<http://mail-archives.apache.org/mod_mbox/lucene-dev/200111.mbox/%3c5.1.0.14.2.20011101140038.03239b30@pop.quiotix.com%3e>

From lucene-dev-return-309-qmlist-jakarta-archive-lucene-dev=jakarta.apache.org@jakarta.apache.org Thu Nov 01 22:10:09 2001

Return-Path: <lucene-dev-return-309-qmlist-jakarta-archive-lucene-dev=jakarta.apache.org@jakarta.apache.org>

Delivered-To: apmail-jakarta-lucene-dev-archive@apache.org

Received: (qmail 33664 invoked from network); 1 Nov 2001 22:10:09 -0000

Received: from unknown (HELO osaka.betaversion.org) (192.18.49.133)

by daedalus.apache.org with SMTP; 1 Nov 2001 22:10:09 -0000

Received: (qmail 27102 invoked from network); 1 Nov 2001 22:12:25 -0000

Received: from nagoya.betaversion.org (192.18.49.131)

by osaka.betaversion.org with SMTP; 1 Nov 2001 22:12:25 -0000

Received: (qmail 3896 invoked by uid 97); 1 Nov 2001 22:10:10 -0000

Delivered-To: qmlist-jakarta-archive-lucene-dev@jakarta.apache.org

Received: (qmail 3878 invoked by uid 97); 1 Nov 2001 22:10:08 -0000

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

Precedence: bulk

List-Unsubscribe: <mailto:lucene-dev-unsubscribe@jakarta.apache.org>

List-Subscribe: <mailto:lucene-dev-subscribe@jakarta.apache.org>

List-Help: <mailto:lucene-dev-help@jakarta.apache.org>

List-Post: <mailto:lucene-dev@jakarta.apache.org>

List-Id: "Lucene Developers List" <lucene-dev.jakarta.apache.org>

Reply-To: "Lucene Developers List" <lucene-dev@jakarta.apache.org>

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

Received: (qmail 3804 invoked from network); 1 Nov 2001 22:10:08 -0000

Message-Id: <5.1.0.14.2.20011101140038.03239b30@pop.quiotix.com>

X-Sender: popbrian@pop.quiotix.com

X-Mailer: QUALCOMM Windows Eudora Version 5.1

Date: Thu, 01 Nov 2001 14:10:06 -0800

To: Doug Cutting <DCutting@grandcentral.com>

From: Brian Goetz <brian@quiotix.com>

Subject: RE: QueryParser should now support non-ASCII characters

Cc: "Lucene Developers List" <lucene-dev@jakarta.apache.org>

In-Reply-To: <4BC270C6AB8AD411AD0B00B0D0493DF0EE7CD9@mail.grandcentral.c

om>

Mime-Version: 1.0

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

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

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

>No. I too have reservations about these and about fuzzy queries, both

>because they make the query syntax even more complicated, and because they

>make it easy for folks to compose really expensive queries. What are your

>thoughts on the matter?

My general thoughts on this sort of thing, after having been through the

same process many times in other O/S projects, is that there should be a

mandatory one-month discussion period ("cooling-off period"?) before adding

new syntax to the query language. I realize that sounds a little silly,

but the alternative is (a) a proliferation of syntax, which complicates the

project and its use, and (b) a poor choice of syntax, as the "good" symbols

get grabbed first, regardless of whether, for example, curly braces would

be better and more intuitive for some other feature.

This has been a huge issue in the WebMacro group, and I've lately taken the

"Read my lips, no new syntax" position, which, despite lots of moaning, has

actually caused us to find much better solutions once the knee-jerk option

is removed.

I don't want us to lose sight of the fact that the query language should be

easy to use by beginners. Any syntax that has the chance of being confused

by a beginner is probably a bad syntax. The +, -, and quote syntaxes are

good; the ^ syntax is more confusing but less likely to be accidentally

used; I'm less sanguine about the braces and brackets, and even less so

about question marks.

That said, I wouldn't mind seeing a syntax for modifying the slop factor in

a phrase query ("Lucene" within 10 words of "search") :)

--

Brian Goetz

Quiotix Corporation

brian@quiotix.com Tel: 650-843-1300 Fax: 650-324-8032

http://www.quiotix.com

--

To unsubscribe, e-mail: <mailto:lucene-dev-unsubscribe@jakarta.apache.org>

For additional commands, e-mail: <mailto:lucene-dev-help@jakarta.apache.org>