

contentstrategy

“Zoeken en Vinden met Lucene”

Seminar *Open Source* enterprise search
en document management

Gilbert Groenendijk

&

Victor van der Lelij

Reekx



Zoeken en Vinden met Lucene

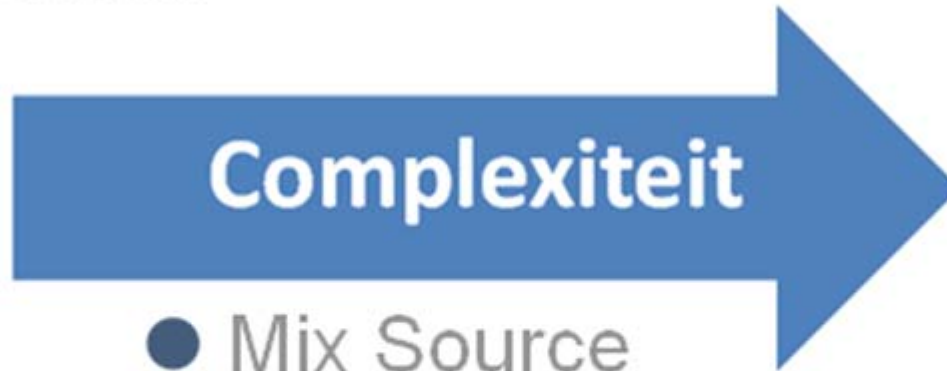
Open Source oplossingen op het gebied van Enterprise Search zijn goed te integreren met bestaande informatiesystemen



- *Open Source* oplossingen
 - Ervaring opdoen met het ontsluiten van informatiebronnen via zoekmachine
 - Niet meteen een keuze hoeven te maken voor een software leverancier
- Er zijn meerdere oplossingsmogelijkheden

Oplossingsmogelijkheden

- Open Source



- Mix Source

- Closed Source



Open Source

- Software is kosteloos te gebruiken
 - Het ideaal is: Verbeteringen en toevoeging aan *Open Source* producten te delen met *Open Source* gemeenschap
- *Open Source* software vraagt om sterke implementatie partij.
 - Er kan niet worden teruggevallen op een software leverancier

Closed Source

- Software leveranciers hebben belang bij een goed product
 - Hoe groter de lijst met requirements hoe complexer de oplossing
 - Meer requirements vs. *Open Source*
 - Meer maatwerk
 - Voor complexe oplossingen, *waar productgaranties een grote rol spelen*, is de keuze voor *Closed Source* sneller gemaakt

Best of both Worlds

- Mix Source
 - Ben bewust van de sterke en zwakke kanten van *Open Source* producten.
- Maak gebruik van beide werelden
- Lucene is een goede zoekmachine
 - Voor zoeken heb je wel content nodig
- Connectors, het extraheren van content is nog steeds een vak apart

Proof of the pudding

- Toepassen van *Open Source* oplossingen in bestaande informatieomgevingen



SharePoint

Meets



open source

Reekx

Organisatie X



- Denkbeeldige Organisatie X
 - Informatiehuishouding bestaat uit
 - SharePoint (als DMS)
 - Gezamenlijke netwerk directory (bv. G-schijf)
 - En vele andere informatiesystemen...
 - Eis ontsluiting via **één** zoekschil
 - Wens van Organisatie X, gebruik zoveel mogelijke *Open Source*



- Gekozen oplossing
 - Lucene als zoekmachine voor alle bronnen
 - SharePoint (Stukje maatwerk voor koppeling)
 - Gezamenlijke netwerk directory (Crawler)

Zoeken en Vinden met Lucene

Seminar Open Source enterprise search en document management

Zoekvraag

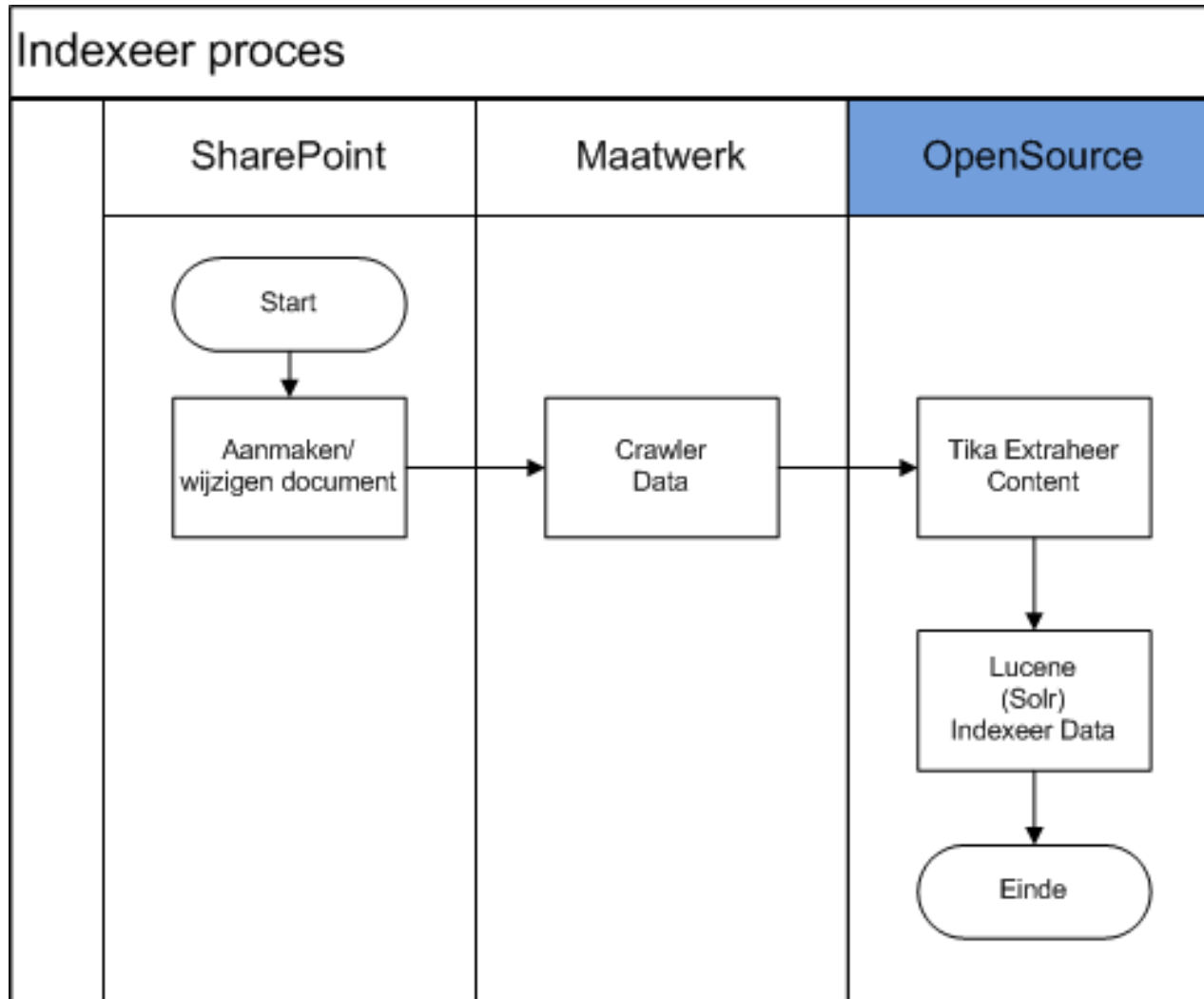
Vrije tekst

Titel

Bron

Zoeken

Indexeerproces - Demo





contentstrategy

Organisatie X



● Demo

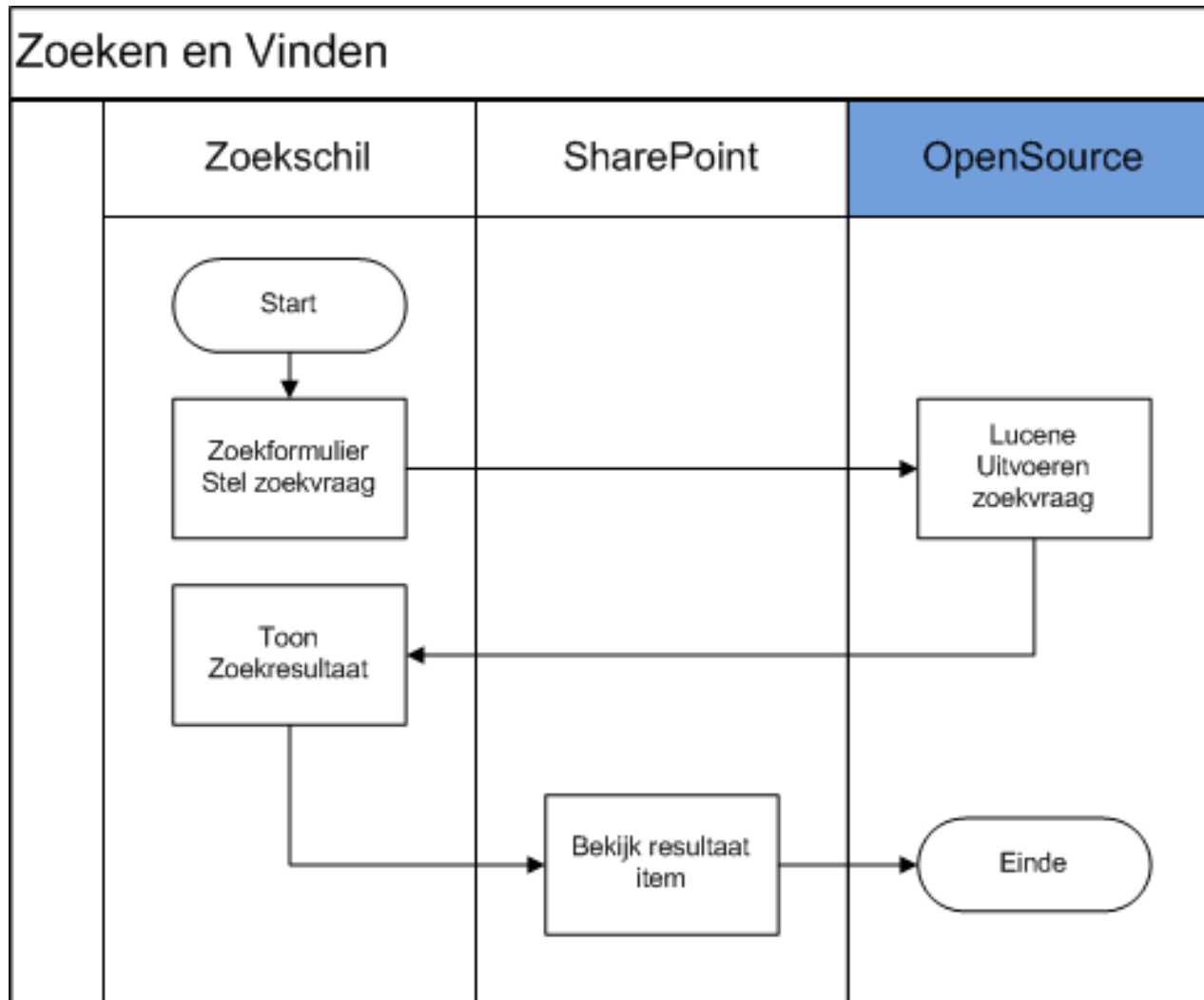
- Introductie zoekschil
 - Stellen van een zoekvraag
- Als informatiemedewerker content aanmaken in DMS (SharePoint)
- Terug naar de zoekschil
 - Medewerker vraagt zich af of zijn document al is terug te vinden

Zoeken en Vinden met Lucene

Seminar Open Source enterprise search en document management

Zoekresultaat

Zoeken en Vinden



Proof of the pudding



- Proof of the pudding: *Het recept*
 - Ontsluiten informatiesystemen
 - Vergaren van content t.b.v. zoekmachine
 - Vergaarde content omzetten in indexbestanden
 - Zoeken in alle data met Lucene
 - Resultaten terug tonen in eigen informatiesystemen

Vragen

