<http://mail-archives.apache.org/mod_mbox/lucene-dev/200110.mbox/%3C3BC7584A.40107@earthlink.net%3E>

From dmitrys@earthlink.net Fri Oct 12 20:53:40 2001

Return-Path: <dmitrys@earthlink.net>

Mailing-List: contact lucene-dev-help@jakarta.apache.org; run by ezmlm

Delivered-To: mailing list lucene-dev@jakarta.apache.org

Received: (qmail 58047 invoked from network); 12 Oct 2001 20:53:40 -0000

Received: from relay3.uswest.net (HELO relay1.uswest.net) (63.226.138.11)

by daedalus.apache.org with SMTP; 12 Oct 2001 20:53:40 -0000

Received: (qmail 5738 invoked by uid 0); 12 Oct 2001 20:53:45 -0000

Received: from unknown (HELO earthlink.net) (65.100.117.194)

by relay3.uswest.net with SMTP; 12 Oct 2001 20:53:45 -0000

Message-ID: <3BC7584A.40107@earthlink.net>

Date: Fri, 12 Oct 2001 14:53:30 -0600

From: Dmitry Serebrennikov <dmitrys@earthlink.net>

User-Agent: Mozilla/5.0 (Windows; U; Windows NT 5.0; en-US; rv:0.9.2) Gecko/20010726 Netscape6/6.1

X-Accept-Language: en-us

MIME-Version: 1.0

To: lucene-dev@jakarta.apache.org

Subject: Question about Filter

Content-Type: text/plain; charset=us-ascii; format=flowed

Content-Transfer-Encoding: 7bit

X-Spam-Rating: daedalus.apache.org 1.6.2 0/1000/N

I though I understood the Filter idea, but it turns out that I had not.

I checked the FAQ and found no useful information in there... Can anyone

provide more information on how it is supposed to be used, please?

Here's what I thought:

- that Filter contains a BitVector that could be used to

"pre-select" documents by number, so that some documents could be

quickly excluded without even evaluating the query on them.

- that a Filter can be created once (expensive) and then be used

with multiple queries (cheap)

The first point seems true, but the second I'm not sure about. The

Filter's interface provides a BitVector given an IndexReader. This seems

to mean that I have to re-create the BitVector every time... Maybe, I'm

supposed to remember the bitVectors created inside the filter (keyed by

IndexReader) and then check if IndexReader has been updated since I last

saw it? Is this the idea? In that case we probably need a lastModified()

method on the IndexReader. The ones that are there are static and

require a Directory or a file.