

Информационные ТЕХНОЛОГИИ В ХИМИИ

А. А. Рагойша

Лекция 2

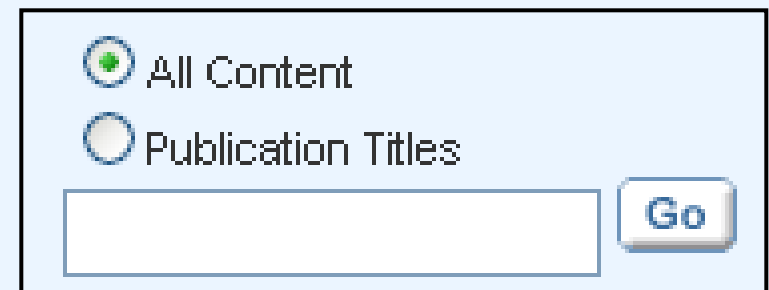
Поисковый бланк и список результатов поиска

Поисковый бланк - 1

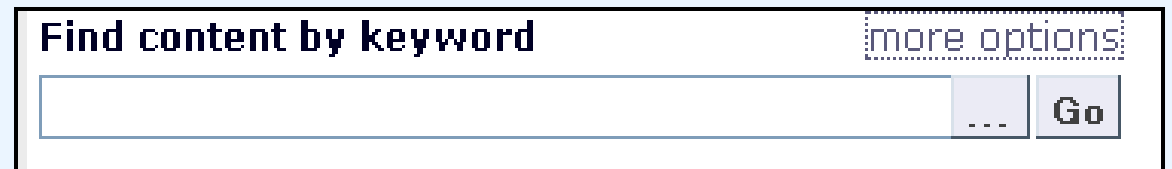
- Простейший, Basic, Quick, Simple
- Расширенный, Advanced, Expert

Классификация условна – в рамках определенной базы данных.

Обычно такие бланки называют **Quick Search**



A search form with two radio buttons: "All Content" (selected) and "Publication Titles". Below the radio buttons is a search input field and a "Go" button.



A search form titled "Find content by keyword" with a search input field, a "more options" link, and a "Go" button.

Поисковый бланк - 2

Пример бланка

Quick Search:

Поиск по двум разным полям; использование булевых операторов; выбор временного интервала.

Query [Help]			
Term 1:	<input type="text"/>	in Field 1:	Title <input type="button" value="v"/>
		AND	<input type="button" value="v"/>
Term 2:	<input type="text"/>	in Field 2:	Inventor Name <input type="button" value="v"/>
Select years [Help]			
	<input type="text" value="1976 to present [full-text]"/>	<input type="button" value="v"/>	<input type="button" value="Search"/> <input type="button" value="Reset"/>

Поисковый бланк - 3

Пример бланка

Advanced Search:

All Sources	Journals	Books	Advanced Search
Term(s):	<input type="text"/>	within:	<input type="text" value="Title"/>
AND	<input type="text"/>	within:	<input type="text" value="Authors"/>
Include:	<input checked="" type="checkbox"/> Journals	<input checked="" type="checkbox"/> All Books	
Source:	<input type="text" value="All sources"/>		
	Select one or more:		
Subject:	<input type="text" value="- All Sciences -"/> Agricultural and Biological Sciences Arts and Humanities Biochemistry, Genetics and Molecular Biology		Hold down the key) to select m
Dates:	<input checked="" type="radio"/> All Years	<input type="radio"/> 1999	to: Present
Search Clear Recall Search			

Элементы бланка

The screenshot shows a search interface with the following elements:

- Navigation tabs: **All Sources** (selected), **Journals**, **Books**.
- Term(s):** An empty text input field.
- AND** (operator) with a dropdown arrow.
- Include:** Two checked checkboxes for **Journals** and **All Books**.
- Source:** A dropdown menu currently showing **All sources**.
- Subject:** A dropdown menu with the text "Select one or more:" above it. The current selection is **- All Sciences -**. Other visible options include **Agricultural and Biological Sci**, **Arts and Humanities**, and **Biochemistry, Genetics and M**.
- Dates:** Radio buttons for **All Years** (selected) and **1999** (with a dropdown arrow), followed by the text "to:".
- Buttons at the bottom: **Search**, **Clear**, and **Recall Sea**.

- Графа бланка (редактируемая графа, редактируемое поле).
- Список.
- Меню (выпадающий список).
- Переключатель.
- Выключатель.

- Текстовые пояснения.
- Ссылка на иной бланк.
- Ссылка на блок инструкций.
- Кнопка начала поиска.

Список результатов поиска

Пользователь получает результаты поиска в форме списка обнаруженных документов.

Список может быть сформирован:

- по алфавиту,
- в хронологическом порядке
(прямом или обратном),
- по релевантности.

Релевантность

Релевантность документа –

степень соответствия его поисковому заданию.

При расчете релевантности учитываются:

- количество поисковых терминов в документе,
- расстояние между ними в тексте,
- число упоминаний каждого из них,
- их плотность,
- их порядок расположения,
- их место – в начале записи или в конце,
- и др.

Особенности ПОИСКОВЫХ СИСТЕМ

Особенности работы

Предназначены для широкой публики

Упростить взаимодействие с системой

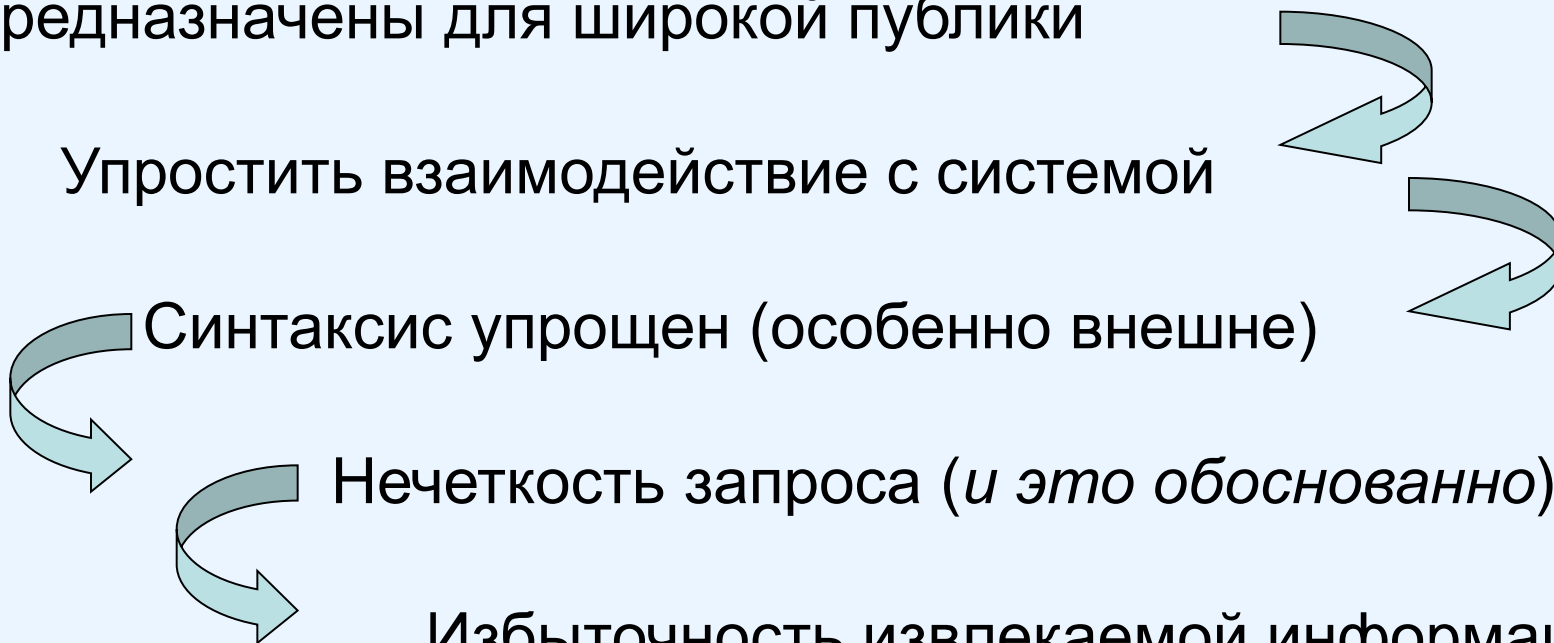
Синтаксис упрощен (особенно внешне)

Нечеткость запроса (*и это обоснованно*)

Избыточность извлекаемой информации

Огромный список результатов поиска

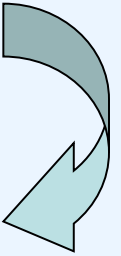
Главное – не величина списка, а **качество сортировки**



Релевантность - 2

Расчет **ранга** веб-страницы – т.е. ее авторитетности

Сколько ссылок с других страниц на данную?
Каков ранг ссылающихся страниц?



Расчет **релевантности**:

+ Вклад 1: поисковые термины на странице
+ Вклад 2: ранг страницы

= Релевантность страницы в списке результатов поиска

Бланки

- Простейший – Усложненный
Простейший – одна графа



A search bar with a yellow border and a 'Найти' button.

- На простейшем бланке:
предполагается, что запрос –
перечень слов, разделенных пробелами
- Оператор по умолчанию:
AND – чаще всего
OR – редко

Логические операторы

- Операторы отображать **ЗАГЛАВНЫМИ** буквами
(если строчными – стоп-слово)

- **AND, + , &**

натрий AND калий
натрий & калий
натрий +калий

Обычно достаточно:
натрий калий

- **OR, | , ...**

натрий **OR** калий
натрий | калий

- **NOT, -**

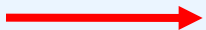
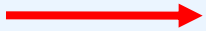
натрий NOT калий
натрий **-калий**

без
пробела!

Регистр букв

- Нечувствительны к регистру - большинство
- Частичная чувствительность – иногда

Пример – Яндекс

- при запросе  **азот**
извлекаются **все** документы
- при запросе  **!Азот**
извлекаются документы со словом *азот*, начинающимся с **Заглавной** буквы

Шаблон

Как правило, нет. Но:

Google, Яндекс

Звездочка * заменяет собой любое **слово**.
Используют внутри кавычек, *например*:

"особенно полезен * чай"

Поиск по словоформам

Google – в основном, нет.

Яндекс – да.

Поиск по полям документов


Возможен, но полезность ограниченная
основная часть информации на веб-странице
находится в одном самом большом поле

Google: поиск по полям - 1

- **intitle:БГУ**
- **intitle:БГУ химфак**
- **intitle:БГУ intitle:химфак**
allintitle:БГУ химфак

intitle:БГУ химфак

About 6,990 results (0.31 seconds)


[Химический факультет БГУ - Главная страница](#)  - [www.chemistry.bsu.by/](#) - *Cached*

Информация о факультете: деканат, кафедры, лаборатории, ко программы.

Абитуриенту	Кафедра не
www.chemistry.bsu.by/.../index.asp?...	www.chemis.../index.asp?
В день этого праздника химфак	электронна

allintitle:БГУ химфак

About 989 results (0.06 seconds)

[БГУ ХИМФАК | Химия и Химики | БГУ ХИМФАК](#)  - [[Translate this page](#)]

[www.himiki.com/bgu-ximfak](#) - *Cached*

Белорусский государственный университет Химический факультет Факультет.

Google: поиск по полям - 2

- **site:bsu.by**
- **site:bsu.by кислород**

[Операция "Кислород" - Энциклопедия кораблей](#) 🔍 - [Translate this page]
ship.bsu.by/text.aspx?guid=4024 - Cached

Операция "Кислород" проведена французскими боевыми пловцами 10 июля 1985 года, в результате которой в новозеландском порту Окленд судно ...

[Открытая студенческая олимпиада 'Кислород-2010'](#) 🔍 - [Translate this page]
www.abc.chemistry.bsu.by/oxygen/default.htm - Cached

19-20 июня 2010 г. на химическом факультете БГУ прошла Первая открытая студенческая олимпиада по химии "Кислород-2010". В состязаниях ...

Google: поиск по полям - 3

- **filetype:pdf** **ext:pdf**

filetype:pdf intitle:БГУ химический факультет

filetype:pdf intitle:БГУ химический факультет|

6 results (0.26 seconds)

[PDF] [ПРОГРАММА ФЕСТИВАЛЯ ФАКУЛЬТЕТОВ БГУ](#) Место проведения: В

- [[Translate this page](#)]

www.ffsn.bsu.by/ffsn.files/news/prog_fest_2011.pdf

File Format: PDF/Adobe Acrobat - [Quick View](#)

16.00-16.30 **Химический факультет**. 16.30-17.00 Биологический **факультет**. 17.00-17.30

Факультет доуниверситетского образования. 17.30-18.00 Военный **факультет** ...

Google: поиск по полям - 4

- **inurl:bsu.by**

[Белорусский государственный университет](#)
[www.bsu.by/](#) - Cached
Ректорат Главные управления Управления Самостояте
институты Общеуниверситетские ...

inurl:bsu inurl:chemistry

[BSU, Faculty of Chemistry](#) 🔍

[wikimapia.org/3360176/BSU-Faculty-of-Chemistry](#) - Cache
World / Belarus / Minsk / Minsk , 2km from center (Минск)
27°32'57"E. BSU, Faculty of Chemistry (Minsk) ...

[www.Chemistry.bsu.by](#) 🔍

[www.markosweb.com/www/chemistry.bsu.by/](#) - Cached

Усложненный бланк (бланк расширенного поиска)

+

- Несколько граф
- Выключатели, переключатели и т. д. для назначения параметров поиска
- Наглядность при формулировании заданий средней и выше чем средней сложности

-

- На усложненном бланке можно реализовать не все синтаксические возможности данной поисковой программы

Advanced Search

Find pages with...

all these words:

this exact word or phrase:

any of these words:

none of these words:

numbers ranging from:

to

Then narrow your results by...

language:

English

Find pages in the language you select.

region:

Belarus

Find pages published in a particular region.

last update:

past 24 hours

Find pages updated within the time you specify.

site or domain:

Search one site (like `wikipedia.org`) or limit your results to a domain like `.edu`, `.org` or `.gov`

terms appearing:

in the title of the page

Источники научной информации

Бесплатная информация – почему?

Коммерческие издательства - как часть **рекламной** стратегии.

moving wall, embargo period

Некоммерческие организации, **финансируемые государством и фондами**, - уже все оплачено.

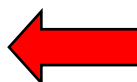
Частные лица – в порядке собственной инициативы, как хобби или для саморекламы.

Проявление результатов инициативы **Open Access**:
чтение, копирование – **без ограничений при условии** четкой ссылки на первоисточник.

- весь журнал *или* отдельные статьи
- платит автор *или* его организация (фонд)

Пример: пометка в оглавлении

[html](#) [pdf](#) [open access](#)



Acta Cryst. (2009). **A65**, 81-108 [doi:10.1107/S0108767308040592]

Three-dimensional Euclidean nets from two-dimensional hyperbolic tilings: kaleidoscopic examples

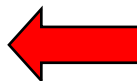
S. J. Ramsden, V. Robins and S. T. Hyde

Synopsis: Systematic enumeration of tilings of the two-dimensional hyperbolic plane leads to *ab initio* construction of triply periodic nets *via* projection of the tilings onto triply periodic minimal surfaces. The technique is described in detail for kaleidoscopic tilings, projected onto the *P*, *D* and *G* (gyroid) minimal surfaces.

Online 28 January 2009

research papers

[html](#) [pdf](#) [buy](#)



Acta Cryst. (2009). **A65**, 109-119 [doi:10.1107/S010876730804186X]

Репозиторий

Репозиторий (Repository) –

хранилище, архив электронных документов.

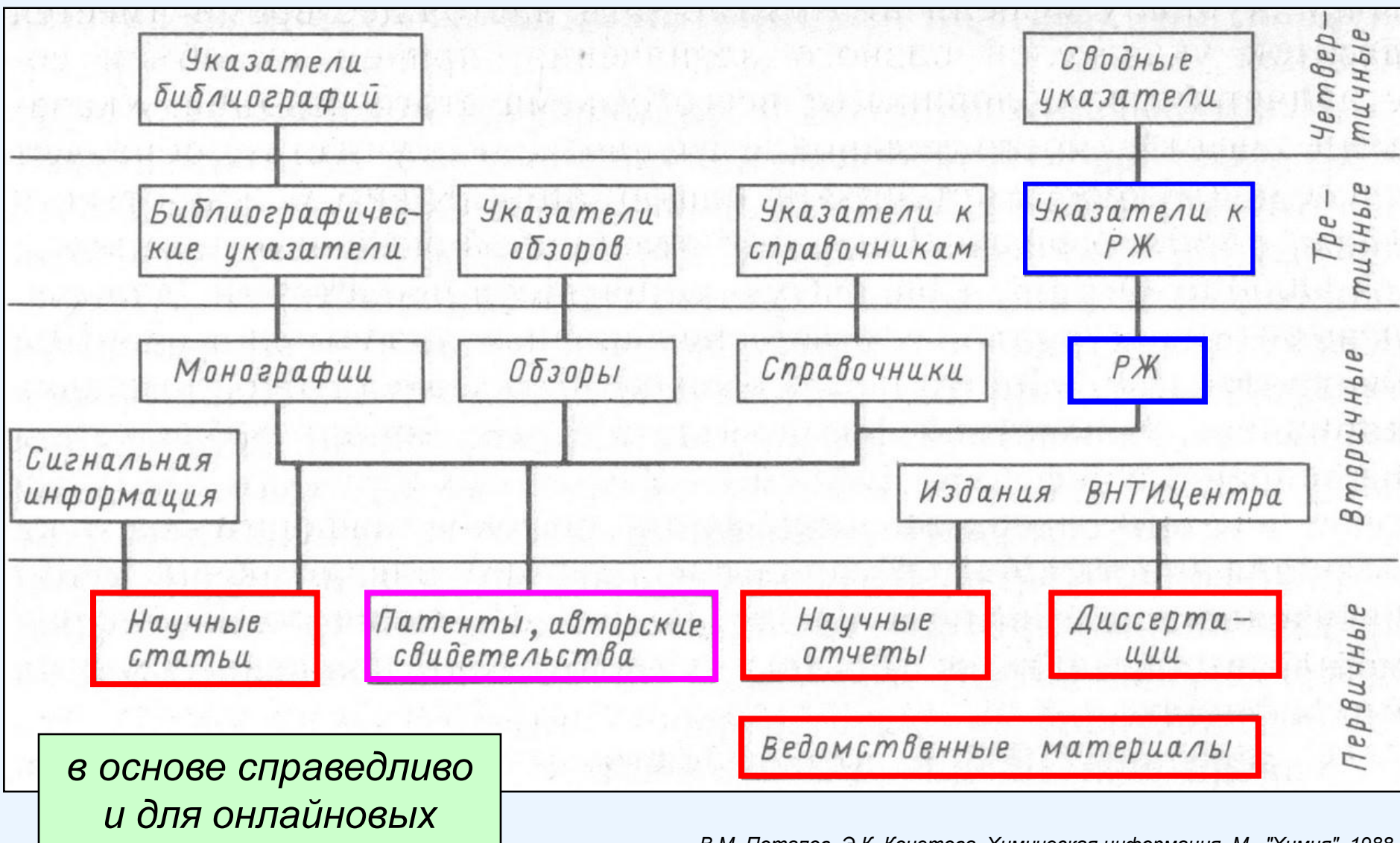
Репозиторий учреждения (IR – Institutional Repository) –

онлайновый архив, как правило открытый для всех, в котором накапливаются результаты деятельности данного учреждения.

В *университетском репозитории*: научные статьи, учебные материалы, административные документы.

Самоархивирование в репозиториях – новая тенденция.

Классификация печатных источников (РФ)



Первичные источники научной информации

Научная статья

Место публикации:

научный журнал, сборник статей, ...

Существенно:

оценка качества работы -
рецензирование

Материалы конференций

Конференция: информация + оценка качества работы

Тезисы доклада

Abstract

Расширенные тезисы доклада

Extended Abstract

Презентация. Текст доклада

Материалы конференции

Proceedings

где?

Сайт конференции (существует недолго)

как найти? по названию конференции  Google

метасайты конференций

Репозитории

(и еще: Платные базы данных)

Диссертации

Ph. D. Dissertation, Doctoral Thesis. Master's Thesis.

Автореферат диссертации.

где? Репозитории. Архивы диссертаций.

▣ Докторских диссертаций (Д)

01.00.00 – физико-математические науки ⁽⁰⁾, 02.00.00 – химические науки ⁽¹⁾, 03.00.00 – биологические науки ⁽⁰⁾, 05.00.00 – технические науки ⁽⁰⁾, ...

▣ Кандидатских диссертаций (К)

01.00.00 – физико-математические науки ⁽⁵⁾ **НОВОЕ**, 02.00.00 – химические науки ⁽⁵⁾ **НОВОЕ**, 03.00.00 – биологические науки ⁽⁴⁾,

*База данных
ВАК Беларуси*

Научные отчеты

(Research) Report

где? Репозитории. (Архивы отчетов).

КНИГИ (как правило, вторичный источник)

Проблема **копирайта** (охраны авторских прав).

где? Архивы и библиотеки (копирайт *нарушен?*).

Форматы: **PDF, DJVU**, (RAR, ZIP), (TXT, EPUB, ...)

как искать?

Специальные средства – по автору, названию.

Через **Google**: **Фрагмент_названия filetype:djvu**

Автор filetype:pdf

Google Books

И еще о терминах

Периодическое издание: научный журнал.
Продолжающееся издание: "Сборник", "Труды",
"Ученые записки", "Бюллетень"

Депонированная рукопись.

Grey (Gray) Literature –
научная, техническая, учебная информация,
распространяемая не через каналы традиционных
коммерческих издательств.

Традиционно: литература временного пользования.

Теперь: в репозиториях.

Терминология: ...print

Препринт (preprint) – документ, имеющий структуру научной статьи, но не прошедший рецензирование.

Одна из **стадий** статьи перед опубликованием *или*

Документ с **неясной** перспективой публикации

Постпринт (postprint) – научный документ, прошедший рецензирование.

Репринт (reprint) – повторно опубликованный документ.

e-print, eprint – электронная форма научного или учебного документа (препринт и постпринт; статья, тезис доклада, презентация, отчет, диссертация и т.д.).

Открытое рецензирование

Рецензирование: и достоинства, и недостатки.

Поиск новых форм. *Пример:*

Atmospheric Chemistry and Physics

ACPD - Volume 9, Number 4, pp. 14263-17849

Est modus in rebus: analytical properties of multi-model ensembles

S. Potempski and S. Galmarini

Page(s) 14263-14314

▣ [Abstract](#) ▣ [Discussion Paper \(PDF, 978 KB\)](#) ▣ [Interactive Discussion](#)
(Open, 2 Comments)

Mixing processes and exchanges in the tropical and the subtropical UT/LS

R. James and B. Legras

Page(s) 25-38

▣ [Abstract](#) ▣ [Final Revised Paper \(PDF, 1156 KB\)](#) ▣ [Discussion Paper \(ACPD\)](#)