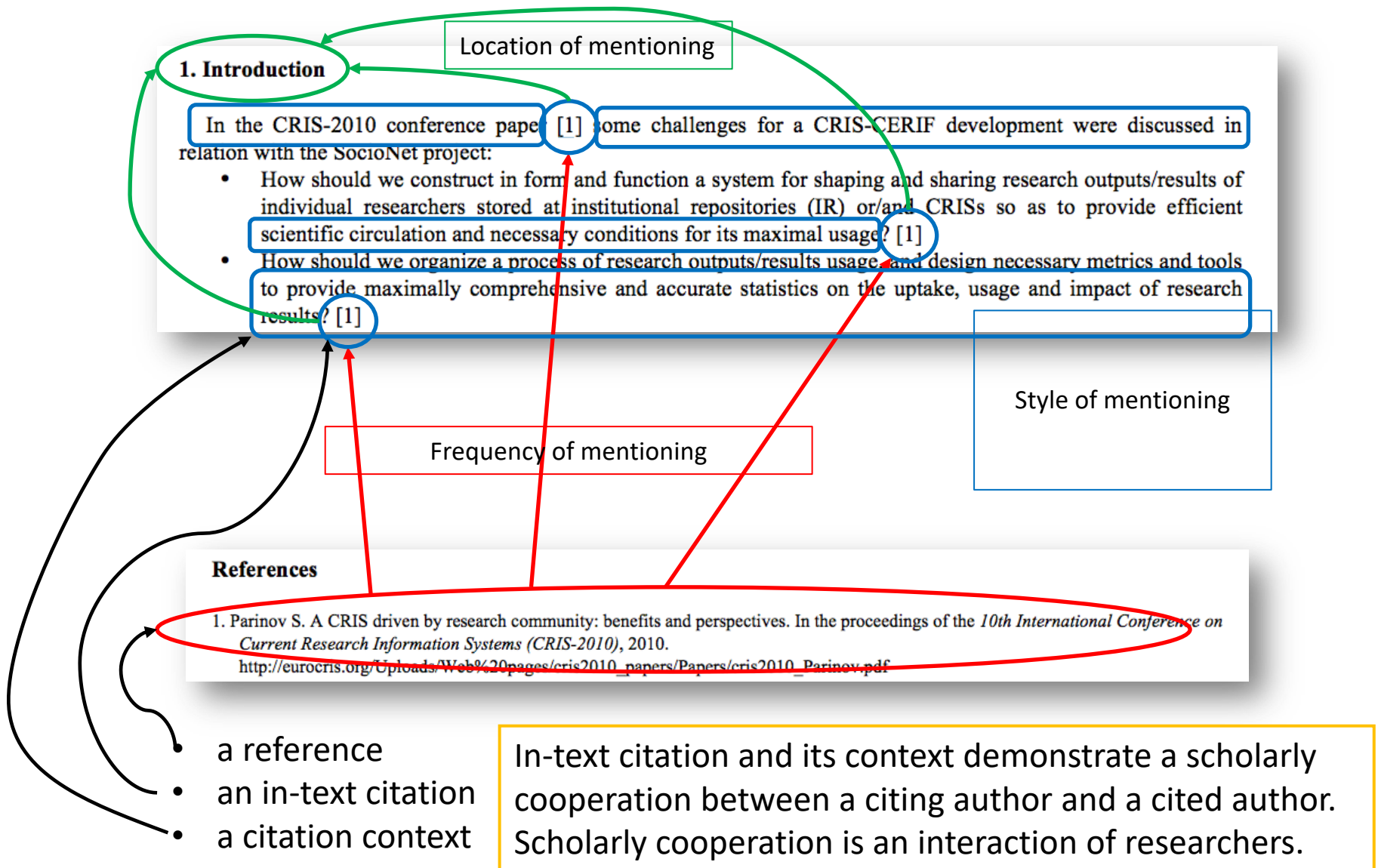


CRIS with in-text citations as interactive entities

Sergey Parinov

CEMI RAS and RANEPa

In-text citations as interactive elements, why?

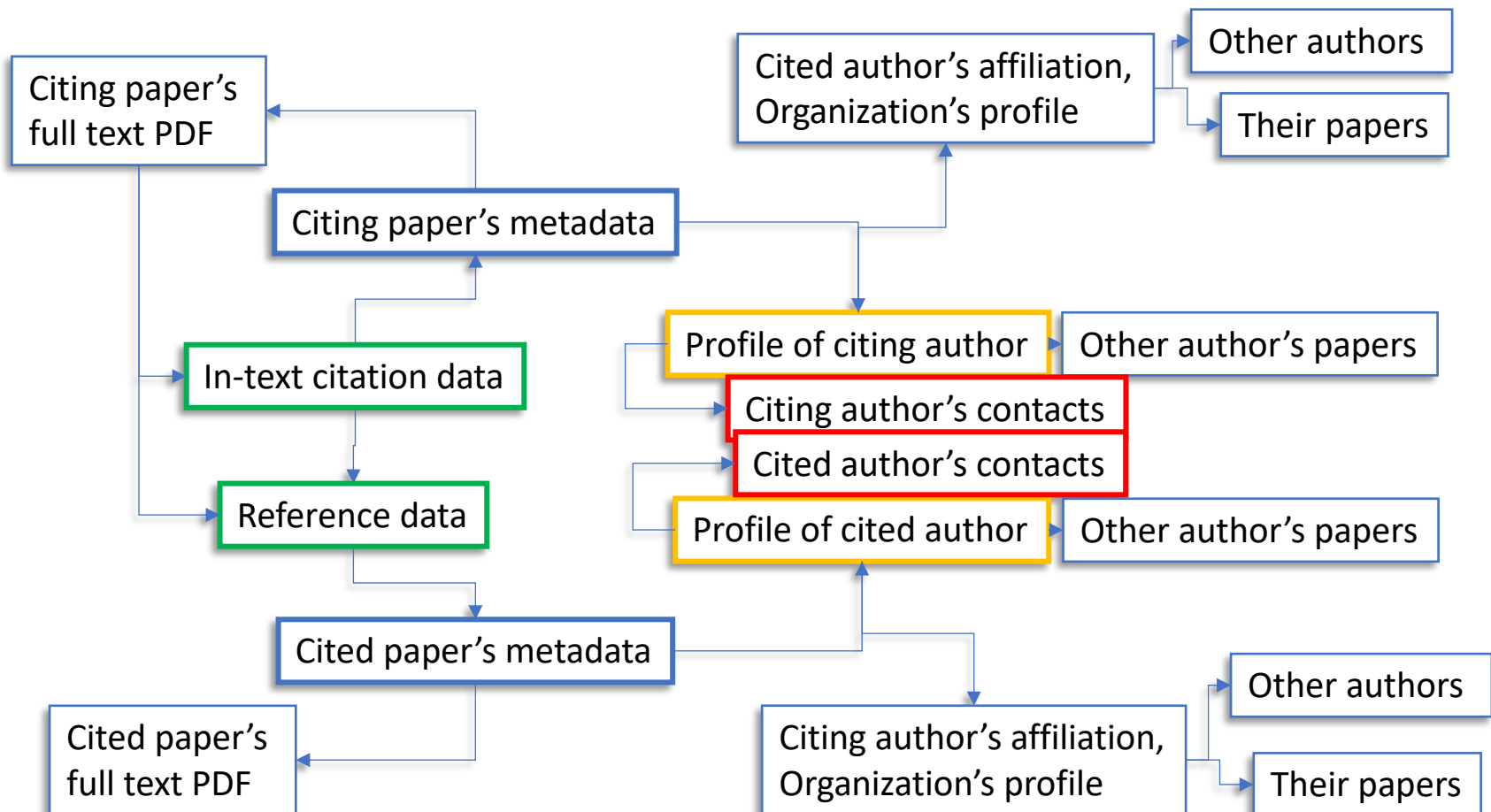


A challenge

- Let's transform the in-text citations into channels for scholarly communication and cooperation.
- Using these channels:
 - the cited authors know who used what of their outputs
 - the cited authors can inform the citing authors about upgrades with cited outputs
 - the citing authors can send requests to cited authors on needed development of cited outputs
- As a result,
 - the research community has wider, than now, scholarly cooperation
 - scholars have better individual research performance

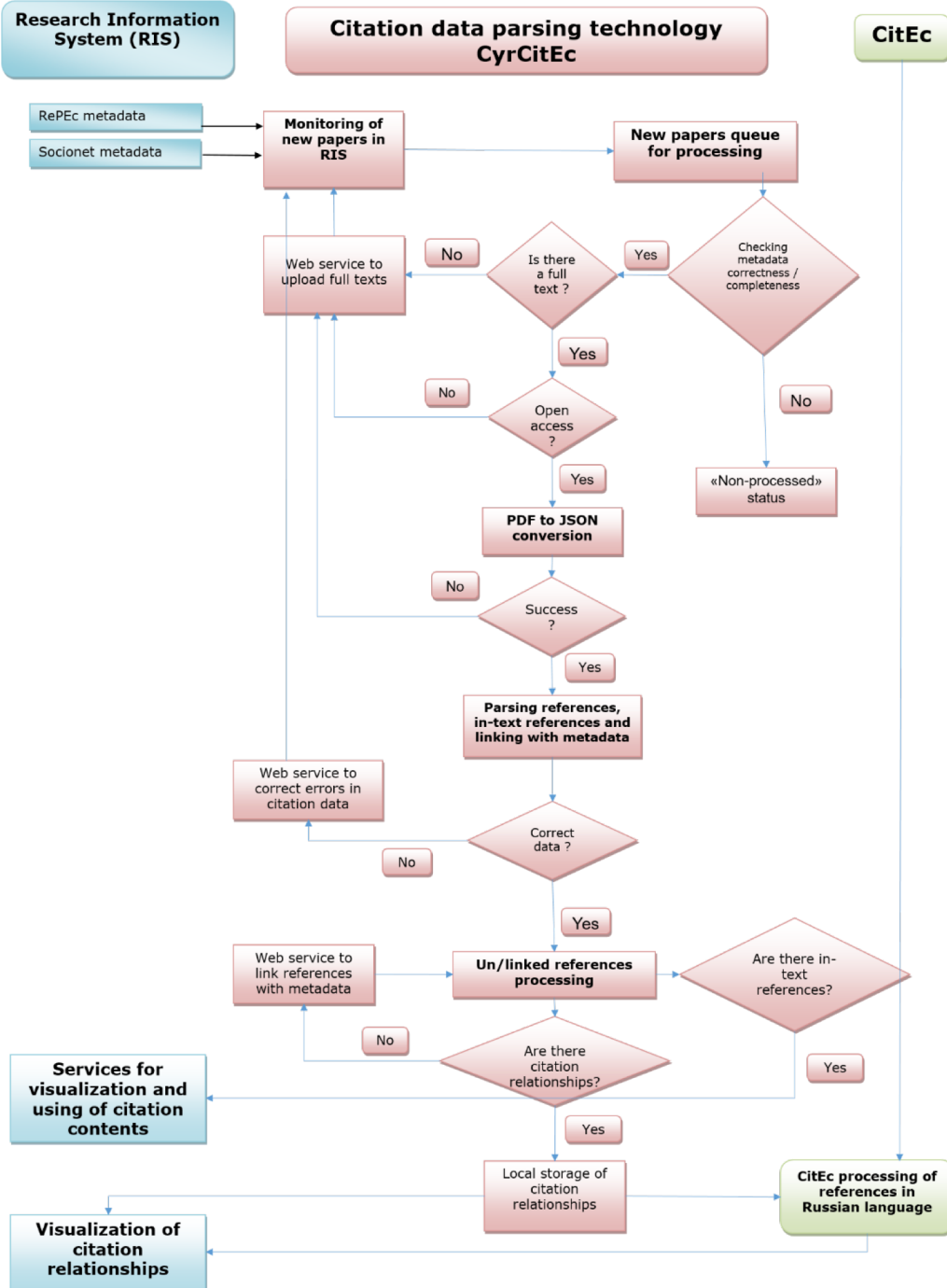
Why we need CRIS ?

If we integrate citation data into CRIS with a rich semantic layer, we can enrich the data by many additional attributes, like citing/cited authors contact data, etc.



Step 1. Citation data parsing by the CyrCitEc project

- Main current outputs of the project:
 - an open source software to parse papers' metadata and full text PDFs;
 - an open service to process papers' PDFs to extract citation data;
 - a dataset of citation data, including citation contexts (currently mostly for papers in Cyrillic);
 - a visualization tool providing for users insight into the citation data extraction process and a some control over results of the citation data parsing



CyrCitEc Technology:

- Takes papers from RePEc and Socionet
- Returns citation data to RePEc/Socionet
- Integrated by data with CitEc/RePEc
- Uses PDF.js+Node.js to convert PDF
- Creates citation data as XML files
- Provides open access to these data

Step 2. Open citation data source

Statistics on 01.06.2018

Column Name	Column Legend	Totals
series	list of processed collections of papers	220
[1]	metadata records available	100553
[2]	records with links to paper's full text	91929
[3]	records at Web ARChive (WARC)	86882
[4]	PDF files in Web ARChive	67048
[5]	PDF files converted to JSON	66802
[6]	JSON files with found reference sections	49129
[7]	JSON with in-text citations	45580
[a]	total references	865042
[b]	total citation contexts	735709
[c]	total mentioned references	806680
[d]	total citation relationships	47052
[e]	total non-mentioned references	129333

Source: <http://citru.repec.org/stats.html>

CyrCitEc: parsed citation data

```
▼<document type="citmap" provider="RANEPA, CitEcCyr project" updated="2018-05-21T07:15:00.000000+00:00"
  created="2018-03-23T00:16:15.000000+00:00">
  ▼<source handle="repec:hig:fsight:v:11:y:2017:i:4:p:84-95">
    ▼<futli url="https://foresight-journal.hse.ru/data/2018/01/14/1160539232/8-Kuzyk-84-95.pdf">
      ▼<version url="rsync://citru.repec.org/warc/RePEc/hig/fsight/v%3A11%3Ay%3A2017%3Ai%3A4%3Ap%3A84-
        95.warc" start="2501" length="258033">
```

```
▼<intextref>
```

```
  ▼<Prefix>
```

Still, in this regard Russia remains far behind not only the countries that traditionally adhere to the university-based R&D model, but even behind some of the former socialist countries and Soviet republics

```
  </Prefix>
```

```
  ▼<Suffix>
```

; Gokhberg, Kuznetsova, 2011]. Russian universities' R&D cooperation with businesses does not look very impressive either (Figure 2). However, the current Russian situation with universities' research and innovation activities is not at all unique.

```
  </Suffix>
```

```
  <Start>8757</Start>
```

```
  <End>8781</End>
```

```
  <Exact>[Gokhberg et al., 2009</Exact>
```

```
  <Reference>20</Reference>
```

```
</intextref>
```

```
▼<reference num="20" start="54464" end="54654" author="Gokhberg Kuznetsova Zaichenko "
  title="Towards a New Role of Universities in Russia Prospects and Limitations" year="2009"
  handle="repec:oup:scippl:v:36:y:2009:i:2:p:121-126">
```

```
  ▼<from_pdf>
```

Gokhberg L., Kuznetsova T., Zaichenko S. (2009) Towards a New Role of Universities in Russia: Prospects and Limitations. Science and Public Policy, vol. 36, no 2, pp. 121–126.

```
  </from_pdf>
```

```
</reference>
```

Step 3. Computer-generated annotations for the in-text citations

- PDF.js module to convert PDF to JSON
- Hypothes.is annotation tool withing Socionet
- forming citation data by the Web Annotation Data Model

```
{ "target": [
  { "source": "https://foresight-journal.hse.ru/data/...",
    "selector": [
      { "type": "TextPositionSelector", "end": 8781, "start": 8757 },
      { "exact": "[Gokhberg et al., 2009]" },
    ],
    "tags": [some tags],
    "text": "some information and statistics about cited reference"
  ]
}
```

Compared with 2000, the higher education sector's share (Figure 2). Still, in this regard Russia remains far behind not only the university-based R&D model, but even behind some other republics [Gokhberg et al., 2009; Gokhberg, Kuznetsova, 2009]. However, the current Russian situation with universities' research

[Gokhberg et al., 2009]

[20] -> Gokhberg Kuznetsova Zaichenko (2009) Towards a New Role of Universities in Russia Prospects and Limitations, citations in the paper -1, total of citations - 1

CyrCitEc Project, RANEPa, 2018-03-12, More...

source: <https://goo.gl/bZJwzZ>

Step 4. A feedback form to express opinion about citation meaning

- Different options depending on user's status:
 - user is the author of the citing paper and the particular browsed citation;
 - user is the author of the cited paper;
 - user is the co-author of the cited papers;
 - user is just a reader, he/she has no relation with the browsed citation and citing/cited papers

Annotating of and Relationship with the object: [What Prevents Universities from Creating Dual Technologies?](#)

Annotating in the PDF: <https://foresight-journal.hse.ru/data/2018/01/11/1234567890.pdf>
// **view PDF with annotations**

Annotation to the selected fragment: ...Still, in this regard Russia remains one of the countries that traditionally adhere to the university-based R&D model, but unlike socialist countries and Soviet republics **[[Gokhberg et al., 2009]** ; Gokhberg et al. (2009) universities' R&D cooperation with businesses does not look very impressive in the current Russian situation with universities' research and innovation activities.

Help

Comments

It is a place for your comments about this citation correctness.

Types of scientist reactions on professional evaluation

☒ disagree with this relationship

☐ agree with this relationship

relationship, semantic linkage for my paper

Save in collection

Public annotations/linkages collection

source: <https://goo.gl/7XrgnE>

Taxonomy of author's reactions

VALUES	FOR CITING AUTHORS	FOR CITED AUTHORS	FOR READERS
agree with this citation, comment	✓	✓	✓
disagree with this citation, comment	✓	✓	✓
ready to improve my paper		✓	
ready to help with taking better effect from using my paper		✓	
propose making a joint paper		✓	
propose a joint development of my results		✓	
misunderstanding of my paper		✓	
protest against style of this citation		✓	

Further development

- Visualization of scholarly cooperation network using the in-text citations and its contexts
 - Who cites whom by what topic with what polarity and/or a function
- Channeling of scholarly communications between citing and cited authors through cooperation network
- Global mechanism of scholarly collaboration based on the “small group” model