CAS STNext® COFFEE LECTURE

PREPRINTS IN CAPLUS

Sarah W. Stokes, FIZ Karlsruhe
What are Preprints?

- Pre-publication versions of articles are often released ahead of print
- Some (but not all) of these preprints are eventually published in scholarly journals
- The published document often differs from the preprint in terms of indexing, authors, etc., depending on editing that occurs prior to publication
- Preprint publications are often the earliest public mention of scientific discovery/invention
Preprint coverage in CAPLUS

- CAS acquires preprint data from a number of third-party preprint servers for inclusion in CAplus
- Over 910K preprint records currently (11/23) included in CAplus
- Major sources of coverage include ARXIV servers, Los Alamos National Laboratory, Beilstein, and open access preprint publications
- COVID-19 specific preprint servers included as well
Why include preprint searching?

- Preprints are often the first public mention of new technology, treatments, or inventions.

- Preprint publications sometimes are cited by research that is published prior to the publication of the eventual document, which can also give early data in IP searching. While the majority of citations will be to the publication, over 100K preprints in CPlus contain citation information for citing documents.

- Since these are early publications, they can be used to follow breakthroughs and early trends in new areas of research – this was shown to be true since the early pandemic with emerging treatments for COVID-19.

=> s preprint/dt and osc.g>=25

910559 PREPRINT/DT
5458469 OSC.G>=25
L1 10348 PREPRINT/DT AND OSC.G>=25
Example: Emerging treatment options

- Find preprint materials in Caplus covering the emerging findings that COVID-19 replication is inhibited by the SSRI fluoxetine – a new potential use for this existing drug.

- Find eventual publications in scholarly journals of these documents.

- Compare and contrast entry date to show earlier availability of preprint materials, as well as indexing and other published data to see differences between preprint and eventual publication.
Broad keyword search combined with /DT

=> S (?FLUOXETIN? (10A) (?SARS? OR ?COVID?)) AND PREPRINT/DT

17181 ?FLUOXETIN?
167683 ?SARS?
218090 ?COVID?
47 ?FLUOXETIN? (10A) (?SARS? OR ?COVID?)
913707 PREPRINT/DT
L1 9 (?FLUOXETIN? (10A) (?SARS? OR ?COVID?)) AND PREPRINT/DT
Preprint search yields highly relevant research

Targeting the endolysosomal host-\textit{SARS}-\textit{CoV-2} interface by the clinically licensed antidepressant \textit{fluoxetine}
Search CAPlus for related publications

=> S (?FLUOXETIN? (10A) (?SARS? OR ?COVID?)) NOT PREPRINT/D

17181  ?FLUOXETIN?
167683  ?SARS?
218090  ?COVID?
  47  ?FLUOXETIN? (10A) (?SARS? OR ?COVID?))
913707  PREPRINT/D
L2  38 (?FLUOXETIN? (10A) (?SARS? OR ?COVID?)) NOT PREPRINT/D
# Compare preprints with publications

<table>
<thead>
<tr>
<th>Title</th>
<th>Accession Number Preprint</th>
<th>Accession Number Pub</th>
<th>Author/Inventor Preprint</th>
<th>Author/Inventor Publication</th>
<th>PA/CS Preprint</th>
<th>PA/CS Pub</th>
<th>Publisher Preprint</th>
<th>Publisher Publication</th>
<th>Source Preprint</th>
<th>Source Publication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antidepressant and antipsychotic drugs reduce viral infection by SARS-CoV-2 and fluoxetine show antiviral activity against the novel variants in vitro</td>
<td>2021:731212 HCAPLUS</td>
<td>2022:1214608 HCAPLUS</td>
<td>Fred, Senem Merve; Kuivanen, Suvi; Ugurlu, Hasan; Casarotto, Plinio Cabrera; Levanov, Lev; Sakse1a, Kalle; Vapalahti, Olli; Castrén, Eero</td>
<td>Fred, Senem Merve; Kuivanen, Suvi; Ugurlu, Hasan; Casarotto, Plinio Cabrera; Levanov, Lev; Sakse1a, Kalle; Vapalahti, Olli; Castrén, Eero</td>
<td>Neuroscience Center - HILIFE, University of Helsinki, Finland</td>
<td>Neuroscience Center-HILIFE, University of Helsinki, Finland</td>
<td>Cold Spring Harbor Laboratory Press</td>
<td>Frontiers Media S.A.</td>
<td>Frontiers in Pharmacology (2021), 12, 755600</td>
<td>CODEN: BIORFP; CODEN: FPRHAU; ISSN: 1663-9812</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ED Preprint</th>
<th>ED Pub</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>30-Mar-21</td>
<td>10-May-22</td>
<td>406</td>
</tr>
<tr>
<td>5-Jan-21</td>
<td>13-Feb-22</td>
<td>404</td>
</tr>
<tr>
<td>26-Oct-20</td>
<td>7-Apr-21</td>
<td>163</td>
</tr>
<tr>
<td>5-Aug-20</td>
<td>29-Dec-20</td>
<td>146</td>
</tr>
<tr>
<td>23-Jun-20</td>
<td>22-Mar-21</td>
<td>272</td>
</tr>
</tbody>
</table>

average: **278.2**
Example: Keeping up with trends

- In the early pandemic, preprints were very useful to researchers trying to find existing drugs that might have activity against COVID-19, but many were quickly found to be ineffective.

- Analysis of preprint data for this type of substance can give us an idea of whether a potential treatment was later found to have any efficacy.

  - The publication patterns of preprint data can shed insight into how viable a candidate a potential treatment turned out to be.
Analyze preprint results for substances by year to see trends

=> S "Hydroxychloroquine" AND PREPRINT/DT AND PY>=2020

8861 "HYDROXYCHLOROQUINE"
3 "HYDROXYCHLOROQUINE"

=> S "Ivermectin" AND PREPRINT/DT AND PY>=2020

7863 "IVERMECTIN"
26 "IVERMECTINS"
7865 "IVERMECTIN"

("IVERMECTIN" OR "IVERMECTINS")

913707 PREPRINT/DT
11599963 PY>=2020

913707 PREPRINT/DT
11599963 PY>=2020

L3 407 "HYDROXYCHLOROQUINE"

L5 105 "IVERMECTIN" AND PREPRINT/DT AND PY>=2020

=> ANA 1- PY; D 1-

L4 ANALYZE L5 1- PY : 4 TERMS
Compare Entry Data to check viability

<table>
<thead>
<tr>
<th>TERM</th>
<th># OCC</th>
<th># DOC</th>
<th>% DOC</th>
<th>PY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>320</td>
<td>320</td>
<td>78.62</td>
<td>2020</td>
</tr>
<tr>
<td>2</td>
<td>63</td>
<td>63</td>
<td>15.48</td>
<td>2021</td>
</tr>
<tr>
<td>3</td>
<td>21</td>
<td>21</td>
<td>5.16</td>
<td>2022</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>3</td>
<td>0.74</td>
<td>2023</td>
</tr>
</tbody>
</table>

********** END OF L4 **********

<table>
<thead>
<tr>
<th>TERM</th>
<th># OCC</th>
<th># DOC</th>
<th>% DOC</th>
<th>PY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>40</td>
<td>40</td>
<td>38.10</td>
<td>2020</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>35</td>
<td>33.33</td>
<td>2021</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
<td>23</td>
<td>21.90</td>
<td>2022</td>
</tr>
<tr>
<td>4</td>
<td>7</td>
<td>7</td>
<td>6.67</td>
<td>2023</td>
</tr>
</tbody>
</table>

********** END OF L6 **********

Hydroxychloroquine

Ivermectin
Emerging (eventually debunked) treatments by year, in preprint publication numbers
Summary

- Preprints can be the earliest available mention of a new technology or substance, which can be crucial for IP searching.
  - Sometimes the preprint is available over a year in advance of the publication!
- Analysis of preprint publications trends gives insight into whether a promising potential invention resulted in publishable results.
- Preprints add even more value to the timely coverage of documents in CAplus.
HELP in CAS STNext

STNext Quick Start

1. Displays a sequential auto-generated transcript name. Hover over the area to access the:
   - **Pencil** icon - Click to edit the transcript name.
   - **Pause/On** button - Click to pause/resume transcript recording.
   - **Download** icon - Click to download the current transcript as a .pdf, .rtf, .txt, or .zip file.
   - **Start New** button - Click to start a new transcript for the session.

2. Use **My Files** to access:
   - **Alerts** - View and edit SDI query alerts.
   - **Transcripts** - Download a transcript, create a report, or record a search.

Download the **STNext Quick Start** reference.
You can also use **Help** to navigate the web interface.
Between problems and progress are connections that matter

Sarah W. Stokes
Senior Product Specialist, IP
sstokes@cas.org

CONTACT

CAS
help@cas.org
cas.org

FIZ Karlsruhe
EMEAHelp@cas.org
stn-international.de