

Google Custom Search Engine as a simple tool for multiple site searching

Irakli Garibashvili

Director, National Scientific Library in Georgia

Plan of a webinar

- ▶ Google Custom Search Engine (CSE) – technology insight
 - ▶ How to create simple engine in 2 steps
 - ▶ Multiple site search
 - ▶ Searching through digital libraries (OPACs, Repositories, eJournals, etc.)
 - ▶ OpenDOAR CSE
- 

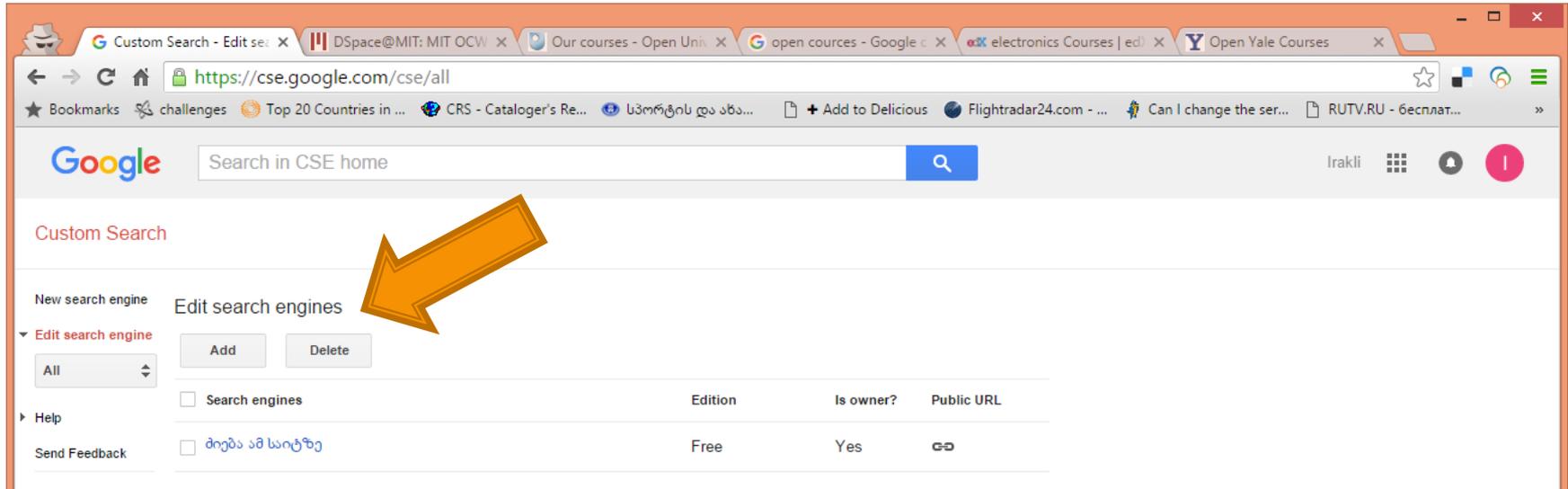
Goal

- ▶ To learn how to construct a simple search engine which will help to find information through
 - A specific site
 - Multiple sites
- ▶ And how to use the created CSE on your library website

Starting a CSE

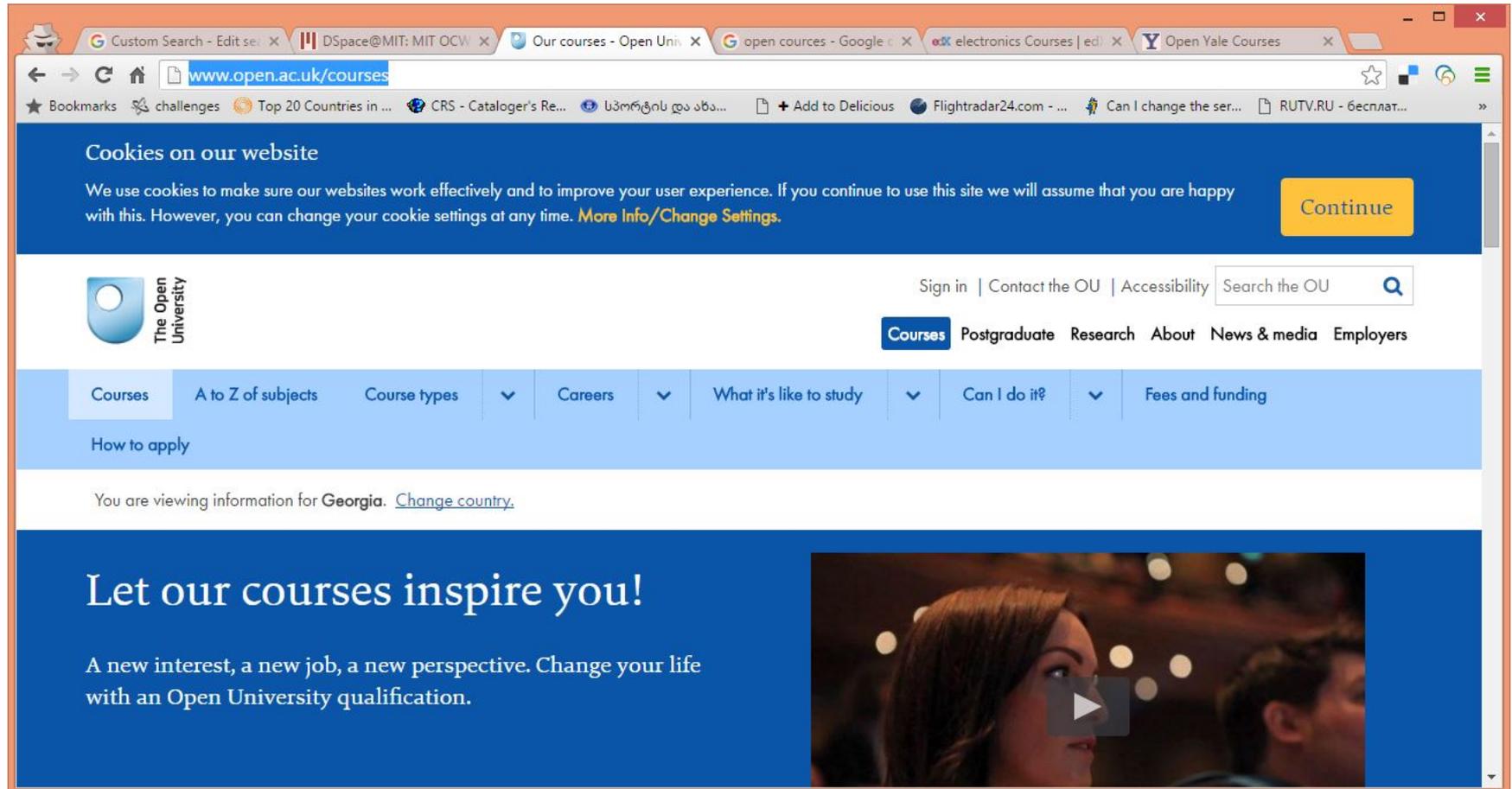
- ▶ Go directly to a <https://cse.google.com/cse/>
- ▶ OR search for “Google Custom Search”
- ▶ Login freely using your Google account

Starting page



Click on Add button to create a first sample engine

Searching for an open course at Open University UK



The screenshot shows a web browser window with the URL www.open.ac.uk/courses. The browser's address bar and tabs are visible at the top. A blue banner at the top of the page contains a cookie notice: "Cookies on our website. We use cookies to make sure our websites work effectively and to improve your user experience. If you continue to use this site we will assume that you are happy with this. However, you can change your cookie settings at any time. [More Info/Change Settings](#)." A yellow "Continue" button is on the right.

Below the banner is the Open University logo on the left and navigation links: "Sign in | Contact the OU | Accessibility | Search the OU" with a search icon. A menu bar includes "Courses" (highlighted), "Postgraduate", "Research", "About", "News & media", and "Employers".

A secondary navigation bar contains: "Courses", "A to Z of subjects", "Course types" (with a dropdown arrow), "Careers" (with a dropdown arrow), "What it's like to study" (with a dropdown arrow), "Can I do it?" (with a dropdown arrow), and "Fees and funding".

Below this is a section titled "How to apply".

A message states: "You are viewing information for **Georgia**. [Change country](#)."

The main content area features a blue banner with the text: "Let our courses inspire you!" and "A new interest, a new job, a new perspective. Change your life with an Open University qualification." To the right of the text is a video player showing a woman's face in profile, with a play button overlay.

Specify a URL of a site to search

Custom Search - Create x DSpace@MIT: MIT OCW x Our courses - Open Uni x open courses - Google c x electronics Courses | ed x Open Yale Courses x

https://cse.google.com/cse/create/new

Search in CSE home

Custom Search

New search engine
Enter the site name and click "Create" to create a search engine for your site. [Learn more](#)

▶ Edit search engine

▼ Help

- Help Center
- Help forum
- Support
- Blog
- Documentation
- Terms of Service
- Send Feedback

Sites to search

www.example.com

www.example.com

You can add any of the following:

- Individual pages: `www.example.com/page.html`
- Entire site: `www.mysite.com/*`
- Parts of site: `www.example.com/docs/*` or `www.example.com/docs/`
- Entire domain: `*.example.com`

If you want to search pages over entire web containing specific schema.org markups, click on "advanced" below.

Language

English

Think about a proper pattern of URL – `http://www.open.ac.uk/courses`

Custom Search

New search engine

Enter the site name and click "Create" to create a search engine for your site. [Learn more](#)

Sites to search

You can add any of the following:

- Individual pages: `www.example.com/page.html`
- Entire site: `www.mysite.com/*`
- Parts of site: `www.example.com/docs/*` or `www.example.com/docs/`
- Entire domain: `*.example.com`

If you want to search pages over entire web containing specific schema.org markups, click on "advanced" below.

Language

English

Indicate a Language and Title of a CSE and just click “Create”

Custom Search - Create x DSpace@MIT: MIT OCW x Our courses - Open Uni x open courses - Google c x electronics Courses | ed x Open Yale Courses x

https://cse.google.com/cse/create/new

Help Center
Help forum
Support
Blog
Documentation
Terms of Service
Send Feedback

http://www.open.ac.uk/courses

www.example.com

You can add any of the following:
Individual pages: www.example.com/page.html
Entire site: www.mysite.com/*
Parts of site: www.example.com/docs/* or www.example.com/docs/
Entire domain: *.example.com

If you want to search pages over entire web containing specific schema.org markups, click on "advanced" below.

Language

English

Name of the search engine

My first search engine

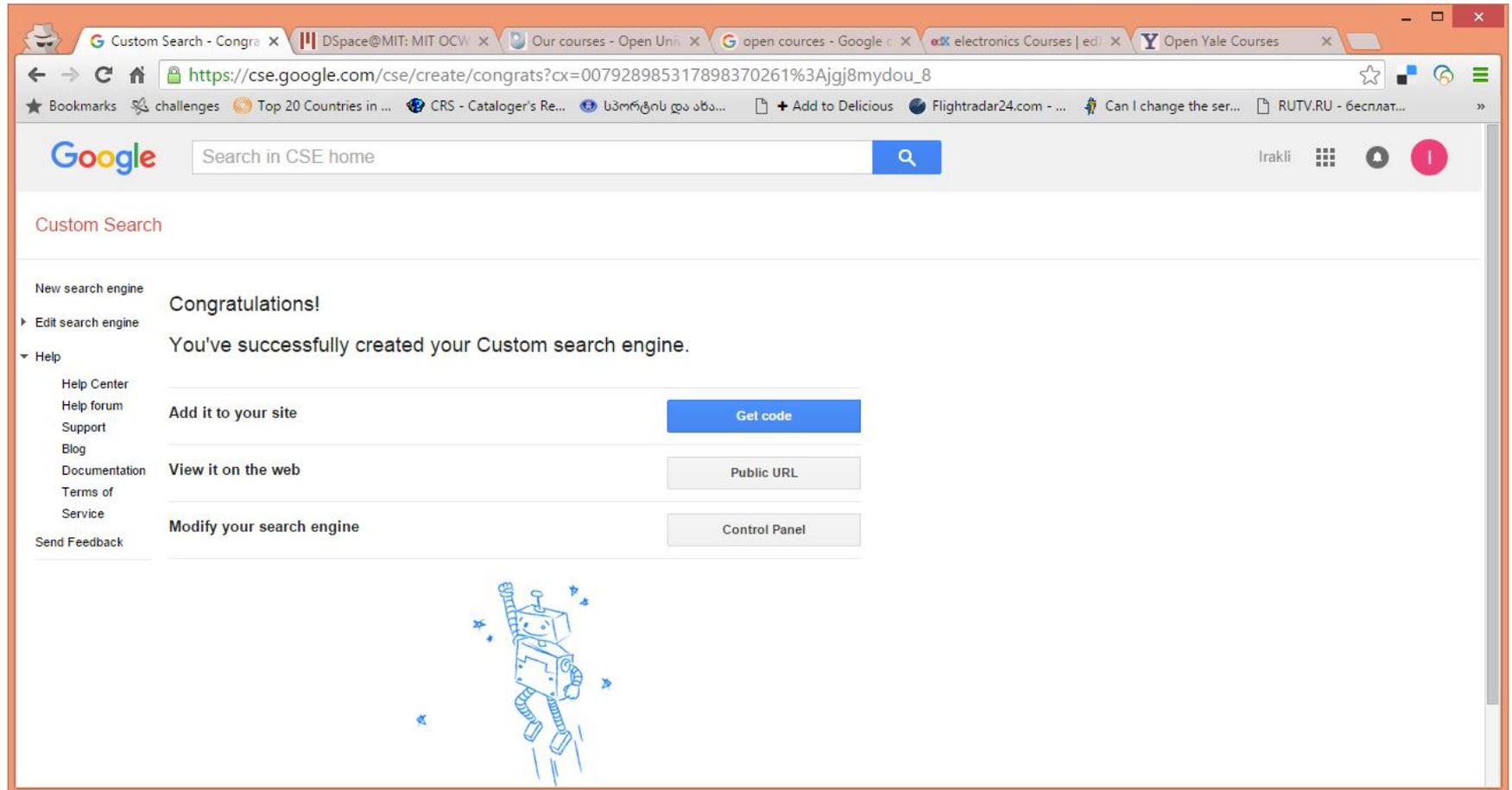
Advanced Options

By clicking 'Create', you agree with the [Terms of Service](#).

CREATE

© 2016 Google - [Google Home](#) - [About Google](#) - [Privacy Policy](#)

Result



The screenshot shows a web browser window with multiple tabs. The active tab is titled "Custom Search - Congra...". The address bar shows the URL: https://cse.google.com/cse/create/congrats?cx=007928985317898370261%3Ajgj8mydou_8. The page content includes:

- Google** logo and search bar with the text "Search in CSE home".
- Custom Search** header.
- New search engine** section with a sub-header "Edit search engine".
- Congratulations!** message: "You've successfully created your Custom search engine."
- Help** menu with links: Help Center, Help forum, Support, Blog, Documentation, Terms of Service, and Send Feedback.
- Add it to your site** button with a **Get code** button.
- View it on the web** button with a **Public URL** button.
- Modify your search engine** button with a **Control Panel** button.
- A blue cartoon robot illustration at the bottom.

Result

- ▶ Congratulations! You have done it.
 - ▶ Three options here:
 - ▶ Public URL – link to a search box (go there and start searching)
 - ▶ Code – simple code which can be embedded to your web page and produce search from your site
 - ▶ Control panel – lets you to modify or refine your engine
- 

Control Panel

The screenshot shows the Google Custom Search Control Panel in a browser window. The browser's address bar displays the URL: `https://cse.google.com/cse/setup/basic?cx=007928985317898370261%3Ajgj8mydou_8`. The page title is "Custom Search".

The interface is divided into several sections:

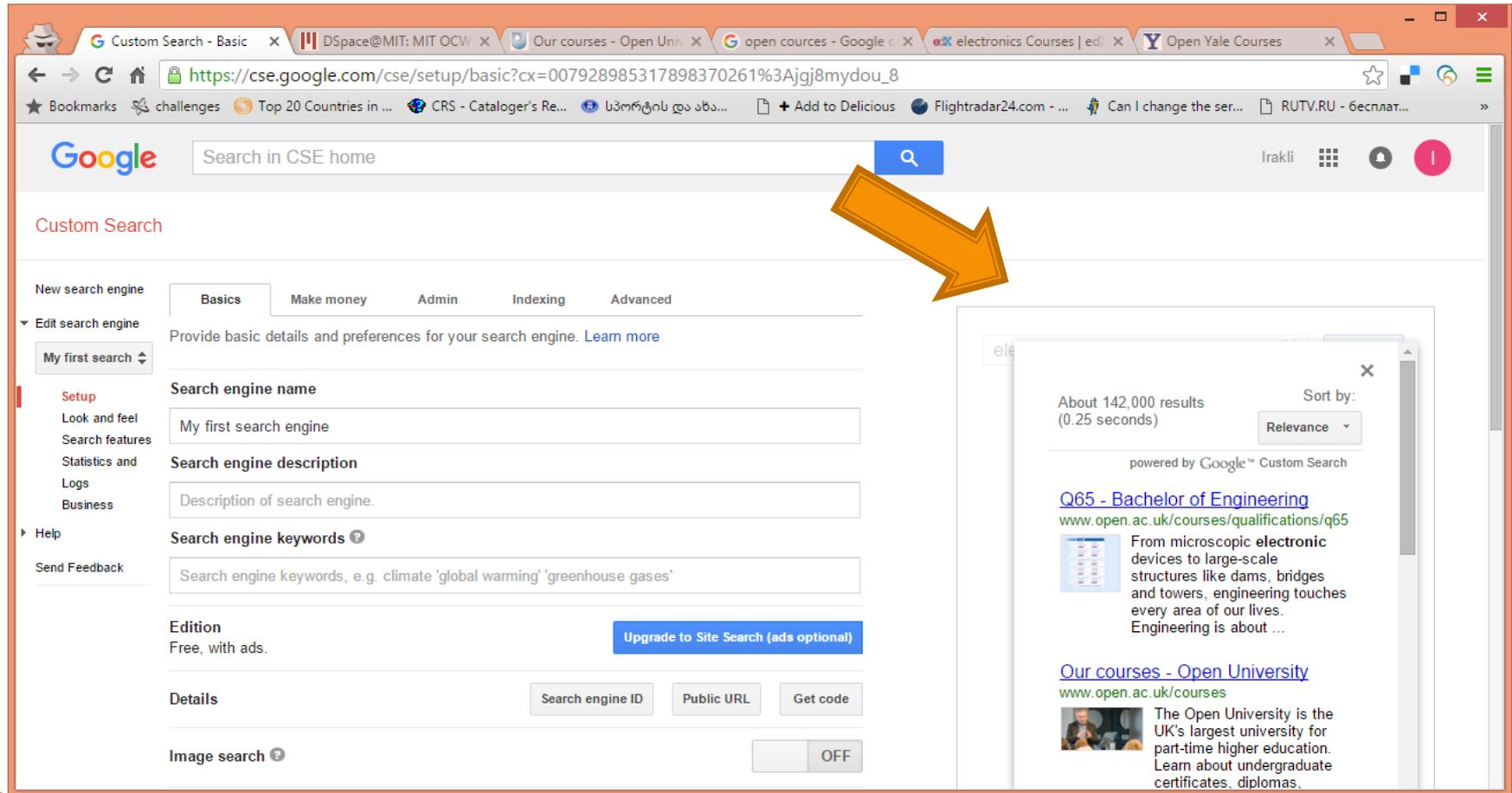
- Navigation:** A top bar with the Google logo, a search box labeled "Search in CSE home", and user profile information for "Irakli".
- Custom Search:** A sub-header for the current page.
- New search engine:** A section with tabs for "Basics", "Make money", "Admin", "Indexing", and "Advanced".
- Edit search engine:** A section with a dropdown menu for "My first search" and a "Send Feedback" link.
- Setup:** A sidebar menu with options: "Setup", "Look and feel", "Search features", "Statistics and Logs", and "Business".
- Form Fields:**
 - Search engine name:** A text box containing "My first search engine".
 - Search engine description:** A text box containing "Description of search engine."
 - Search engine keywords:** A text box containing "Search engine keywords, e.g. climate 'global warming' 'greenhouse gases'".
- Edition:** A section showing "Free, with ads." and a blue button labeled "Upgrade to Site Search (ads optional)".
- Details:** A section with three buttons: "Search engine ID", "Public URL", and "Get code".
- Image search:** A section with a toggle switch currently set to "OFF".

On the right side of the page, there is a search box containing the text "electronics" and a search button.

Control panel

- ▶ A control panel includes
- ▶ Tools for modification of your engine – on the left
- ▶ Demo search box, where you can make searches and test various modification options – on the right

Demo – just enter a search term and try



The image shows a browser window displaying the Google Custom Search setup interface. The browser's address bar shows the URL: https://cse.google.com/cse/setup/basic?cx=007928985317898370261%3Aajg8mydou_8. The page title is "Custom Search - Basic".

The main content area is titled "Custom Search" and includes a search bar with the text "Search in CSE home" and a magnifying glass icon. Below this, there are tabs for "Basics", "Make money", "Admin", "Indexing", and "Advanced".

The "Basics" tab is active, showing the "Edit search engine" section. It includes a dropdown menu for "My first search" and a "Setup" section with the following fields:

- Search engine name:** My first search engine
- Search engine description:** Description of search engine.
- Search engine keywords:** Search engine keywords, e.g. climate 'global warming' 'greenhouse gases'

At the bottom of the setup section, there is an "Edition" field set to "Free, with ads." and a blue button labeled "Upgrade to Site Search (ads optional)".

Below the setup section, there are "Details" buttons for "Search engine ID", "Public URL", and "Get code". At the very bottom, there is an "Image search" toggle set to "OFF".

A large orange arrow points from the search bar area towards the right side of the page. On the right side, a search results overlay is visible, showing "About 142,000 results (0.25 seconds)" and "Sort by: Relevance". The results include:

- Q65 - Bachelor of Engineering**
www.open.ac.uk/courses/qualifications/q65
From microscopic **electronic** devices to large-scale structures like dams, bridges and towers, engineering touches every area of our lives. Engineering is about ...
- Our courses - Open University**
www.open.ac.uk/courses
The Open University is the UK's largest university for part-time higher education. Learn about undergraduate certificates, diplomas.

A result page

The screenshot shows a web browser window with the URL www.open.ac.uk/courses/qualifications/q65. The page features the Open University logo and a navigation menu with options like 'Courses', 'Postgraduate', 'Research', 'About', 'News & media', and 'Employers'. A secondary menu includes 'Courses', 'A to Z of subjects', 'Course types', 'Careers', 'What it's like to study', 'Can I do it?', and 'Fees and funding'. A message states 'You are viewing information for Georgia. [Change country.](#)'

The main content area displays a large image of a person wearing safety glasses in a workshop. Below the image, the text 'Bachelor of Engineering (Honours)' is prominently displayed. To the right of the image are social media icons for Facebook, Twitter, and Google+. A light blue tooltip box is overlaid on the page, containing the following information:

- Degree ?
- Course code Q65
- Credits ?

A 'Back to top' button is visible at the bottom right of the tooltip.

Modification of results layout – the left menu – “Look and feel”

The image shows a browser window displaying the Google Custom Search 'Look and feel' settings page. The browser's address bar shows the URL: `https://cse.google.com/cse/lookandfeel/layout?cx=007928985317898370261:jgj8mydou_8`. The page title is 'Custom Search'. The main content area is titled 'New search engine' and includes tabs for 'Layout', 'Themes', 'Customize', and 'Thumbnail'. Under the 'Layout' tab, there is a section for 'Edit search engine' with a search bar and a 'Web search' button. Below this, there are six layout options: 'Overlay' (highlighted with a blue border), 'Two page', 'Full width', 'Two column', 'Compact', and 'Results only'. The 'Overlay' option shows a grid of search results with a central search bar. The 'Two page' option shows two columns of results with a plus sign between them. The 'Full width' option shows a single wide row of results. The 'Two column' option shows two columns of results. The 'Compact' option shows a single column of results with a mobile phone icon. The 'Results only' option shows a single column of results without a search bar. On the right side of the page, there is a preview of the search engine with a search bar and a search button.

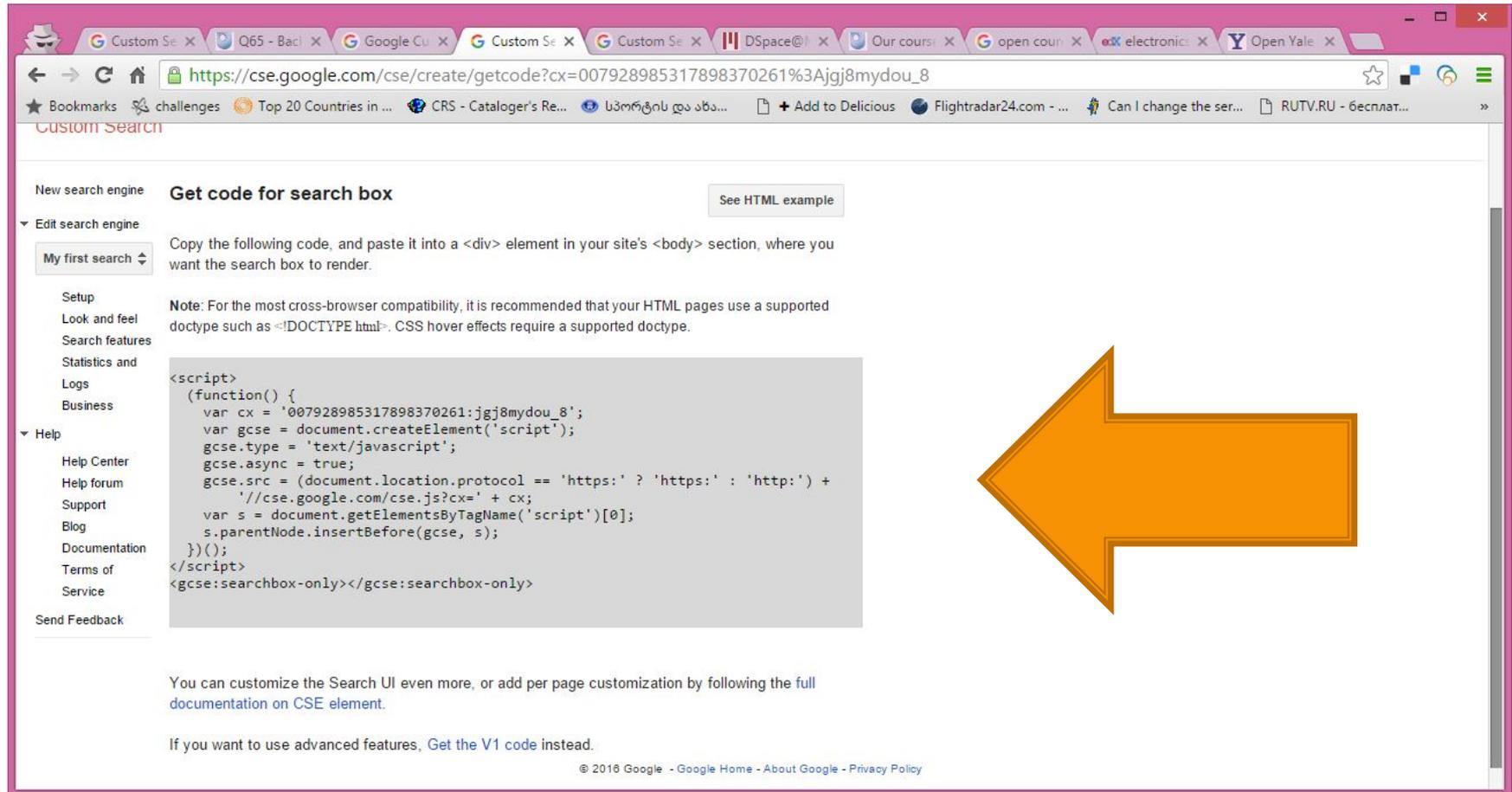
Look and feel

- ▶ Gives you several options how the results will be shown:
- ▶ On top of a search box
- ▶ In a new browser Tab
 - Opens a new Tab and keeps your site open
 - Opens a result screen in the same Tab
- ▶ In another page of a site
- ▶ Etc.
- ▶ You can choose any option and test results on the right side of the Control Panel page
- ▶ When you are finished (found an option which best suits to your goals) put it you're your website to let users make searches.

How to put a resulting CSE on your website

- ▶ Public URL – just put a link with a proper place and text
- ▶ Embed a code to your website to let users make searches directly from your site

Get code – click on a button to find the code



Custom Search

New search engine **Get code for search box** [See HTML example](#)

▼ Edit search engine

My first search ▾

Setup
Look and feel
Search features
Statistics and
Logs
Business

▼ Help

Help Center
Help forum
Support
Blog
Documentation
Terms of
Service

Send Feedback

Copy the following code, and paste it into a <div> element in your site's <body> section, where you want the search box to render.

Note: For the most cross-browser compatibility, it is recommended that your HTML pages use a supported doctype such as <!DOCTYPE html>. CSS hover effects require a supported doctype.

```
<script>
(function() {
  var cx = '007928985317898370261:jgj8mydou_8';
  var gcse = document.createElement('script');
  gcse.type = 'text/javascript';
  gcse.async = true;
  gcse.src = (document.location.protocol == 'https:' ? 'https:' : 'http:') +
    '//cse.google.com/cse.js?cx=' + cx;
  var s = document.getElementsByTagName('script')[0];
  s.parentNode.insertBefore(gcse, s);
})();
</script>
<gcse:searchbox-only></gcse:searchbox-only>
```

You can customize the Search UI even more, or add per page customization by following the [full documentation on CSE element](#).

If you want to use advanced features, [Get the V1 code](#) instead.

© 2016 Google - [Google Home](#) - [About Google](#) - [Privacy Policy](#)

Embedding a code

- ▶ Just copy the code
 - ▶ Note two versions of a code V1 & V2 – you have to choose a one which will work properly on your site
 - ▶ The version depends on your CMS and you can just test both of them.
- 

Sample – embedding a code to my site – to a blog entry

The screenshot shows a web browser window with the URL www.sciencelib.ge/en/node/add/blog. The page header features the National Scientific Library - Georgia logo and a navigation bar with the text "ქართული | Ask librarian".

The main content area is titled "Home » Create content" and "Submit Blog entry". It includes a "Title: *" field and a "Categories" dropdown menu. The categories list includes:

- რეზიუმეები
- ადმინისტრაცია და ორგანიზაცია
- არქიტექტურა
- ასტრონომია და ასტროფიზიკა
- ბიზნესი
- ბიოგრაფია
- ბიოლოგია ზოგადად
- ბოტანიკა
- განათლება
- გეოგრაფია

Below the categories is a field for "Universal Decimal Classification (UDC): *".

The left sidebar contains a "Search the site" section with a search box and a "CAPTCHA" link, and an "Irakli" section with a list of links:

- About the library
 - Slides from our library
 - Location
- eArchive of rare photos
- Georgian Universities on Youtube
- My blog
- Create content
 - Blog entry
 - Event

The right sidebar features a "Search our catalogue" section with a search box and a "Go" button, a "Random Library Photo" section with a photo of a library aisle, and a "Ministry of Education and Science" section.

Embedding...

The screenshot shows a web browser window with multiple tabs. The active tab is 'www.sciencelib.ge/en/node/add/blog'. The browser's address bar and various icons are visible at the top. On the left side, there is a navigation menu with categories like 'Site building', 'Site configuration', 'User management', and 'Logs'. The main content area is a form for adding a blog post. It features a 'Body:' section with a rich text editor toolbar and a code editor area containing JavaScript code for embedding a Google Custom Search widget. Below the code editor is an 'Input format' section. At the bottom, there is a 'Log message:' section with a rich text editor toolbar and a text input area. On the right side of the page, there are several widgets: 'Events' showing a calendar for January 2016, 'Upcoming library events' with a message that no events are available, and 'eIFL.net - innovative libraries worldwide' with a list of news items.

Body: *

Hide HTML H1 H2 H3 H4 P S E @ [v] ←

```
<script type="text/javascript" async="" src="http://cse.google.com/cse.js?cx=007928985317898370261:jg18mydou_8"></script>
<script>
  (function() {
    var cx = '007928985317898370261:jg18mydou_8';
    var gcse = document.createElement('script');
    gcse.type = 'text/javascript';
    gcse.async = true;
    gcse.src = (document.location.protocol == 'https:' ? 'https:' : 'http:') +
      '//cse.google.com/cse.js?cx=' + cx;
    var s = document.getElementsByTagName('script')[0];
    s.parentNode.insertBefore(gcse, s);
  })();
</script>
```

→ **Input format**

Log message:

Choose style: ▼ Font: ▼ Font size: ▼

B I U | [text alignment icons] | [list icons] | [undo/redo icons] | [insert icons] | [check/refresh icons]

Events

« January 2016

Mon	Tue	Wed	Thu	Fri	Sat	Sun
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

Upcoming library events

- No upcoming events available

[more](#)

eIFL.net - innovative libraries worldwide

- News from the last SCCR of 2015
- EIFL and IDS webinar: DSpace repositories
- EIFL webinar: Google Custom Search
- EIFL-funded open access workshops in Palestine attract 100 people

Code V2

Custom Search

New search engine **Get V1 code for search box** [See HTML example](#)

▶ Edit search engine

▼ Help

- Help Center
- Help forum
- Support
- Blog
- Documentation
- Terms of Service
- Send Feedback

Copy the following code, and paste it into a <div> element in your site's <body> section, where you want the search box to render.

Note: For the most cross-browser compatibility, it is recommended that your HTML pages use a supported doctype such as <!DOCTYPE html>. CSS hover effects require a supported doctype.

```
<div id='cse-search-form' style='width: 100%;'>Loading</div>
<script src='//www.google.com/jsapi' type='text/javascript'></script>
<script type='text/javascript'>
google.load('search', '1', {language: 'en', style: google.loader.themes.V2_DEFAULT});
google.setOnLoadCallback(function() {
  var customSearchOptions = {};
  var orderByOptions = {};
  orderByOptions['keys'] = [{label: 'Relevance', key: ''}, {label: 'Date', key: ''}];
  customSearchOptions['enableOrderBy'] = true;
  customSearchOptions['orderByOptions'] = orderByOptions;
  var customSearchControl = new google.search.CustomSearchControl('007928985317898370261');
  customSearchControl.setResultSetSize(google.search.Search.FILTERED_CSE_RESULTSET);
  var options = new google.search.DrawOptions();
  options.enableSearchboxOnly('https://www.google.com/cse?cx=007928985317898370261');
```

You can customize the look and feel even more or extend the functionality of the Custom Search

Resulting page with a CSE search box

The screenshot shows a web browser window with the URL www.sciencelib.ge/en/node/70. The page is titled "National Scientific Library - Georgia" and features a header with a logo and a banner image of a book and a pen. The main content area displays a confirmation message: "Your Blog entry has been created." with a large orange arrow pointing to it. The message is submitted by "Irakli" on Mon, 01/18/2016 - 05:27. Below the message, there is a search bar and a link to "You are watching this post, click to stop watching". The page also includes a sidebar with a search box and a list of links, and a right sidebar with a search box and a "Random Library Photo" section.

Test post1 | Natli x Google Custom x etheses.whiteros x კომპონი: მუღი x WebCite - Wikip x Managing in the x How to Recogni: x google custom s x

www.sciencelib.ge/en/node/70

Apps ★ Bookmarks challenges Top 20 Countries in ... CRS - Cataloger's Re... სპორტის და ახა... + Add to Delicious Flightradar24.com - ... Can I change the ser...

 National Scientific Library - Georgia

ბძელო ლექსი, ნაძვალ მწიგნობრობა. ამისთვისაა
ღმრთებელი ამ ცნობა, ვითარცა სერვისი, რესურსი უნებია
ა იგი ამ ყოველივეს, გრძობენთან სავსეა.

ქართული | Ask librarian

 Search the site

→ CAPTCHA

Search

 Irakli

- ▼ About the library
 - Slides from our library
 - Location
- eArchive of rare photos
- Georgian Universities on Youtube
- My blog
- ▶ Create content
- Invite your friends and colleagues

Home » Blogs » My blog

Test post1

[view](#) [edit](#) [track](#)

Your Blog entry has been created.

Submitted by **Irakli** on Mon, 01/18/2016 - 05:27.

 You are watching this post, click to stop watching

» [Irakli's blog](#) | [Add new comment](#) | [SHARE](#) 

 Search our catalogue

Library

Random Library Photo

www.flickr.com





 Ministry of Education and Science

Results of a search – “nano technology”

The image shows a screenshot of a web browser displaying search results for 'nano technology' on Google Custom Search. The browser's address bar shows the URL: https://cse.google.com/cse?cx=007928985317898370261:jgj8mydou_8&q=nano%20technology&.... The search bar contains the text 'nano technology'. The results section shows 'About 5 results (0.48 seconds)' and a 'Sort by: Relevance' dropdown menu. The first section is 'Ads by Google' and contains four entries:

- Nanotechnology Engineering**
www.asme.org/nanotechnology
Submit papers on application of nano-devices, nano-scale structures.
- Electron Detectors**
www.surface-concept.com/
for diffraction and spectroscopy nanotechnology
- +99.99% MWCNT Producer**
www.ntherma.com/
+99.99% Purity MWCNT Leading Hi Purity MWCNT Provider
- Electrospin Nanospinner**
www.inovenso.com/
High throughput electrospinning machines for nanofiber production
References/Projects - Contacts

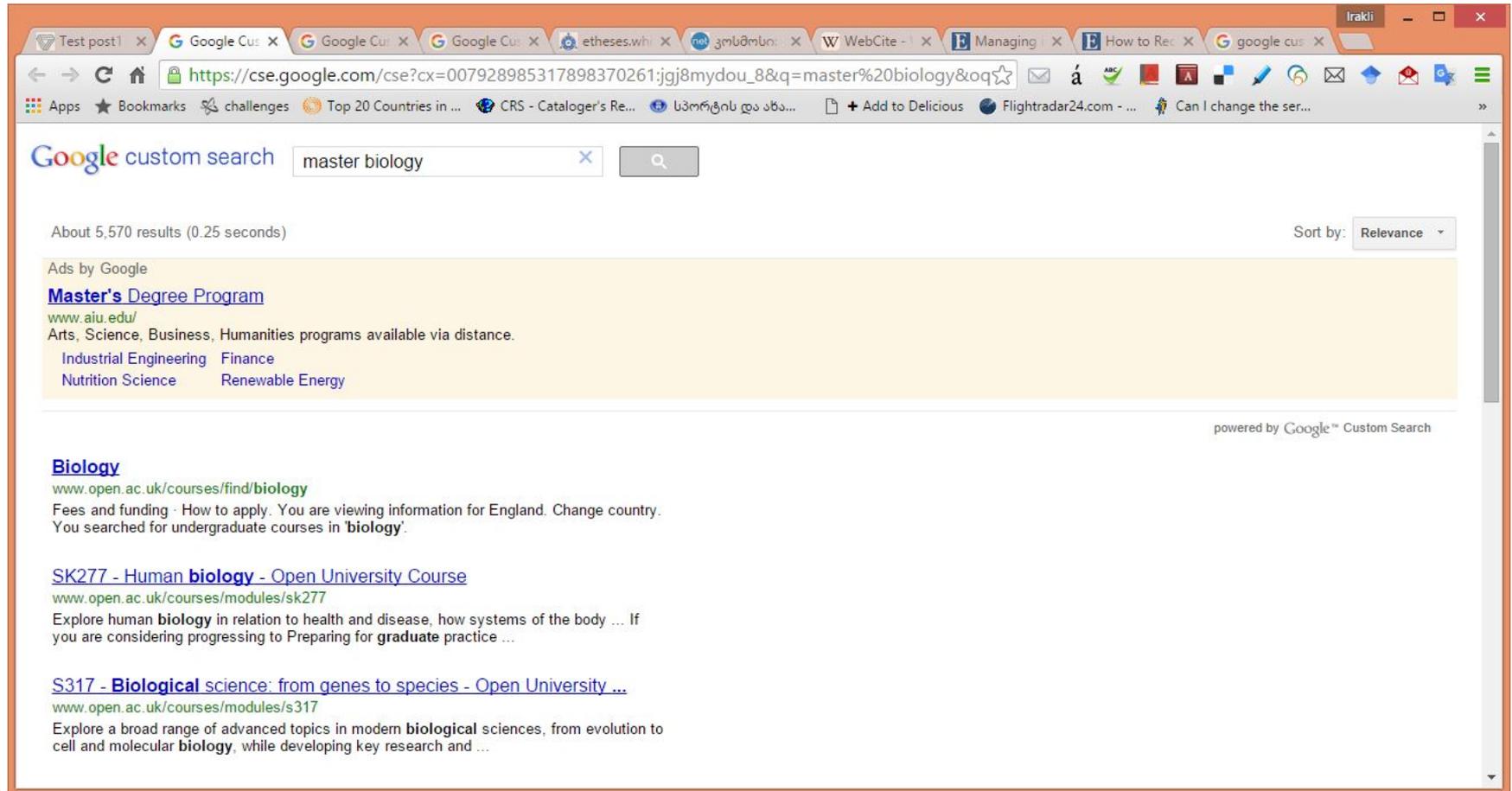
Below the ads, it says 'powered by Google™ Custom Search'. Further down, it says 'Showing results for [nanotechnology](#)
Search instead for [nano technology](#)'. The first organic result is:

- T356 - Engineering small worlds: micro and nano technologies ...**
www.open.ac.uk/courses/modules/t356
The printed study materials are divided into three main themes together with a supporting text book as an introduction to the subject of nanotechnology.

Re-modifying your engine

The image shows a web browser window displaying the Google Custom Search Engine (CSE) lookandfeel layout editor. The browser's address bar shows the URL: https://cse.google.com/cse/lookandfeel/layout?cx=007928985317898370261:jgj8mydou_8. The page features several layout preview options, including "Two column", "Compact", and "Results only". A "Google hosted" option is highlighted with a blue border, showing a search bar and the Google logo. Below the preview options, there are radio buttons for "Search results appear in:" with "Current window" selected and "New window" unselected. At the bottom, there are two buttons: "Save" and "Save & Get Code". The footer of the page reads "© 2016 Google - Google Home - About Google - Privacy Policy".

Now results are in a new Tab



The screenshot shows a web browser window with multiple tabs. The active tab is titled "google cus" and displays the Google Custom Search results for the query "master biology". The search results are sorted by "Relevance" and show approximately 5,570 results. The first result is an advertisement for a "Master's Degree Program" from a.iu.edu, listing various fields of study. The second result is a link to "Biology" courses from open.ac.uk, including "SK277 - Human biology" and "S317 - Biological science: from genes to species".

Test post | Google Cus | Google Cus | Google Cus | etheses.whi | 3m8mLn: | WebCite - | Managing | How to Rec | google cus | Iraki

https://cse.google.com/cse?cx=007928985317898370261:jgj8mydou_8&q=master%20biology&oq=

Apps ★ Bookmarks challenges Top 20 Countries in ... CRS - Cataloger's Re... 3m8mLn: 3m8mLn: + Add to Delicious Flightradar24.com - ... Can I change the ser...

Google custom search master biology

About 5,570 results (0.25 seconds) Sort by: Relevance

Ads by Google

Master's Degree Program
www.aiu.edu/
Arts, Science, Business, Humanities programs available via distance.
Industrial Engineering Finance
Nutrition Science Renewable Energy

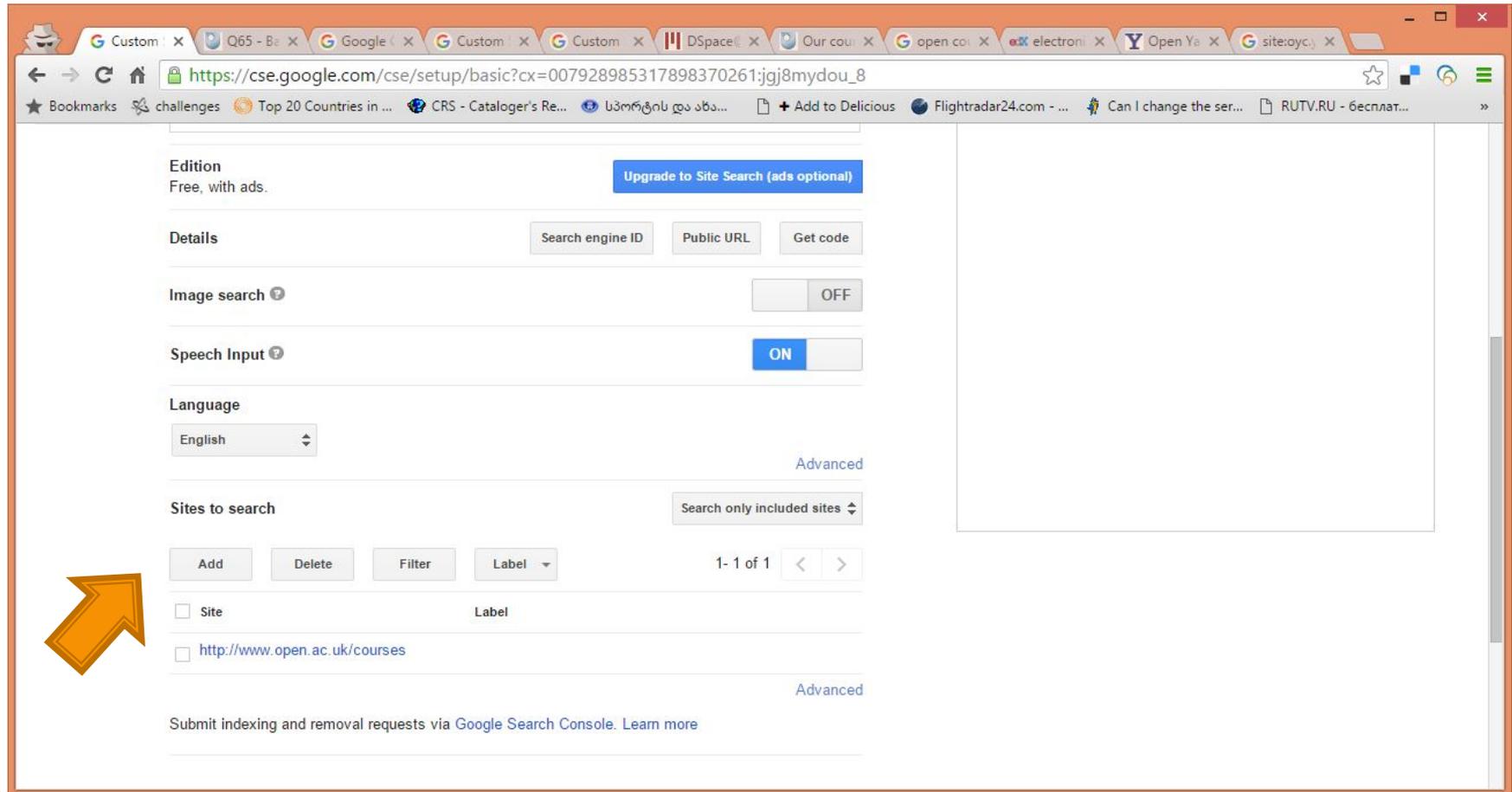
powered by Google™ Custom Search

Biology
www.open.ac.uk/courses/find/biology
Fees and funding · How to apply. You are viewing information for England. Change country.
You searched for undergraduate courses in 'biology'.

SK277 - Human biology - Open University Course
www.open.ac.uk/courses/modules/sk277
Explore human **biology** in relation to health and disease, how systems of the body ... If you are considering progressing to Preparing for **graduate** practice ...

S317 - Biological science: from genes to species - Open University ...
www.open.ac.uk/courses/modules/s317
Explore a broad range of advanced topics in modern **biological** sciences, from evolution to cell and molecular **biology**, while developing key research and ...

Adding a site for search from the CSE Control panel – “Add” button



The screenshot displays the Google Custom Search Control Panel interface. The browser's address bar shows the URL: https://cse.google.com/cse/setup/basic?cx=007928985317898370261:jjj8mydou_8. The page is divided into several sections:

- Edition:** Free, with ads. A blue button labeled "Upgrade to Site Search (ads optional)" is visible.
- Details:** Includes buttons for "Search engine ID", "Public URL", and "Get code".
- Image search:** A toggle switch currently set to "OFF".
- Speech Input:** A toggle switch currently set to "ON".
- Language:** A dropdown menu set to "English". A link for "Advanced" settings is present.
- Sites to search:** A section with a dropdown menu set to "Search only included sites". It contains a table with columns for "Site" and "Label".

Site	Label
<input type="checkbox"/> http://www.open.ac.uk/courses	

An orange arrow points to the "Add" button in the "Sites to search" section. At the bottom of the page, there is a link: "Submit indexing and removal requests via [Google Search Console](#). [Learn more](#)".

Adding a URL

The screenshot shows the Google Custom Search setup interface. A modal dialog is open, titled "Include site individually | [Include sites in bulk](#)". The dialog contains a text input field with the URL "http://oyc.yale.edu/" and a "Label" dropdown menu. Below the input field are three radio button options:

- Include all pages whose address contains this URL
- Include just this specific page or URL pattern I have entered
- Dynamically extract links from this page and include them in my search engine

Under the third option, there are three sub-options:

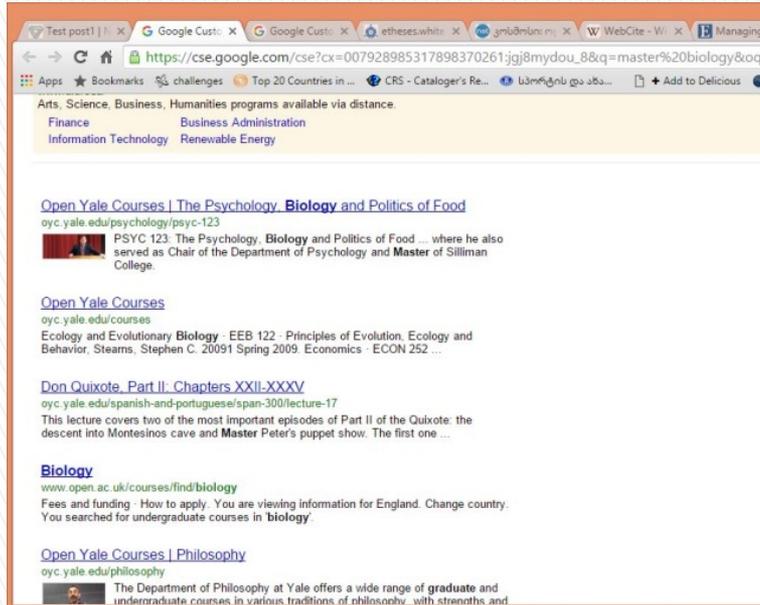
- Include all pages this page links to
- Include all partial sites this page links to
- Include all sites this page links to

At the bottom of the dialog are "Save" and "Cancel" buttons. The background page shows the "Sites to search" section with a table containing one entry: "http://www.open.ac.uk/courses".

Adding a URLs

- ▶ Several options:
- ▶ Including each URL separately or adding a list in bulk
- ▶ Add a page with specified address
- ▶ Add a pattern specifying a group of pages
- ▶ Extracting links from a specified page (good for online list of publications, where you can extract each of them individually)
- ▶ CSE has some Help documentation and you can just start with very basic options testing everything on your Control Panel before you put the results online

Result



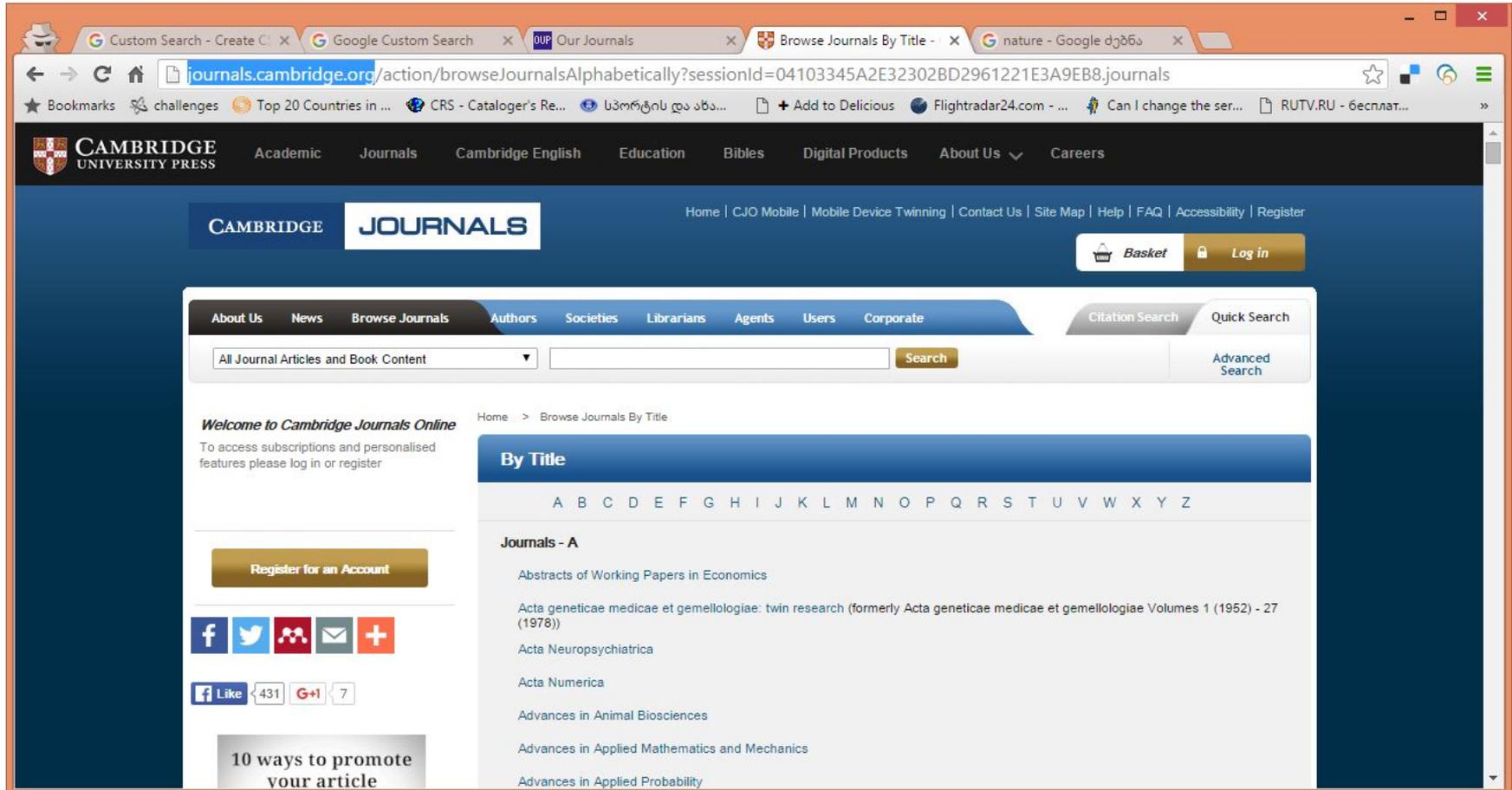
- ▶ Now our CSE is searching in multiple sites indicated in the list – e.g. open courses from various Unis



Another sample

- ▶ Goal: CSE searching through e-journals from various publishers

Oxford journals – see a proper pattern



The screenshot shows the Cambridge Journals Online website. The browser address bar displays the URL: journals.cambridge.org/action/browseJournalsAlphabetically?sessionId=04103345A2E32302BD2961221E3A9EB8&journals. The page features a navigation menu with links for Academic, Journals, Cambridge English, Education, Bibles, Digital Products, About Us, and Careers. The main header includes the Cambridge University Press logo and the text "CAMBRIDGE JOURNALS". A search bar is present with a dropdown menu set to "All Journal Articles and Book Content" and a "Search" button. The page content is titled "Browse Journals By Title" and includes a "By Title" section with an alphabetical index (A-Z). The "Journals - A" section lists several journals, including "Abstracts of Working Papers in Economics", "Acta geneticae medicae et gemellologiae: twin research (formerly Acta geneticae medicae et gemellologiae Volumes 1 (1952) - 27 (1978))", "Acta Neuropsychiatrica", "Acta Numerica", "Advances in Animal Biosciences", "Advances in Applied Mathematics and Mechanics", and "Advances in Applied Probability".

Custom Search - Create C... Google Custom Search Our Journals Browse Journals By Title - nature - Google d...
journals.cambridge.org/action/browseJournalsAlphabetically?sessionId=04103345A2E32302BD2961221E3A9EB8&journals
Bookmarks challenges Top 20 Countries in ... CRS - Cataloger's Re... სპორტის და ან... + Add to Delicious Flightradar24.com - ... Can I change the ser... RUTV.RU - бесплат...
CAMBRIDGE UNIVERSITY PRESS Academic Journals Cambridge English Education Bibles Digital Products About Us Careers
Home | CJO Mobile | Mobile Device Twinning | Contact Us | Site Map | Help | FAQ | Accessibility | Register
CAMBRIDGE JOURNALS Basket Log in
About Us News Browse Journals Authors Societies Librarians Agents Users Corporate Citation Search Quick Search
All Journal Articles and Book Content Search Advanced Search
Welcome to Cambridge Journals Online Home > Browse Journals By Title
To access subscriptions and personalised features please log in or register
Register for an Account
f t m e +
f Like 431 G+1 7
10 ways to promote your article
By Title
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Journals - A
Abstracts of Working Papers in Economics
Acta geneticae medicae et gemellologiae: twin research (formerly Acta geneticae medicae et gemellologiae Volumes 1 (1952) - 27 (1978))
Acta Neuropsychiatrica
Acta Numerica
Advances in Animal Biosciences
Advances in Applied Mathematics and Mechanics
Advances in Applied Probability

Browser tabs: Custom Search - Create C..., Google Custom Search, Our Journals, Browse Journals By Title - ..., Home : Nature a - z index

Address bar: www.nature.com/siteindex/index.html

Navigation: nature.com | Publications A-Z index | Browse by subject

Special Section and Web Focus: Access the selected articles free online!

Journal of Human Genetics

My account, Submit Manuscript, Register, Subscribe

nature.com A-Z index

Search [input] go Advanced search

Publications A - M	Publications N - Z
A	N
▶ Acta Pharmacologica Sinica	▶ Nature
▶ American Journal of Gastroenterology	▶ Nature Arabic Edition
▶ American Journal of Gastroenterology Supplements	▶ Nature Biotechnology
B	▶ Nature Cell Biology
▶ BDJ Open	▶ Nature Chemical Biology
▶ BDJ Team	▶ Nature Chemistry
▶ Blood Cancer Journal	▶ Nature Climate Change
▶ Bone Marrow Transplantation	▶ Nature Communications
▶ BoneKEy Reports	▶ Nature Digest
▶ Bone Research	▶ Nature Energy
▶ British Dental Journal	▶ Nature Genetics
▶ British Journal of Cancer	▶ Nature Geoscience
C	▶ Nature Immunology
▶ Cancer Gene Therapy	▶ Nature Materials

RSS Web Feeds

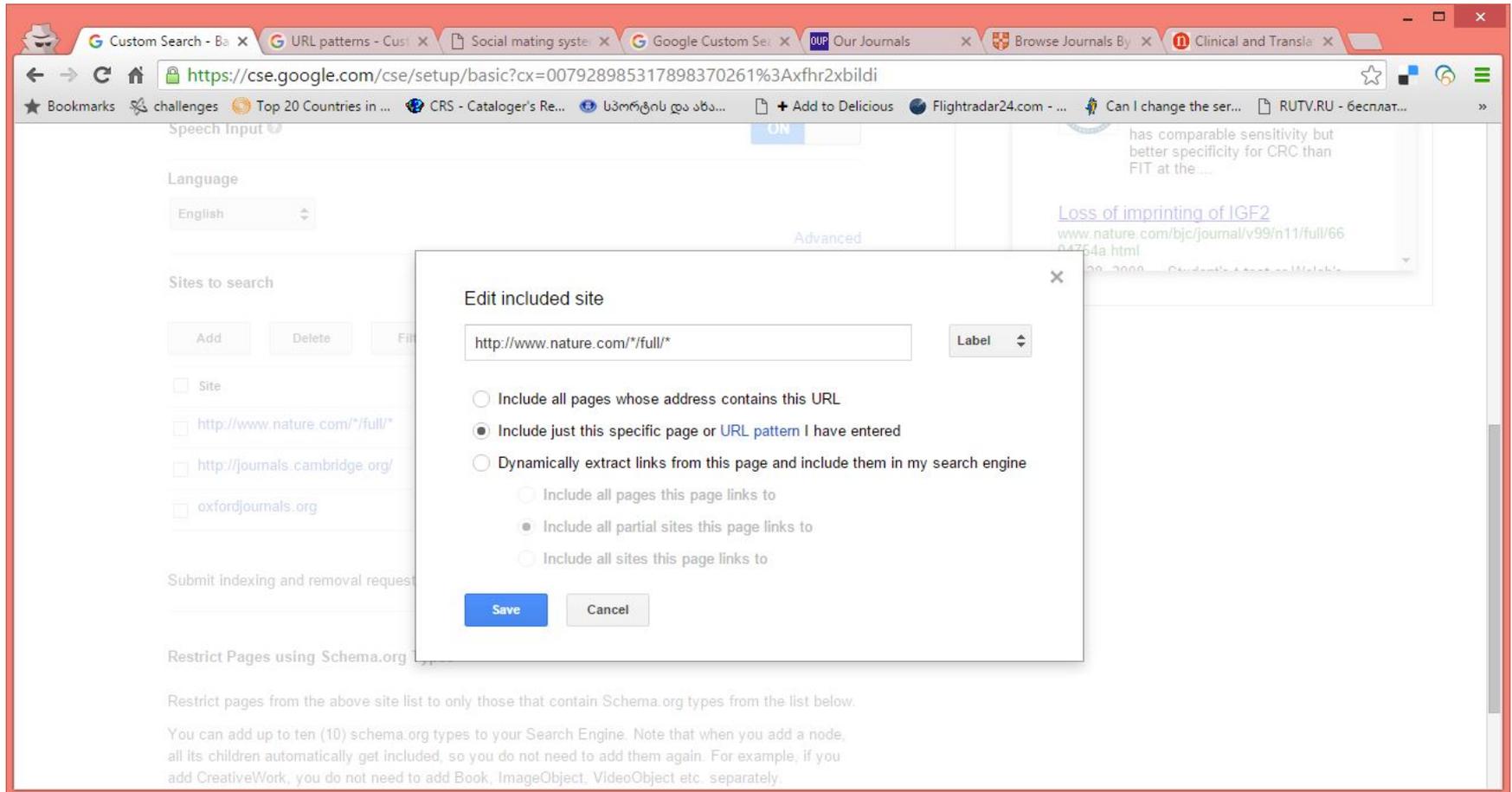
Gateways and databases

- ▶ Application Notes
- ▶ Bioentrepreneur
- ▶ BioPharma Dealmakers
- ▶ Nature Index
- ▶ Nature Milestones
- ▶ Nature Network
- ▶ Protocol Exchange
- ▶ All Gateways and databases

Productivity and information tools

- ▶ Nature.com developers
- ▶ Nature.com mobile applications

Adding URLs



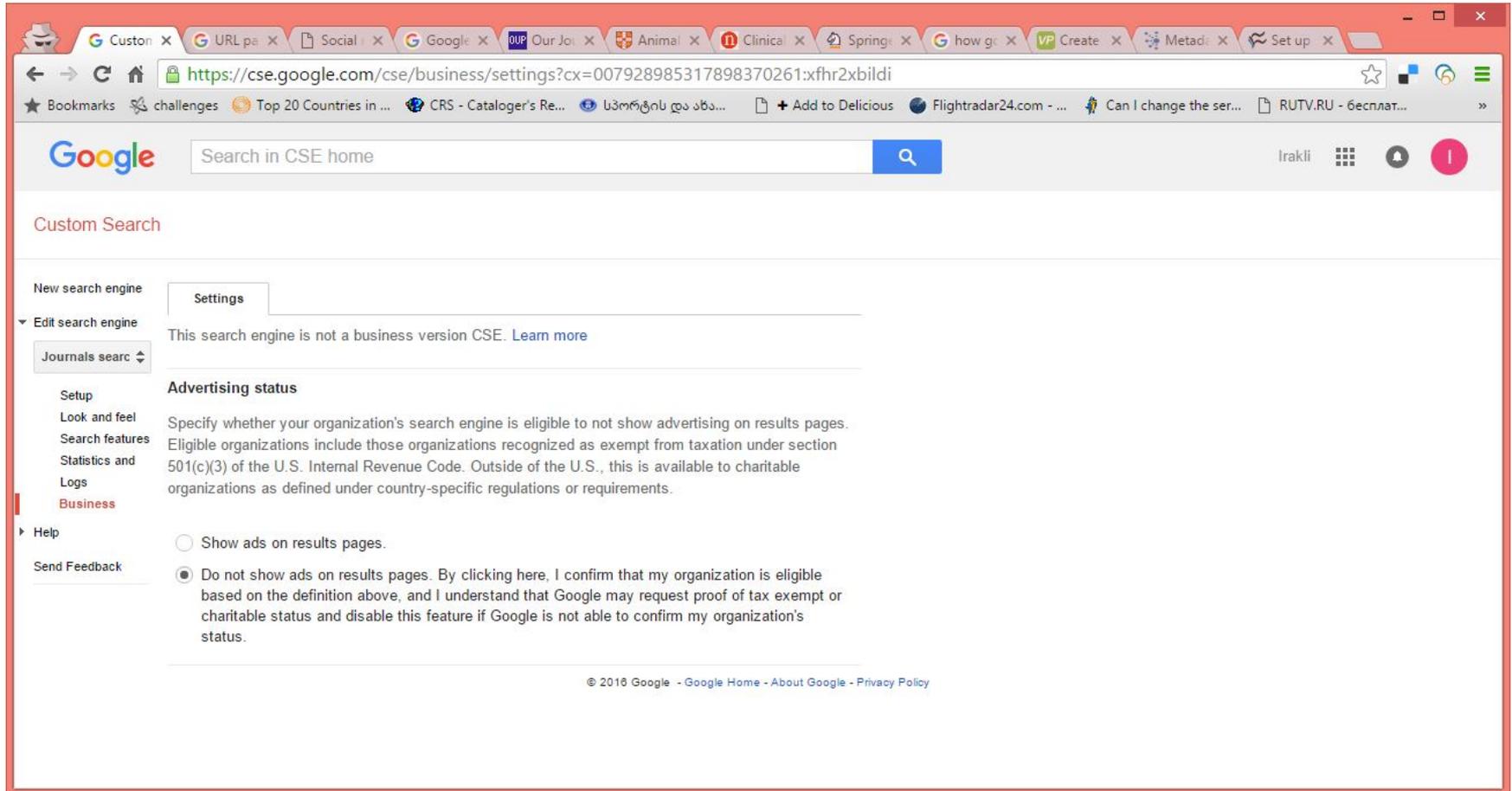
The screenshot shows a web browser window with multiple tabs. The active tab is titled "Our Journals" and displays the URL <https://cse.google.com/cse/setup/basic?cx=007928985317898370261%3Axhfr2xbildi>. The page content includes a "Language" dropdown set to "English", a "Speech Input" toggle set to "ON", and a "Sites to search" section with buttons for "Add", "Delete", and "Fill". A list of sites is shown, including "http://www.nature.com/**/*.full/*", "http://journals.cambridge.org/", and "oxfordjournals.org".

An "Edit included site" dialog box is open, centered on the screen. It contains the following elements:

- A text input field containing the URL: `http://www.nature.com/**/*.full/*`
- A "Label" dropdown menu.
- Three radio button options:
 - Include all pages whose address contains this URL
 - Include just this specific page or URL pattern I have entered
 - Dynamically extract links from this page and include them in my search engine
- Under the "Dynamically extract links" option, three sub-options:
 - Include all pages this page links to
 - Include all partial sites this page links to
 - Include all sites this page links to
- "Save" and "Cancel" buttons at the bottom.

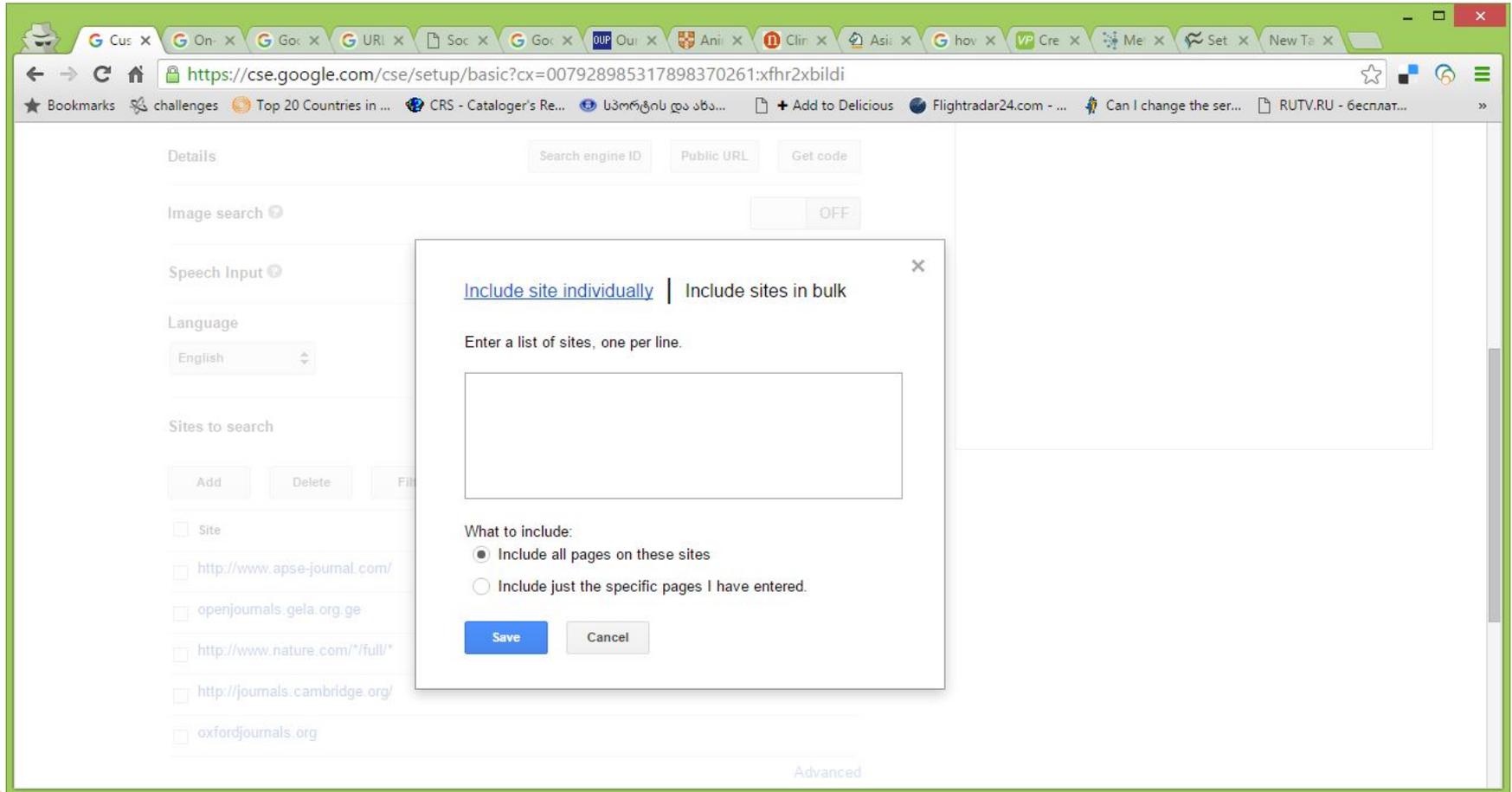
Below the dialog box, the page text reads: "Restrict Pages using Schema.org" and "Restrict pages from the above site list to only those that contain Schema.org types from the list below. You can add up to ten (10) schema.org types to your Search Engine. Note that when you add a node, all its children automatically get included, so you do not need to add them again. For example, if you add CreativeWork, you do not need to add Book, ImageObject, VideoObject etc. separately."

A new option – you can switch off advertising from results



The screenshot shows a web browser window with multiple tabs open. The active tab is titled "Custom Search" and the address bar shows the URL: <https://cse.google.com/cse/business/settings?cx=007928985317898370261:xfhr2xbildi>. The page content includes a search bar with the text "Search in CSE home" and a search button. Below the search bar, there is a "Custom Search" heading. The main content area is divided into sections: "New search engine" with a "Settings" tab, "Edit search engine" with a "Journals search" dropdown, and "Advertising status". The "Advertising status" section contains the following text: "Specify whether your organization's search engine is eligible to not show advertising on results pages. Eligible organizations include those organizations recognized as exempt from taxation under section 501(c)(3) of the U.S. Internal Revenue Code. Outside of the U.S., this is available to charitable organizations as defined under country-specific regulations or requirements." Below this text, there are two radio button options: "Show ads on results pages." (which is unselected) and "Do not show ads on results pages. By clicking here, I confirm that my organization is eligible based on the definition above, and I understand that Google may request proof of tax exempt or charitable status and disable this feature if Google is not able to confirm my organization's status." (which is selected). At the bottom of the page, there is a copyright notice: "© 2016 Google - Google Home - About Google - Privacy Policy".

Adding URLs in bulk



The screenshot shows the CSE Google setup page in a browser. The URL in the address bar is <https://cse.google.com/cse/setup/basic?cx=007928985317898370261:xfhr2xbildi>. The page has several sections: Details, Image search (OFF), Speech Input, Language (English), and Sites to search. The Sites to search section contains a list of sites with checkboxes: Site, <http://www.apse-journal.com/>, openjournals.gela.org.ge, http://www.nature.com/**/full/, <http://journals.cambridge.org/>, and oxfordjournals.org. A modal dialog is open in the center, titled "Include site individually | Include sites in bulk". The dialog has a close button (X) in the top right corner. It contains the text "Enter a list of sites, one per line." and a large empty text input field. Below the input field, there is a section "What to include:" with two radio buttons: "Include all pages on these sites" and "Include just the specific pages I have entered." At the bottom of the dialog are "Save" and "Cancel" buttons. The word "Advanced" is visible at the bottom right of the page.

Several tips for note

- ▶ CSE is not expanding a general Google search but limiting it to a specific sites
- ▶ CSE is producing not very accurate results – meaning that it does not eliminate entirely junk results but is minimizing results to ones which are more practical to you – eventually you have to make decision which one you choose
- ▶ CSE cannot search in specified metadata fields (at least in the way described above), but it is searching through all metadata (unlike engines of ILSs or Digital repository software)
- ▶ CSE normally does not recognize truncations (which might be useful for specific languages in specific software systems). E.g KOHA can search for truncated terms (like phys or math), Dspace can search for phys* or math* .

Some more thoughts

- ▶ Searching through Websites (meaning a combination of web pages) is more or less intuitive – you have specify a pattern carefully
 - ▶ But CSE also makes possible to search through Websites produced via various databases (CMSs, ILSs, etc..)
 - ▶ For example you can create a search through Drupal sites or Wikipedia pages
 - ▶ Or you can search through OPACs of various ILSs (not all)
- 

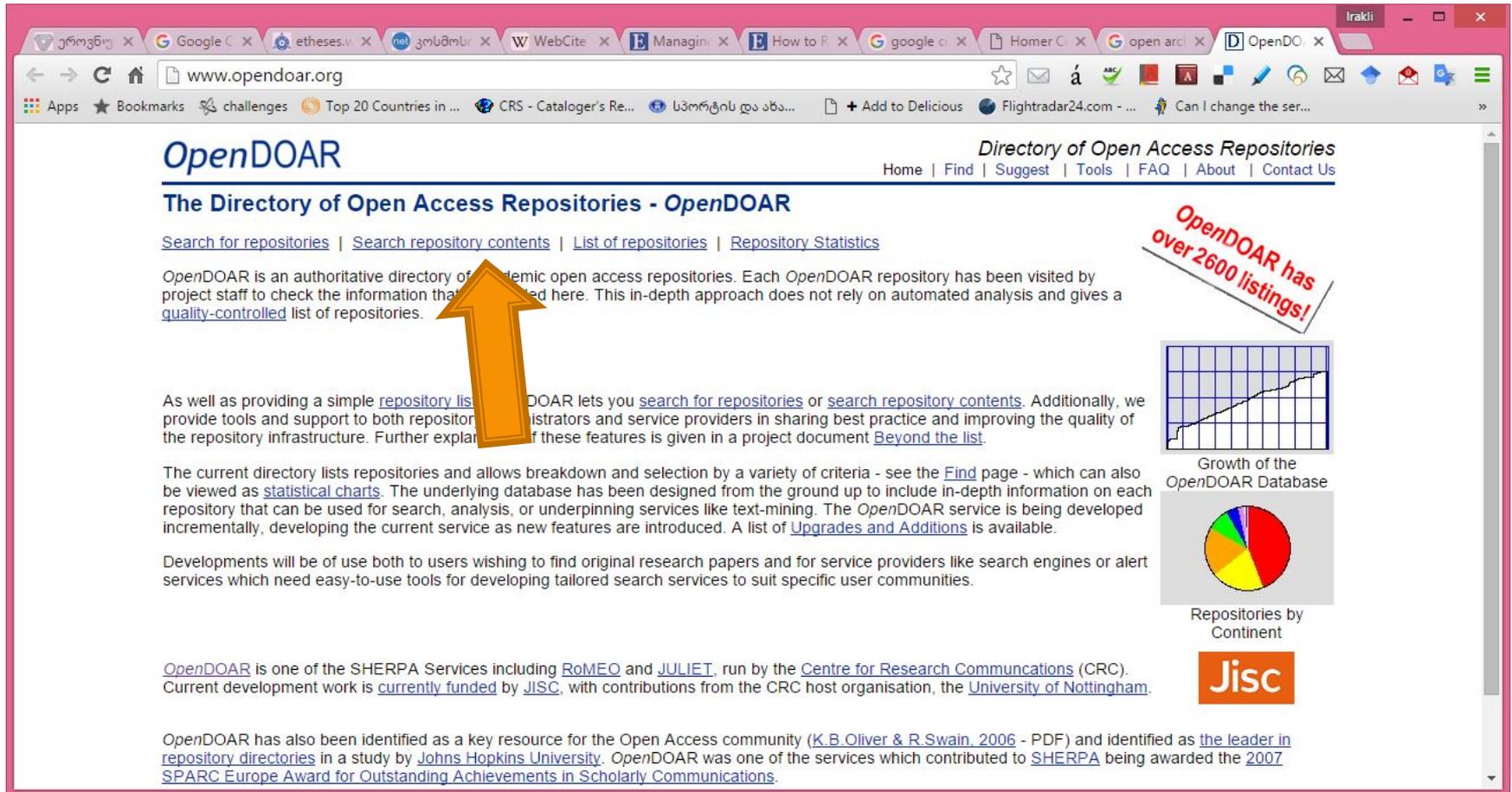
Searching Koha OPACs

- ▶ You can create a search engine producing results of searches through multiple online catalogues produced via KOHA ILS (which is popular in EIFL countries and libraries)
 - ▶ As a result you can have on your website single search box through all KOHA OPACs in your country
 - ▶ All you need is just to create a standard CSE and add URLs like `opac.library1.com`, `opac.library2.com`, `opac.library3.ge`, etc.
 - ▶ We could even create a union OPAC of EIFL libraries this way.
- 

CSE in Open repositories or Open e-journals

- ▶ CSE works well also in software systems like Dspace or Open Journal Systems, which are also popular in EIFL countries and libraries
- ▶ In these cases the engine is searching through metadata as well as through full text files.

OpenDOAR CSE



The screenshot shows the OpenDOAR website homepage. The browser's address bar displays www.opendoar.org. The page features a navigation menu with links for Home, Find, Suggest, Tools, FAQ, About, and Contact Us. A prominent orange arrow points to the search links: [Search for repositories](#), [Search repository contents](#), [List of repositories](#), and [Repository Statistics](#). The main text describes OpenDOAR as an authoritative directory of academic open access repositories, visited by project staff to check information. A callout box on the right states "OpenDOAR has over 2600 listings!". Below this, there are two charts: a line graph titled "Growth of the OpenDOAR Database" and a pie chart titled "Repositories by Continent". The Jisc logo is visible at the bottom right. The footer contains information about SHERPA Services (RoMEO and JULIET) and mentions that OpenDOAR was identified as a key resource and awarded the 2007 SPARC Europe Award.

OpenDOAR
Directory of Open Access Repositories
Home | Find | Suggest | Tools | FAQ | About | Contact Us

The Directory of Open Access Repositories - OpenDOAR

[Search for repositories](#) | [Search repository contents](#) | [List of repositories](#) | [Repository Statistics](#)

OpenDOAR is an authoritative directory of academic open access repositories. Each OpenDOAR repository has been visited by project staff to check the information that is listed here. This in-depth approach does not rely on automated analysis and gives a [quality-controlled](#) list of repositories.

As well as providing a simple [repository list](#), OpenDOAR lets you [search for repositories](#) or [search repository contents](#). Additionally, we provide tools and support to both repository administrators and service providers in sharing best practice and improving the quality of the repository infrastructure. Further explanation of these features is given in a project document [Beyond the list](#).

The current directory lists repositories and allows breakdown and selection by a variety of criteria - see the [Find](#) page - which can also be viewed as [statistical charts](#). The underlying database has been designed from the ground up to include in-depth information on each repository that can be used for search, analysis, or underpinning services like text-mining. The OpenDOAR service is being developed incrementally, developing the current service as new features are introduced. A list of [Upgrades and Additions](#) is available.

Developments will be of use both to users wishing to find original research papers and for service providers like search engines or alert services which need easy-to-use tools for developing tailored search services to suit specific user communities.

OpenDOAR is one of the SHERPA Services including [RoMEO](#) and [JULIET](#), run by the [Centre for Research Communications](#) (CRC). Current development work is [currently funded](#) by [JISC](#), with contributions from the CRC host organisation, the [University of Nottingham](#).

OpenDOAR has also been identified as a key resource for the Open Access community ([K.B. Oliver & R. Swain, 2006](#) - PDF) and identified as [the leader in repository directories](#) in a study by [Johns Hopkins University](#). OpenDOAR was one of the services which contributed to [SHERPA](#) being awarded the [2007 SPARC Europe Award for Outstanding Achievements in Scholarly Communications](#).

OpenDOAR has over 2600 listings!

Growth of the OpenDOAR Database

Repositories by Continent

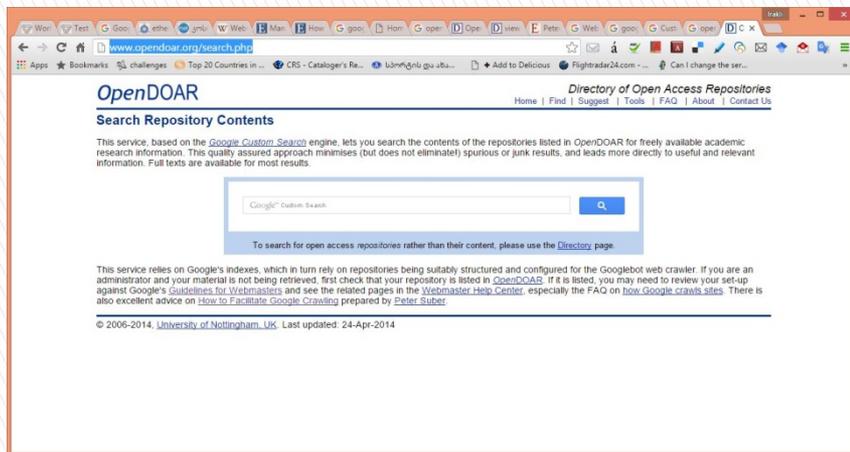
Jisc

OpenDOAR

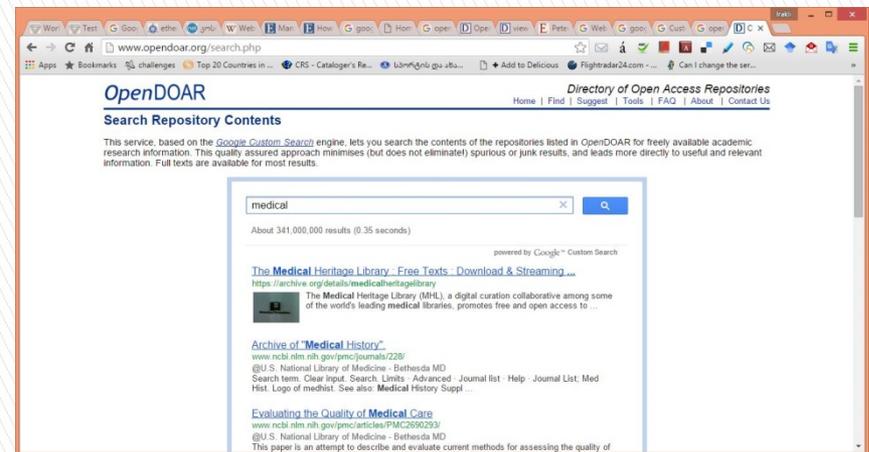
- ▶ OpenDOAR is a Directory of Open Repositories including thousands of OA repositories of Universities, scholarly societies, etc. worldwide.
 - ▶ More than hundred countries
 - ▶ Several languages (English, French, Russian,..)
 - ▶ It includes searching for repositories (where you can find OA repositories worldwide by your interest)
 - ▶ As well as CSE for searching through all these databases
- 

Direct Link to OpenDOAR CSE

http://www.opendoar.org/search.php



The screenshot shows the OpenDOAR search page. The browser address bar displays www.opendoar.org/search.php. The page title is "OpenDOAR Directory of Open Access Repositories". Below the title, there is a navigation menu with links: Home | Find | Suggest | Tools | FAQ | About | Contact Us. The main heading is "Search Repository Contents". A paragraph explains that the service is based on the Google Custom Search engine and provides academic research information. A search bar with the placeholder text "Google Custom Search" and a magnifying glass icon is present. Below the search bar, a note states: "To search for open access repositories rather than their content, please use the Directory page". At the bottom, there is a disclaimer: "This service relies on Google's indexes, which in turn rely on repositories being suitably structured and configured for the Googlebot web crawler. If you are an administrator and your material is not being retrieved, first check that your repository is listed in OpenDOAR. If it is listed, you may need to review your set-up against Google's Guidelines for Webmasters and see the related pages in the Webmaster Help Center, especially the FAQ on how Google crawls sites. There is also excellent advice on How to Facilitate Google Crawling prepared by Peter Suber." The footer contains the copyright information: "© 2006-2014, University of Nottingham, UK. Last updated: 24-Apr-2014".

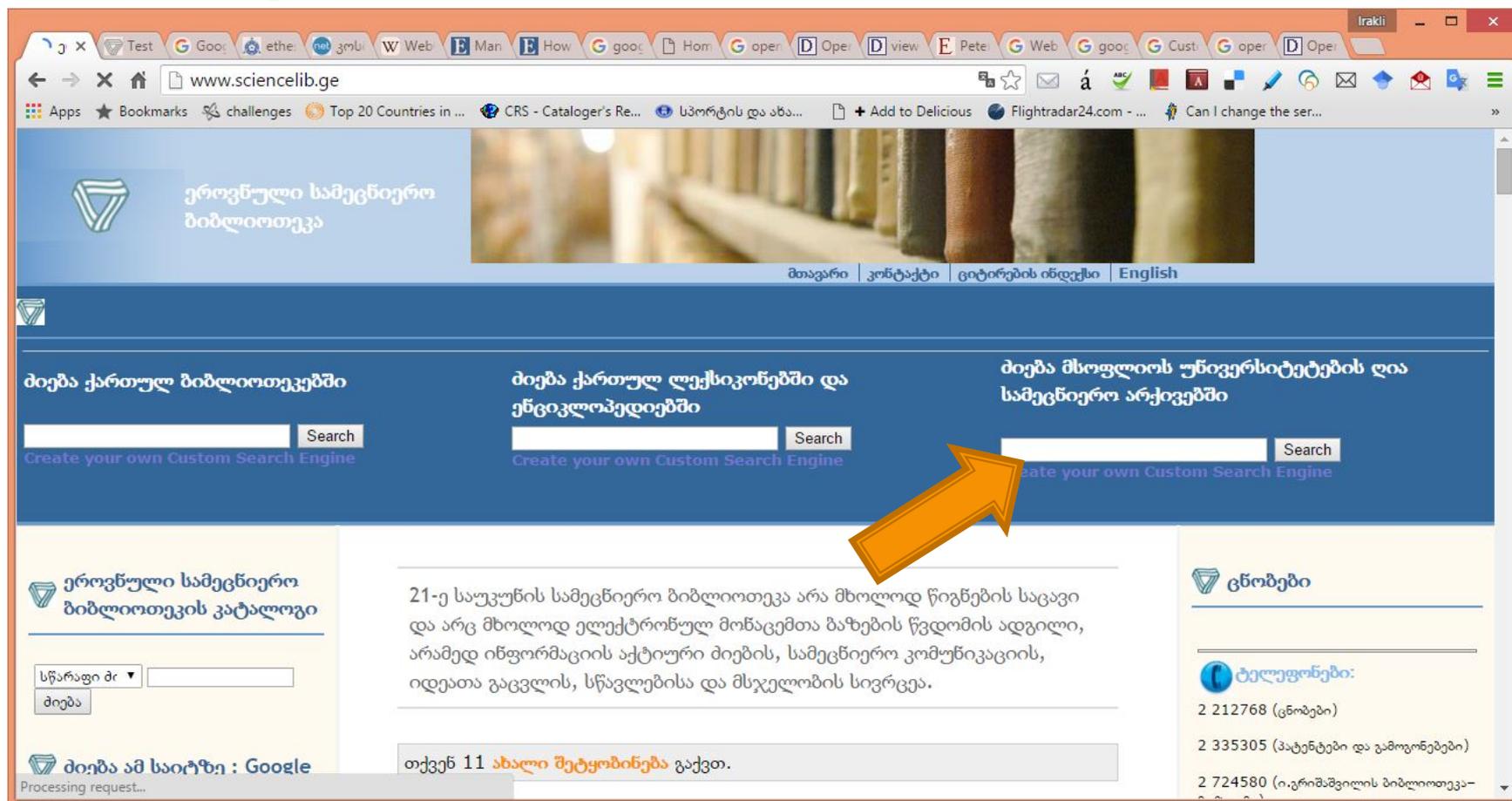


The screenshot shows the OpenDOAR search page with search results for the term "medical". The browser address bar displays www.opendoar.org/search.php. The page title is "OpenDOAR Directory of Open Access Repositories". Below the title, there is a navigation menu with links: Home | Find | Suggest | Tools | FAQ | About | Contact Us. The main heading is "Search Repository Contents". A paragraph explains that the service is based on the Google Custom Search engine and provides academic research information. A search bar with the text "medical" and a magnifying glass icon is present. Below the search bar, it shows "About 341,000,000 results (0.35 seconds)" and "powered by Google Custom Search". The search results list several items:

- [The Medical Heritage Library: Free Texts, Download & Streaming...](https://archive.org/details/medicalheritagelibrary)
The Medical Heritage Library (MHL), a digital curation collaborative among some of the world's leading medical libraries, promotes free and open access to ...
- [Archive of "Medical History"](http://www.ncbi.nlm.nih.gov/pmc/journals/228/)
@U.S. National Library of Medicine - Bethesda MD
Search term: Clear Input Search Limits Advanced Journal list Help Journal List Med Hist. Logo of medhist. See also: Medical History Suppl ...
- [Evaluating the Quality of Medical Care](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2590293/)
@U.S. National Library of Medicine - Bethesda MD
This paper is an attempt to describe and evaluate current methods for assessing the quality of

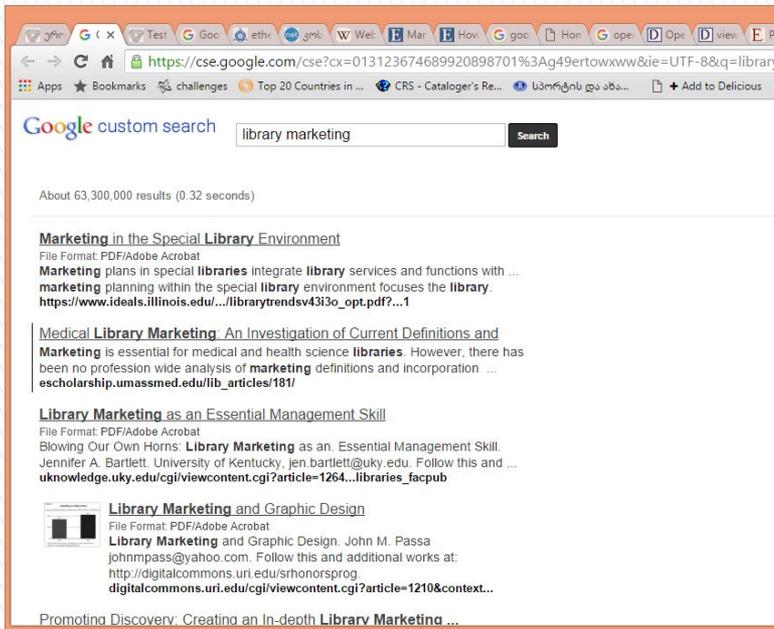


Or you can embed a code to your library website



The screenshot shows the website www.sciencelib.ge in a browser window. The page features a header with the logo and text "ეროვნული სამეცნიერო ბიბლიოთეკა" (National Scientific Library) and a navigation menu with options like "მთავარი" (Home), "კონტაქტი" (Contact), "ციტირების ინდექსი" (Citation Index), and "English". Below the header, there are three search sections, each with a search bar and a "Search" button. The first section is for "ძიება ქართულ ბიბლიოთეკებში" (Search in Georgian libraries), the second for "ძიება ქართულ ლექსიკონებში და ენციკლოპედიებში" (Search in Georgian dictionaries and encyclopedias), and the third for "ძიება მსოფლიოს უნივერსიტეტების და სამეცნიერო არქივებში" (Search in world universities and scientific archives). An orange arrow points to the search bar in the third section. Below the search sections, there is a central text block and a sidebar on the right with a "ცნობები" (News) section. The central text block contains the following text: "21-ე საუკუნის სამეცნიერო ბიბლიოთეკა არა მხოლოდ წიგნების საცავი და არც მხოლოდ ელექტრონულ მონაცემთა ბაზების წვდომის ადგილი, არამედ ინფორმაციის აქტიური მიების, სამეცნიერო კომუნიკაციის, იდეათა გაცვლის, სწავლებისა და მსჯელობის სივრცეა." Below this text, there is a grey box with the text "თქვენ 11 ახალი შეტყობინება გაქვთ." (You have 11 new messages). The sidebar on the right contains a "ცნობები" (News) section with a "ტელეფონები:" (Phone numbers:) sub-section, listing three numbers: 2 212768 (ცნობები), 2 335305 (პატენტები და გამოგონებები), and 2 724580 (ი.გრომშვილის ბიბლიოთეკა-...).

A result screen



- ▶ The code can be produced from OpenDOAR CSE page with just a basic knowledge of software design
- ▶ Or I can share it with you



▶ Thank you!

▶ Questions?

▶ Irakli Garibashvili igar@hotmail.com