

## D4Science Infrastructure - Incident #6316

### gCube WIKI (wiki.gcube-system.org) seems not indexed by Google

Jan 10, 2017 02:48 PM - Massimiliano Assante

<b>Status:</b>	Closed	<b>Start date:</b>	Jan 10, 2017
<b>Priority:</b>	High	<b>Due date:</b>	Jan 13, 2017
<b>Assignee:</b>	_InfraScience Systems Engineer	<b>% Done:</b>	100%
<b>Category:</b>	System Application	<b>Estimated time:</b>	2.00 hours
<b>Target version:</b>	UnSprintable		
<b>Infrastructure:</b>	Production		
<b>Description</b>			
<p>I realised this when trying to look for some of our components together with Luca, none of the WIKI pages were returned, and if you type "gCube WIKI" on google no results points to the most obvious one.</p> <p>Not sure why this is happening, one reason could be the missing robot.txt file: <a href="https://wiki.gcube-system.org/robots.txt4">https://wiki.gcube-system.org/robots.txt4</a></p> <p>It would be VERY beneficial if Google indexed our WIKI.</p>			

#### History

##### #1 - Jan 10, 2017 02:48 PM - Massimiliano Assante

- Description updated

##### #2 - Jan 12, 2017 02:14 PM - Andrea Dell'Amico

- Status changed from New to In Progress

When a robot.txt file is not present the search engine spiders are free to collect and index all the public accessible pages, so it has to be something different.

I verified that no noindex or nofollow directive is present in the local configurations, so there should not be any obstacle to the spiders activity.

I even see some accesses, here are the latest:

```
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [11/Jan/2017:18:45:43 +0100] "GET /index.php?title=Developer_Guide&printable=yes HTTP/1.1" 200 17555 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [11/Jan/2017:20:02:00 +0100] "GET /index.php?title=File:GCore.jpg&action=history HTTP/1.1" 200 21242 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.64.4 - - [11/Jan/2017:20:49:21 +0100] "GET /gCube/index.php/Image:Model.pdf HTTP/1.1" 301 5 "-" "Googlebot/2.1 (+http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.64.62 - - [11/Jan/2017:23:54:48 +0100] "GET /index.php?title=Administrator_Guide&oldid=1722 HTTP/1.1" 200 34730 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.64.4 - - [12/Jan/2017:01:04:39 +0100] "GET /index.php?title=Developer_Guide&action=history HTTP/1.1" 200 47301 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [12/Jan/2017:02:20:25 +0100] "GET /index.php?title=Status_%26_Downloads&oldid=1710 HTTP/1.1" 200 13223 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.32 - - [12/Jan/2017:03:36:12 +0100] "GET /index.php?title=Administrator_Guide&action=edit HTTP/1.1" 200 27179 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [12/Jan/2017:04:52:14 +0100] "GET /index.php?title=Administrator_Guide&action=info HTTP/1.1" 200 14007 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.32 - - [12/Jan/2017:06:07:47 +0100] "GET /index.php?title=Status_%26_Downloads&action=edit HTTP/1.1" 200 11872 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [12/Jan/2017:06:11:02 +0100] "GET /gCube/Index.php/Image:Model.pdf HTTP/1.1" 301 178 "-" "Googlebot/2.1 (+http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.63 - - [12/Jan/2017:08:39:20 +0100] "GET /index.php?title=Status_%26_Downloads&action=info HTTP/1.1" 200 13917 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html)"
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [12/Jan/2017:09:55:07 +0100] "GET /index.php?title=Administrator_Guide&printable=yes HTTP/1.1" 200 33273 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.go
```

```
ogle.com/bot.html) "
gcore.wiki.gcube-system.org_access.log:66.249.76.32 - - [12/Jan/2017:11:10:54 +0100] "GET /index.php?title=Fau
lt_Management&action=edit HTTP/1.1" 200 10993 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.
com/bot.html) "
gcore.wiki.gcube-system.org_access.log:66.249.76.32 - - [12/Jan/2017:11:27:16 +0100] "GET /apple-app-site-asso
ciation HTTP/1.1" 301 5 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html) "
gcore.wiki.gcube-system.org_access.log:66.249.76.34 - - [12/Jan/2017:11:27:17 +0100] "GET /gCube/Apple-app-sit
e-association HTTP/1.1" 404 10313 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.com/bot.html
) "
gcore.wiki.gcube-system.org_access.log:66.249.76.63 - - [12/Jan/2017:12:27:13 +0100] "GET /index.php?title=Res
ource_Model&printable=yes HTTP/1.1" 200 10774 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google.
com/bot.html) "
gcore.wiki.gcube-system.org_access.log:66.249.66.239 - - [12/Jan/2017:13:42:49 +0100] "GET /index.php?title=0.
6_Release_Notes&oldid=1571 HTTP/1.1" 200 19230 "-" "Mozilla/5.0 (compatible; Googlebot/2.1; +http://www.google
.com/bot.html) "
```

### #3 - Jan 21, 2017 12:29 AM - Andrea Dell'Amico

One thing I noticed is that the wiki URLs are not referenced from anywhere. Not even [www.gcube-system.org](http://www.gcube-system.org) has a link to the wiki. Or at least I cannot find one.

I've found this instead, still alive: <http://wiki.gcore.research-infrastructures.eu/>. It needs to die, I guess?

### #4 - Jan 23, 2017 10:13 AM - Leonardo Candela

This page of the gCube site contain three links to the wiki <https://www.gcube-system.org/documentation>

Re wiki.gcore.research-infrastructures.eu I guess it is still there because we are a bit nostalgic

### #5 - Feb 02, 2017 01:30 PM - Andrea Dell'Amico

Leonardo Candela wrote:

This page of the gCube site contain three links to the wiki <https://www.gcube-system.org/documentation>

Re wiki.gcore.research-infrastructures.eu I guess it is still there because we are a bit nostalgic

But it points to a wrong url, or am I wrong?

### #6 - Feb 02, 2017 01:33 PM - Andrea Dell'Amico

- Status changed from *In Progress* to *Feedback*

- % Done changed from 0 to 100

A problem with the sites ssl configuration emerged, see [#6782](#).

It manifested itself with the latest chrome only, but I guess that the google spider could be affected by the same behaviour. The sites configuration is now fixed, let me know if google starts to index the sites.

### #7 - Feb 07, 2017 06:44 PM - Andrea Dell'Amico

- Status changed from *Feedback* to *Closed*

Andrea Dell'Amico wrote:

A problem with the sites ssl configuration emerged, see [#6782](#).

It manifested itself with the latest chrome only, but I guess that the google spider could be affected by the same behaviour. The sites configuration is now fixed, let me know if google starts to index the sites.

It was that indeed, now the wiki has been harvested and indexed correctly.