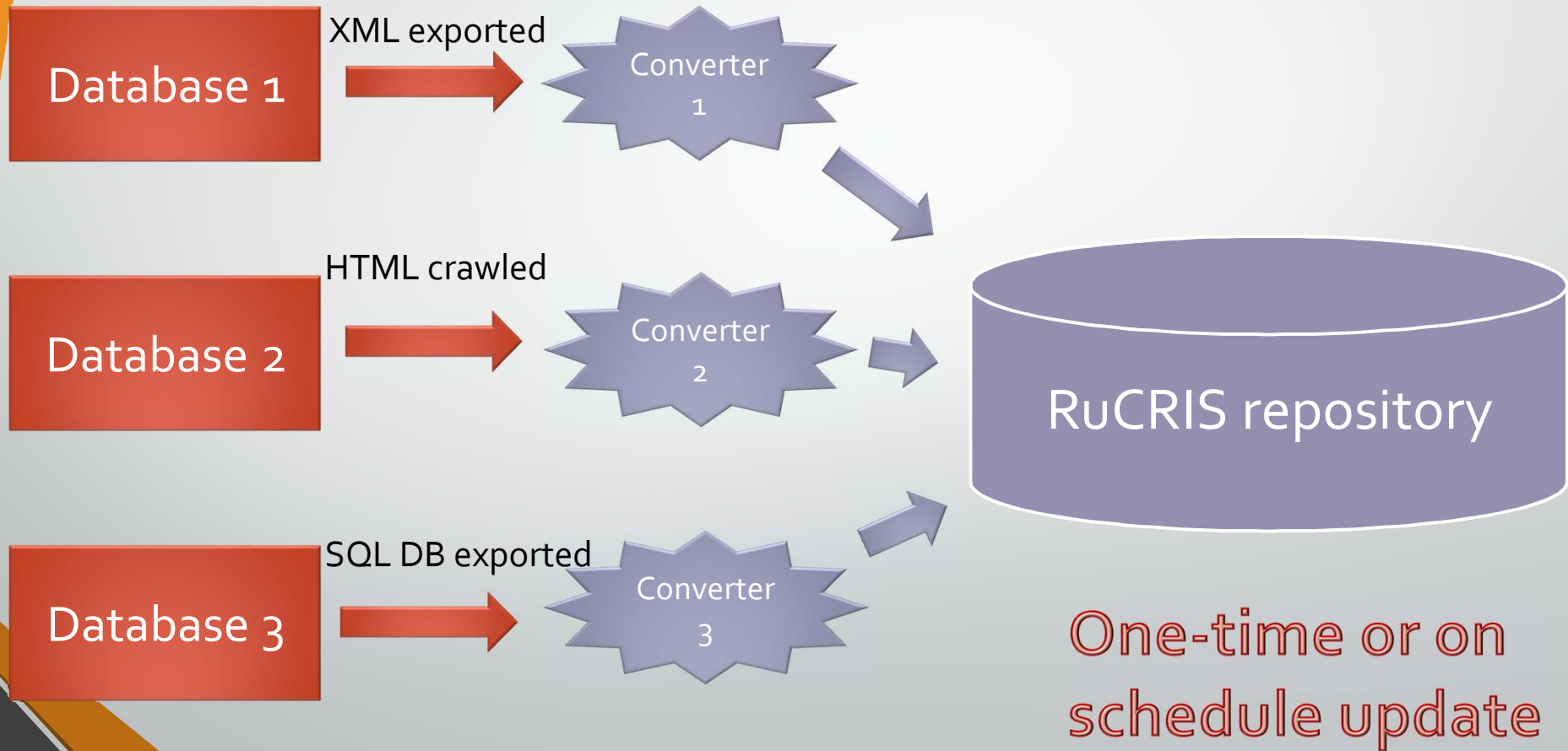
Problems

Problem	Description
Low collaboration level	Data owners aren't engaged into integration efforts
Syntactic heterogeneity	There is no common way to collect the original data in preassigned systems
Semantic heterogeneity	
Different data schemas	Different systems use their own data schemas and don't provide any API.
Different classifiers	Similar classifiers contain very different classes (e.g. science fields)
Duplicate entities	The same entity exists in different sources.

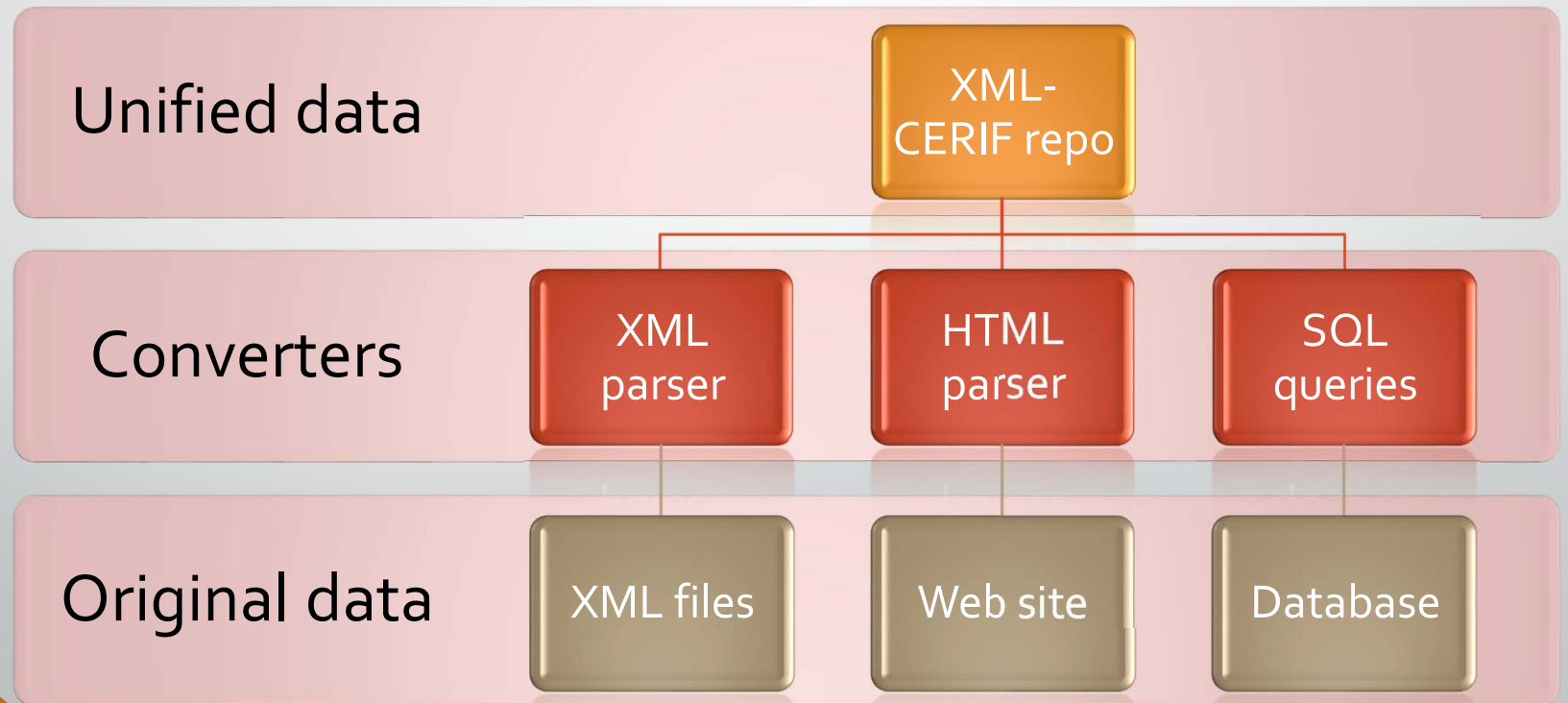
Access data sources. Plan A



Access data sources. Plan B



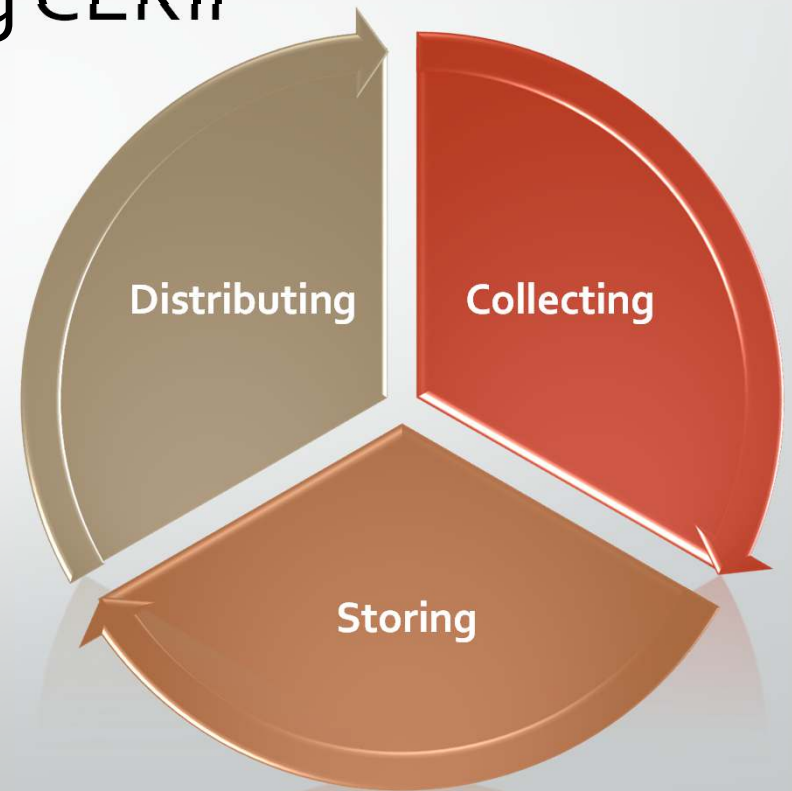
Problem: syntactic heterogeneity
no common way to collect the original data



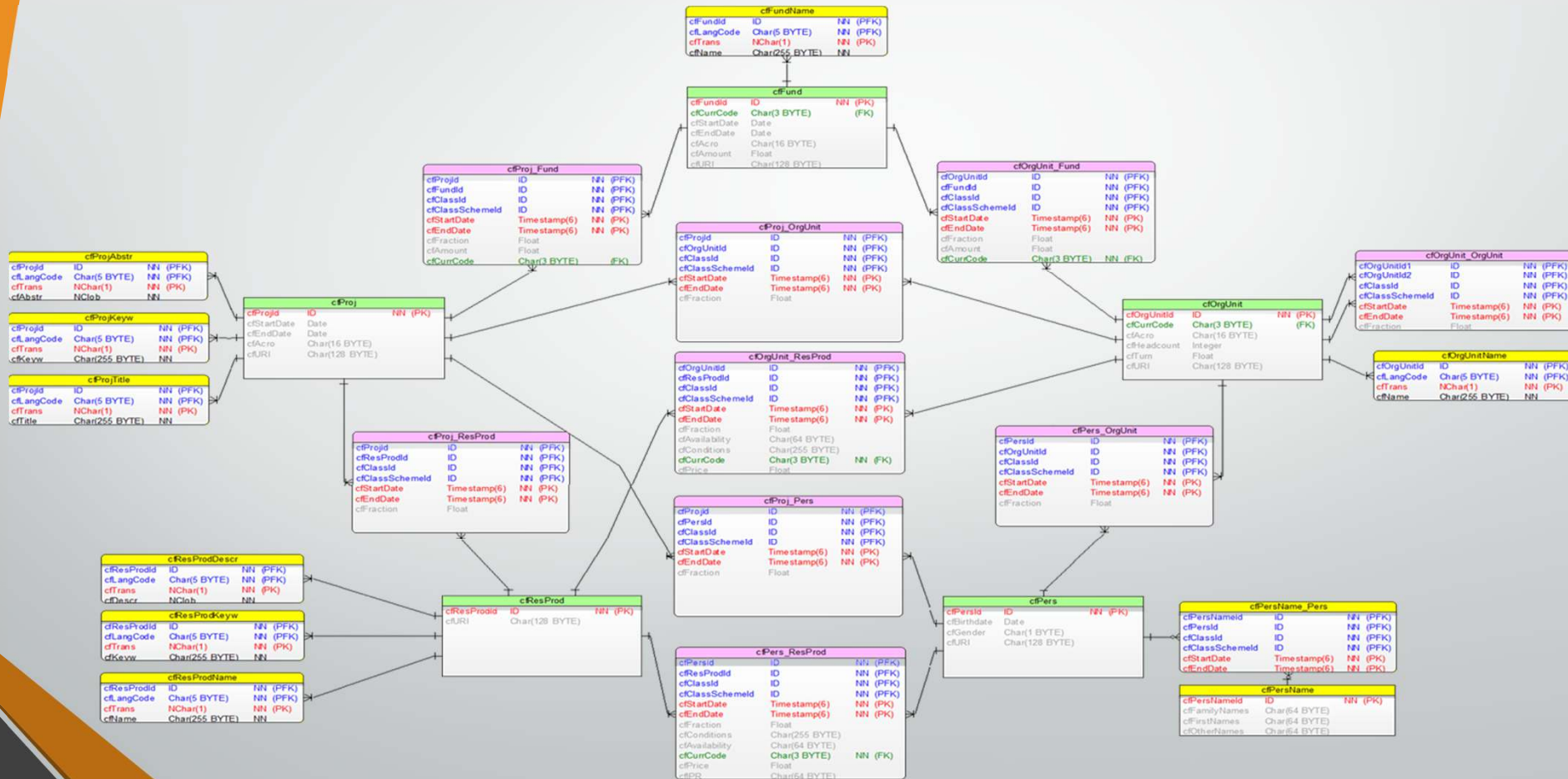
Benefits of using CERIF

Research project management system have to be of national scale and even more – have strong connections with European CRIS systems.

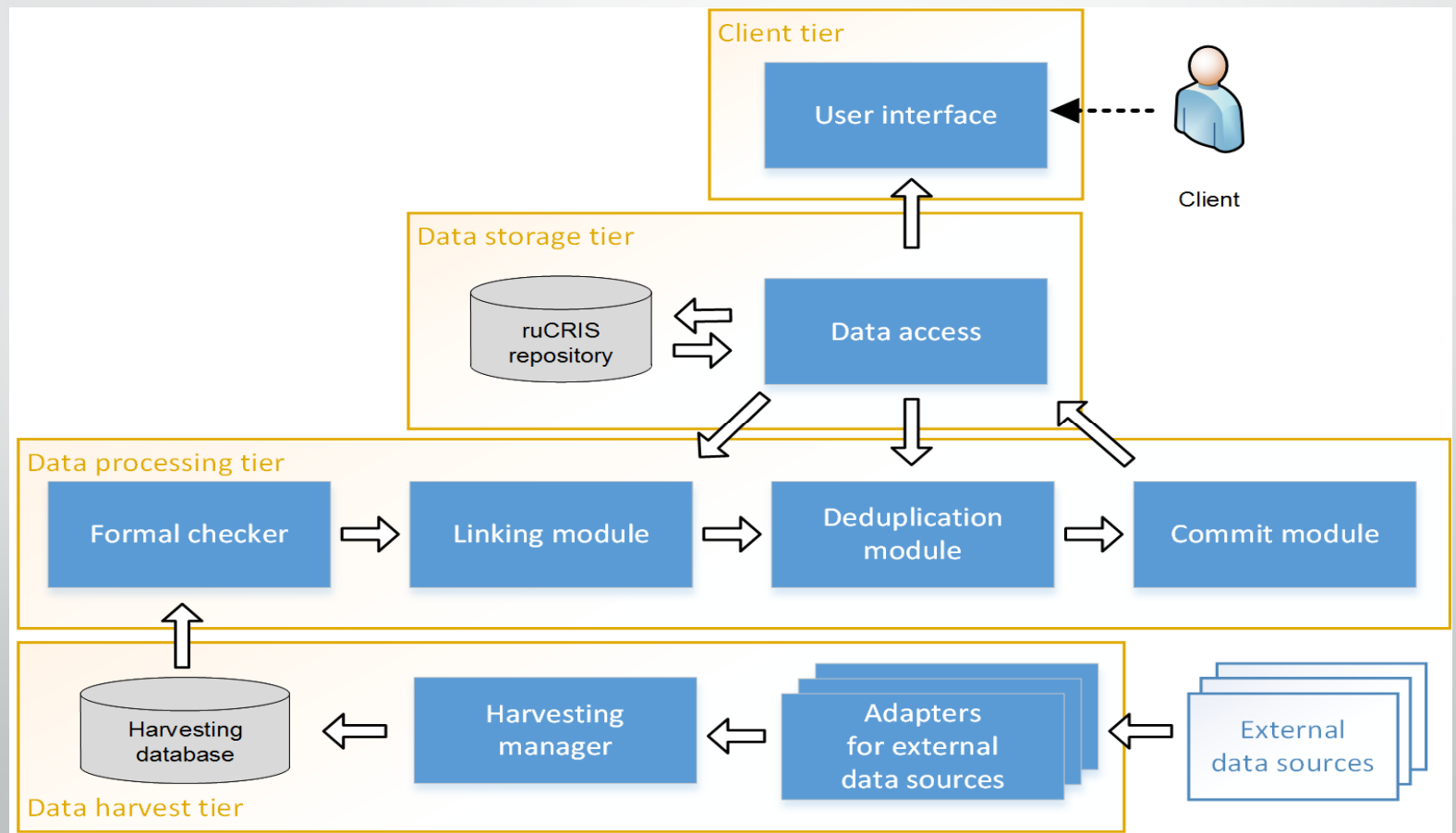
Using CERIF for data storage and exchange is huge benefit in achieving of this goal.



Essential parts of CERIF



ruCRIS architecture



Duplicates problem:

Is it the same John Smith or not?

Entity	Preferable way	Alternative way
Person	Check global IDs (e.g. ORCID)	Check name, degree, affiliation, email
Organization	Check global IDs (e.g. tax payer ID)	Check names, location
Classifier	Use unified classifiers	Generalization

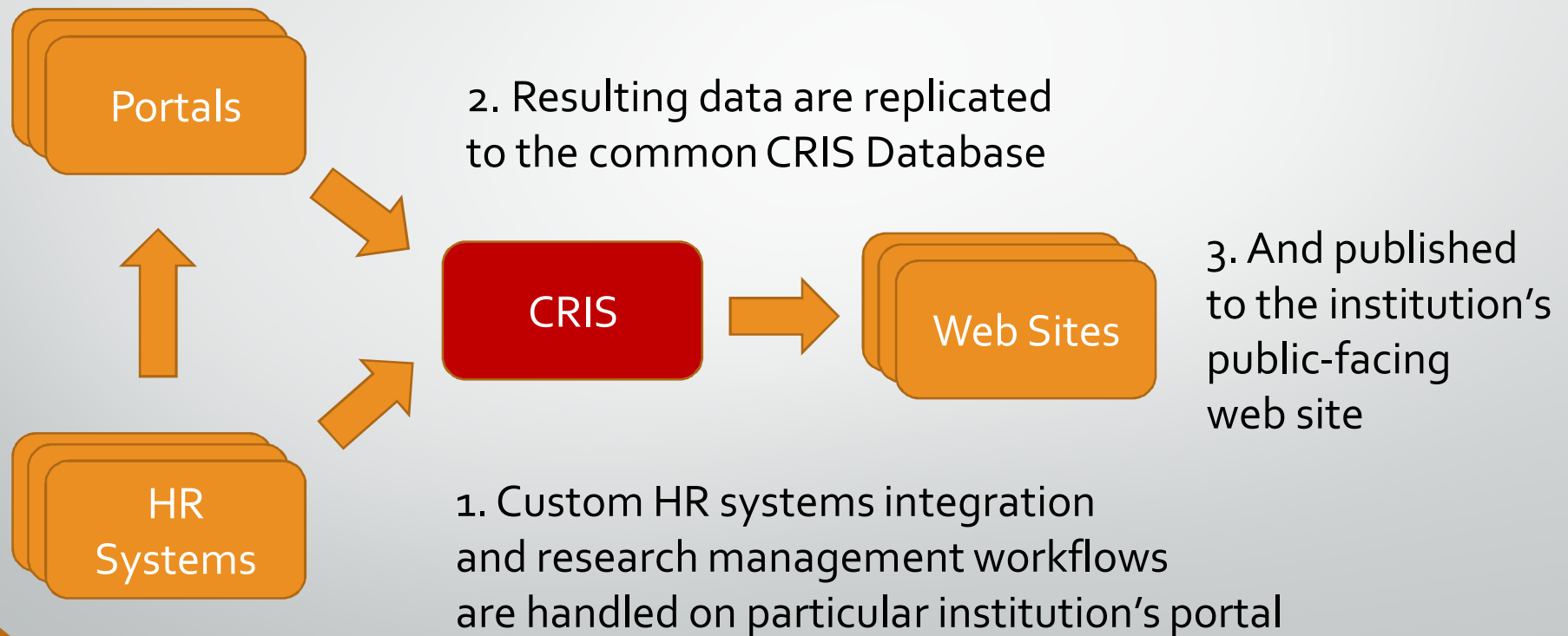
RuCRIS statistics

- Projects: 22 130
- Persons: 9 300
- OrgUnits: 2 700



Source	Projects
RFBR	9 587
RFHR	3 929
CIT	4 829
RRSCCR	2 247
FP	176
SBRAS	1 369
TOTAL	22 130

Common scenario of current research data movement



CRIS/CERIF entities representation

All main CERIF entities are represented as corresponding web pages with unique and persistent URLs:

- OrgUnits
- Persons
- Projects
- Publications
- Et cetera

Программа VIII.68.1. Стратиграфия, палеобиогеография и экосистемные реконструкции протерозоя и фанерозоя Сибири и российского сектора Арктики

ИНГГ ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ ВОСПОМОГАТЕЛЬНОЕ УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ НЕФТЕГАЗОВОЙ ГЕОЛОГИИ И ГЕОФИЗИКИ ИМ. А.А. ТРОФИМУКА
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК

Новости Институт Наука Образование Сотрудничество

Лаборатория палеонтологии и стратиграфии докембрия

В 1963г. академиком Б.С.Сколовым, в рамках лаборатории палеозоя Института геологии и геофизики, была создана подкаембрийская группа в составе Б.В.Хоментовский, В.Ю.Шенфиль, М.С.Яшин. Деятельность этой группы, возглавляемой Б.В.Хоментовским, началась с изучения стратиграфии позднего докембрия и кембрия Платонского нагорья и среднего течения р.Лена (1962-1964 гг.), в конце 60-х годов область исследований расширилась на все образование Сибирской платформы. В это время состав группы пополнился Е.Л.Булаковым, А.С.Губишери. Результатом первого этапа исследований явилась монография "Опорные разрезы отложений докембрия и кембрия Сибирской платформы" вышедшая в 1972 году.

В 70-ые годы было широко поставлено изучение практически всех осадочностей позднего докембрия. Расширилась география проведения работ за счет включения в сферу исследований группы районов складчатого обмывания Сибирской платформы (Саяно-Алтайская складчатая область, Таймыр) и ее внутренних районов. Значительно вырос состав группы за счет появления новых сотрудников: Б.Т.Литвиленко (1972 г.), аспирантов А.А.Постникова и А.А.Теряева (1977), позже - А.Н.Диденко, Г.А.Карловой, К.Л.Пала, С.С.Брагина.

В 1977 году группа была выделена в лабораторию стратиграфии и палеонтологии позднего докембрия под руководством Б.В.Хоментовского. Сотрудники лаборатории активно участвовали практически во всех межрегиональных и международных совещаниях по проблемам стратиграфии и палеонтологии позднего докембрия, а также в работах СибРМКС в создании всех унифицированных стратиграфических схем позднего докембрия Сибири.

В 1997 году, в связи с выделением самостоятельного Института геологии нефти и газа в составе СВНТИМ СО РАН, изменился кадровый состав лаборатории: часть сотрудников перешла в другие структуры, был значительно обновлен состав коллектива за счет молодежи (Б.В.Хомчич, К.Е.Наговицин, М.Ш.Яебуллин). Заведующим выбран к.т.-м. А.А.Постников. В 1998 г. произошло объединение с кембрийской группой сотрудников, возглавляемой В.А.Луценковой, и переименование созданного коллектива в лабораторию палеонтологии и стратиграфии позднего докембрия и кембрия.

С 80-х годов основной задачей стало углубление исследований с целью разработки подробных региональных стратиграфических схем и комплексному обоснованию подразделений Сибирской шкалы докембрия и кембрия. Ведущая роль в исследованиях лаборатории попрежнему остается за палеонтологическими методами, охватывающими практически все группы органических остатков этого времени: трилобиты, археоцаты, брахиоподы, водоросли, мелкоресничатая фауна, фитолиты, микрофоссилии и проблематичные докембрийские остатки.

По плану НИР Института лаборатория проводит исследования по программе 27.2 «Стратиграфия, биогеохронология и типизация экосистемных пестроек в протерозойско-фанерозойской истории осадочных бассейнов Сибири, связь с

Заведующий лабораторией


Постников Анатолий Александрович
кандидат геолого-минералогических наук, старший научный сотрудник

Контактная информация
Рабочий телефон: +7(383) 333-33-29
e-mail: PostnikovAA@igg.sbras.ru

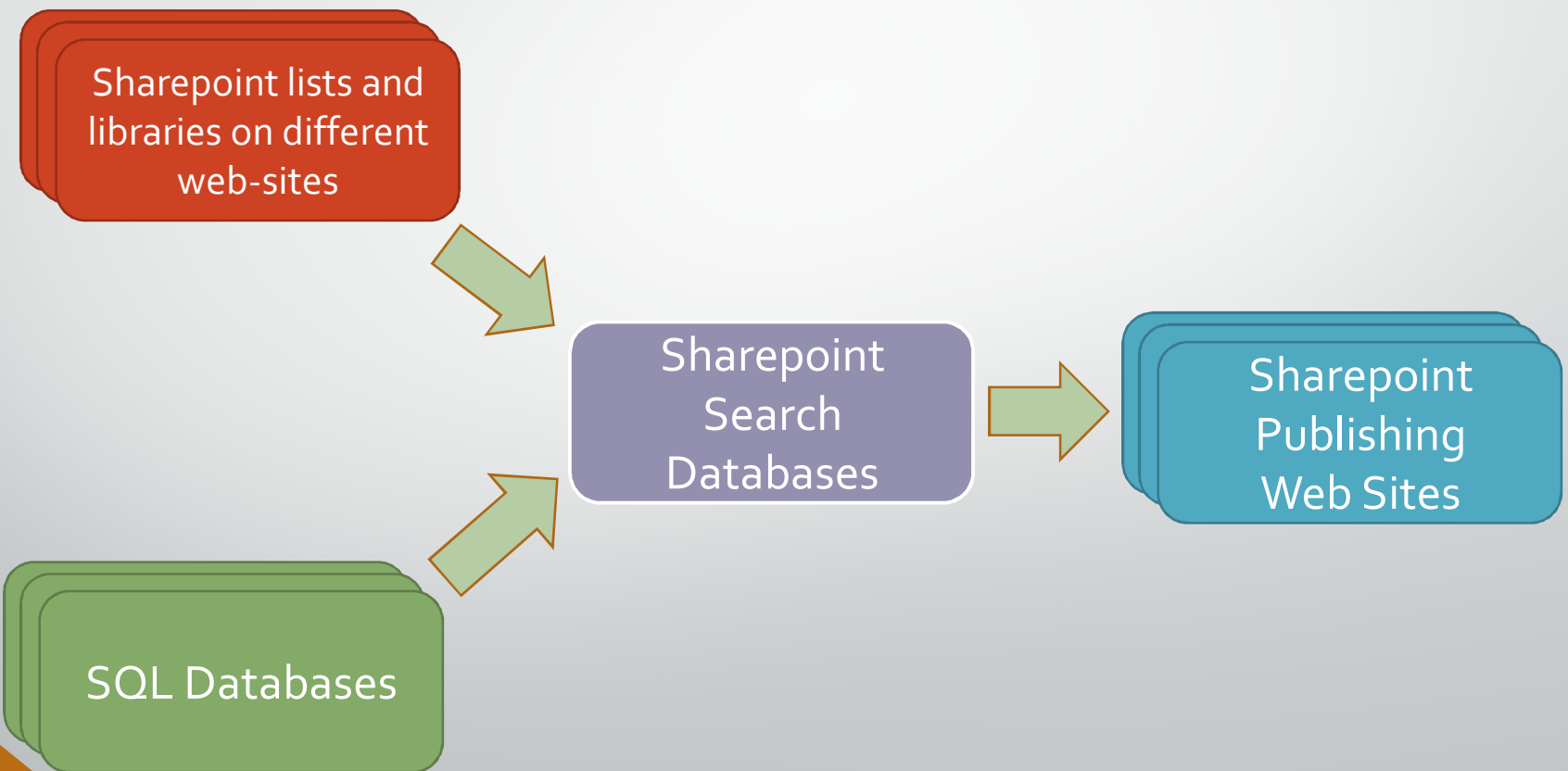
Сотрудники


Быкова Наталья Владимировна
младший научный сотрудник
тел. +7(383) 353-91-91
e-mail: BykovaNV@igg.sbras.ru


Гражданкин Дмитрий Владимирович
доктор геолого-минералогических наук
ведущий научный сотрудник
тел. +7(383) 333-33-29
тел. +89117871081
e-mail: GrazhdankinDV@igg.sbras.ru

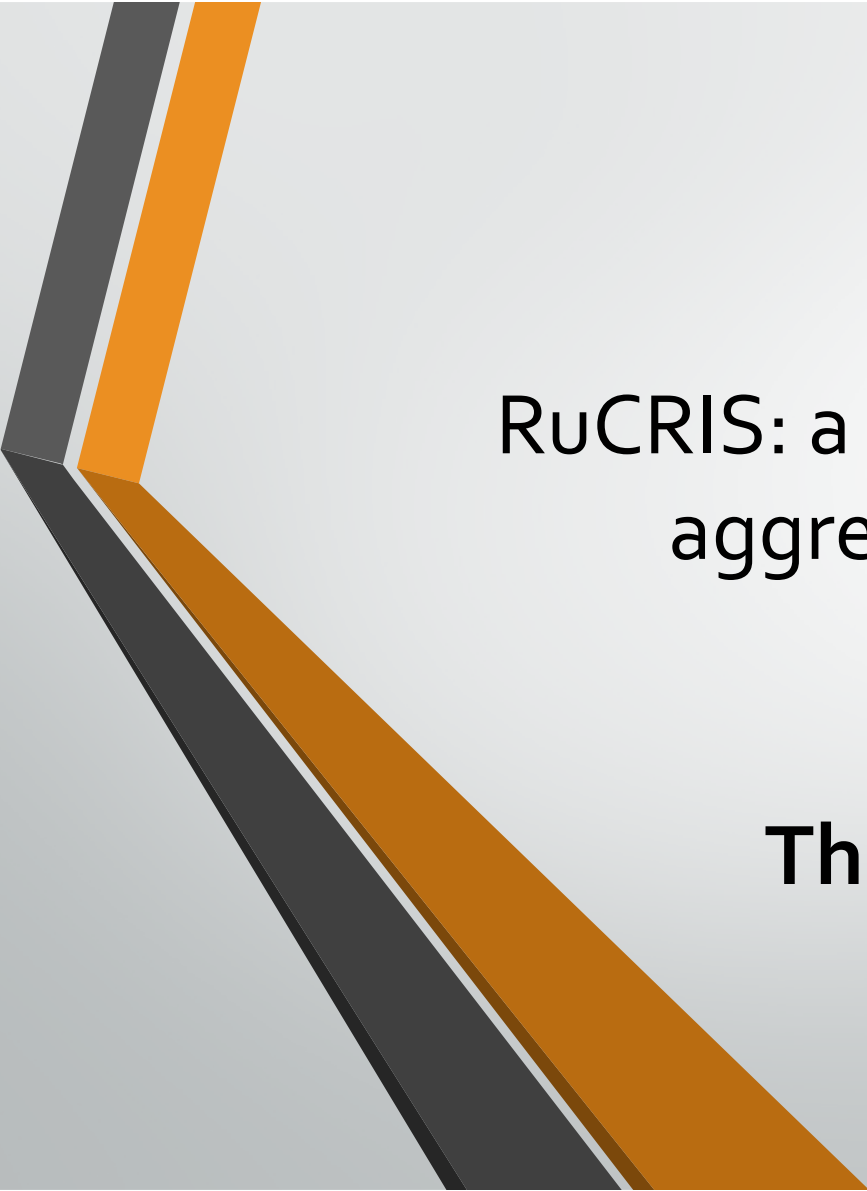
Publishing infrastructure

Based on Microsoft Sharepoint 2013 Cross-site Publishing features



Conclusion

- Mission completed: **pilot system works**
- Research open data in Russia: **still weak**
- Requirements for national CRIS:
 - global IDs
 - unified classification system
 - API for data access



RuCRIS: a pilot CERIF-based system to
aggregate heterogeneous data of
Russian research projects

Thank you for your attention!