



‘Als je onbekend bent met Solr, moet je dit boek hebben gelezen’

Meer dan ooit tevoren gaat het in onze industrie om het snel ontsluiten van massief grote datacollecties. En wat blijkt? In veel gevallen is de traditionele manier om data te doorzoeken niet langer de meest efficiënte manier. Er is iets anders nodig. Solr probeert een deel van het antwoord te leveren.

Solr is een Apache-project dat voortkomt uit Lucene. Lucene is een populaire en inmiddels al meer dan 10 jaar oude zoekmachine, gebouwd in Java. Solr bundelt Lucene met een verzameling uitbreidingen en andere Lucene-projecten in een samenhangend stukje middleware. Zie het maar zo: waar Lucene een transparante blokkendoos is, is Solr toch vooral een blackbox gebouwd uit Lucene-bouwstenen. Solr is een product dat je installeert, Lucene is een library die je meebundelt met je eigen software.

In ‘Loosely Coupled’ stelt Doug Kaye dat een oplossing definitief is gearriveerd als O’Reilly er een boek over heeft gepubliceerd. Het is dan weliswaar geen O’Reilly boek, maar met de publicatie van ‘Solr 1.4; Enterprise Search Server’ lijkt Solr in ieder geval een nieuwe stap op de technology adoption curve te maken. Het vult zeker een leemte op.

‘Solr 1.4: Enterprise Search Server’ is op een logische manier opgebouwd. Je kunt het zonder probleem van voor naar achteren lezen zonder de draad kwijt te raken, maar het leent zich ook prima als offline referentie documentatie; het is zeker geen slepende tutorial, en ieder hoofdstuk beschrijft één specifiek en afgerond Solr-onderwerp tot in detail.

Om een gevoel te geven voor de scope van het boek: het omvat de configuratie, deployment, de verschillende vormen van text analyse die Solr out of the box ondersteunt, verschillende indexerings mogelijkheden, verschillende integratiemogelijkheden, faceting, en tenslotte - last but not least - scalability.

Zoals gezegd, zo op het oog is het een prima boek. Toch zijn er onderdelen waarvan je zou willen dat er wat meer aandacht aan zou zijn besteed, zoals bijvoorbeeld het gebruik van multi-value velden, om maar eens een dwarstraat te noemen. Wordt de volgorde gewaarborgd? Het is op basis van dit boek

niet te zeggen. Kun je één specifieke waarde in de lijst wijzigen? Het boek geeft geen antwoord.

Iets anders dat in het boek onderbelicht blijft is de extensibility van Solr zelf. Dat is misschien het enige belangrijke onderwerp waarvoor geen apart hoofdstuk is gereserveerd. Er wordt af en toe gezinspeeld op mogelijkheden om een eigen custom component te implementeren, maar in veel gevallen blijft het daarbij.

‘Solr 1.4, Enterprise Search Server’ is een prima boek. Als je onbekend bent met Solr, dan is dit het boek dat je moet hebben. Het is een prima introductie, en het beantwoordt veel vragen. Het roept aan de andere kant ook wel eens vragen op, maar het biedt dan in ieder geval een aangrijpingspunt om op het web verder te zoeken.

Ik zou me kunnen voorstellen dat het boek ook wat te bieden heeft voor meer gevorderde gebruikers. Dat gevoel blijft in ieder geval hangen na het lezen van sommige delen van het boek, zoals bijvoorbeeld het deployment hoofdstuk.

Sterker nog, er is eigenlijk geen alternatief, behalve urenlang rondwaren op de Solr web site, en dat is geen aantrekkelijke optie. Voor de volledigheid is het goed om er bij te zeggen dat het boek is gepubliceerd, voordat Solr 1.4 goed en wel was gereleased, en dat was misschien wat al te gretig, maar tot problemen heeft dat zo op het oog niet geleid. «

Titel: Solr 1.4: Enterprise Search Server
Auteur: David Smiley en Eric Pugh
ISBN-13: 978-1-847195-88-3
Uitgever: Packt Publishing

Waardering



Wilfred Springer
is Chief Scientist bij Xebia.