



お客様事例

株式会社セブンネットショッピング

オープンソース検索エンジン Solr/Lucene の採用で、検索精度とパフォーマンスの大幅な向上を実現。

The screenshot shows the homepage of the 7net shopping website. At the top, there's a navigation bar with links for 'セブンネットショッピング' (7net Shopping), 'ログイン' (Login), '会員登録(無料)' (Free Member Registration), 'マイページ' (My Page), 'カートを見る' (View Cart), and 'お届け状況' (Delivery Status). Below the navigation is a search bar with placeholder text 'すべての商品' (All Products) and a magnifying glass icon. A banner at the top right says '宅配送料無料!' (Free Shipping!) and 'おまけ! 受取りも送料・手数料無料' (Free shipping and handling fees for pickup). The main content area features a '商品カテゴリ' (Product Category) sidebar with links for 書籍, 雑誌, CD, DVD・Blu-ray, TV・携帯ゲーム, 楽器・ホビー, 文具・事務用品, 金品, ドリンク・お酒, ホーム・キッチン, コスメ・ヘルスケア, ファッション, ベビー・キッズ, 食品・ギフト, and WiFiサービス「セブンスポット」. To the right, there are several product cards with titles like 'いらっしゃいませ。会員の方はログインしてください。初めての方はご新規登録! 2000円以上お買い上げで、nanaco1000ポイントが当たるチャンス!', '新しい年、新しい自分、新しい考え方を見つける! 冊', and '映画『マジック・ツリーハウス』の主題歌! 植村花菜のニューシングル発売!'. On the far right, there are promotional sections for 'goodday 2012 SPRING' (featuring a striped orange and yellow jacket) and 'WINTER FAIR 開催中!' (featuring a yellow puffer jacket). There's also a 'S.M. ENTERTAINMENT' section featuring 'AKB48'.

セブンネットショッピング

ショッピング こだわり専門店 みんなのクチコミ ネットスーパー e.デパート トラブル チケットひあ 検索 メール登録 マイページ カートを見る お届け状況

すべての商品

宅配送料無料! 受取りも送料・手数料無料

いらっしゃいませ。会員の方はログインしてください。初めての方はこちら

2000円以上お買い上げで、nanaco1000ポイントが当たるチャンス!

商品カテゴリ

書籍 雑誌 CD DVD・Blu-ray TV・携帯ゲーム 食品・ホビー 文具・事務用品 金品 ドリンク・お酒 ハーム・キッズ コスメ・ヘルスケア ファッション ベビー・キッズ 産直・ギフト

選択の科学

選ばれの科学。コロナ大変
学び直ススクール特別編
著者 和田裕美 / 著
シーカー・エインガー / 著
横井祐子 / 著
1,700円

誰でもリーダーになれる3つの
内実 サイバイン
著者 木村義洋 / 著
1,300円

カモメになったペンギン
ジョンP.コターノ著
ホルガース・スザン著
周和博訳 西村辰氏著
1,250円

選ぶのは自分自身
選ぶのはあなた
著者 佐藤志
1,050円

◎ その他の自己啓発に関する本はコチラ

映画「マジック・ツリーハウス」の主題歌! 植村花菜のニューシングル発売!

人気のランキング

1時間ごとに更新

基準 石丸 斎藤 鮎川 性別 性別

人気の専門店

1. 植村花菜

セブンプレミアム

セブンネットショッピング

会員登録(無料) ログイン 会員サービス ご利用が
簡単! お届け地図 マイページ カートを見る お届け状況

ツイートする ブックマーク

good day
2012 SPRING

詳しくはこちら

SM
WINTER FAIR
開催中!

詳しくはこちら

AKB48
オフィシャルグッズ、
限定商品、販売中!
詳しくはこちら

正月大感謝!

ビカビヤロウ
正月太りしおろす解消! せんか?
人気のダイソーカード部分
ダイヤドームで貰ふと合った大作
上手撮影いたします。

詳しくはこちら

わたしのクチコミ
クチコミとは?

プロフィール

株式会社ヤブンインターネットラボ



ソーシャルソリューション本部
システム2部
シニアリーダー
北林洋太さん



ソーシャルソリューション本部
システム 1 部
リーダー
都筑和仁 様

〒100-0005 東京都千代田区丸の内 1-8-3 丸の内トラストタワー本館 20 階

T:03-5288-5927

E-03-5288-5928



ソーシャルソリューション本部
システム2部
システム3チーム チーフ
本間文康さん



ソーシャルソリューション本部
システム2部
システム3チーム
田中晴菜さん

検索精度とパフォーマンスの向上を目的としたリニューアルを実施。

■今回のプロジェクトの概要について教えてください。

北林／セブンネットショッピングはセブン＆アイ・ホールディングスの総合ネット通販として 2009 年 12 月にサービス・イン。アイテム数約 500 万、会員数 900 万人という国内有数の規模を誇る EC サイトです。検索エンジンとして、当初は国産の商用ソフトを採用していましたが、検索精度やパフォーマンスの改善を課題としていました。

具体的には文章解析に N-gram 方式のものが用いられていたため、お客様がサザンオールスターズの CD を購入しようとして「サザン」と入力しているのに「サザンカ」に関連する商品が上位に表示されてしまっていたのです。

また画面生成のために多数のクエリーが走る構造だったため、検索結果の表示に時間を要することがありました。

こうした問題を解決するため、新しい検索エンジンの検討に入ったのが 2010 年 11 月のことです。最終的な候補として残ったのは、商用ソフトと、オープンソース (OSS) の Solr/Lucene のふたつ。検索精度、実装の容易性、保守性、コストなどさまざまな観点から比較検証しました。検証環境で実験してみると、Solr/Lucene はサーバーなどのスペックが低くても、パフォーマンスが低下しにくいことが分かりました。

こうしたことから Solr/Lucene の採用が決まり、2011 年 5 月から開発がスタート。11 月 18 日に新しい検索エンジンへの切り替えが完了しました。

OSS ならではの開かれた開発体制に、商用ソフト以上の安心感があった。

■OSS を採用することに不安はありませんでしたか。

北林／商用ソフトは、メーカーという閉じた環境で開発やバージョンアップが行われています。機能を知るためにメーカーが用意した説明書を読むしかありません。

以前の検索エンジンを使った開発では、不明点をメーカーに尋ねてもタイムリーな回答を得られなかつたことがありました。

それに対して、OSS はオープンな環境。特に Solr/Lucene はコミュニティが非常に活発に活動しています。何か疑問点が生じても、マーリングリストに質問を投げれば、誰かが答えてくれます。またバージョンアップもメーカー主導で行われるものではないため、更新の頻度が落ちるということはありません。

EC サイトという性格上、お客様のニーズを仮説検証し、常に辞書登録を更新し、設定をチューニングしていく必要があります。こうした点を総合的に考えると、むしろ OSS である Solr/Lucene を選択するほうが安心感がありました。

都築／もちろん、ロンウイットという会社が存在していたのも安心材料のひとつでした。日本人唯一のコミッターとして Solr/Lucene の開発に参加している関口さんから、直接、トレーニングやサポートを受けることができるわけですから、OSS を不安視する必要はまったくないだろうという判断になったのです。

検索時の負荷を分散するため、複数のサーバーで機能を分担。

■実際の開発においてはどのような苦労がありましたか。

北林／Solr/Lucene を採用したことによる苦労は特にありません。今回、私たちが重視していたのは、とにかくパフォーマンスを向上させること。99%の確率で、1 秒以内に検索結果を返す、というのが目標でした。

そのため、Solr のキャッシュに加えてバックエンドにインデックス生成用の別サーバーを立てるなどの工夫をしています。

今回のプロジェクトでは実装に入る前に 1 カ月半ほどかけてプロトタイプを使って性能検証を行い、その性能検証結果を

受けて、ハードウェア構成を決定するという手順を取っています。

都築／プロトタイプの開発時には、特定のキーワードで検索結果がヒットしない、という問題も起きました。ロンワットに質問を投げたところ、非常に短時間で回答をもらうことができ、設定を変更することで解決しました。また、Solr のマルチコア、マージを活用した検索インデックス切り替え方式の妥当性検証などにもロンワットに支援していただきました。

北林／検索対象、スコアの重み付けなどは、稼動後にも頻繁に変更の必要があると考えていたので、XML ファイルでこれらを変更することで迅速に対応できるしくみにしました。また、カテゴリーごとにサーバーを分けるなど、かなりの部分で分散化を図っています。



Solr の実力を引き出すためにも、丁寧なチューニングは不可欠。

■テスト段階での問題点などはありましたか。

北林／今回、テストには約 2 ヶ月の期間をかけています。関係者のみに公開される検証用の仮サイトを設け、100 名体制で様々なキーワードを使ったテストを実施しました。

ここでも課題になったのは、パフォーマンスの問題。例えば、セブンネットショッピングには専門店というカテゴリーがあるのですが、データが一般商品とは異なる階層にあります。そのため、クエリーが複数、走ることになり、検索結果を表示するまでに時間がかかるてしまう。そこで検索実行を非同期にするなどして、パフォーマンスの向上を図っています。

都築／EC サイトとしての使い勝手の良さを実現するため、テスト段階になって仕様が変更になる部分はかなりありました。

スコアの重み付けなども、当初は Solr がデフォルトで持ってくるデータを重視していたのですが、売れ筋や予約受付中のものを重視すべきだ、といった指摘を受けて設定を変更しています。

Solr/Lucene を採用すれば、それだけで良い検索結果が得られる、というわけではなく、やはりチューニングは必要。むしろ、そうした設定の変更にも柔軟に対応できるという点が、Solr/Lucene を採用したことの大きなメリットでした。

検索結果表示の大幅な高速化を実現。検索精度のさらなる向上を目指す。

■リニューアルによって、どの程度の効果がありましたか。

都築／今回は形態素解析器としてやはり OSS である lucene-gosen を採用しています。採用を決定する際には性能比較を行いました。その結果、lucene-gosen の方が約 3 倍も高速であることがわかりました。

北林／検索結果の表示に 1 秒以上かかることがリニューアル後は 0.4% にまで低減し、高速化を実現しています。

ただし、検索精度を向上させるためには、手作業で辞書登録していくしかない。そこは保守運用のスタッフの尽力によるものです。

田中／辞書登録は、商品名や人名などを中心に追加しており、最終的には lucene-gosen のデフォルト約 49 万語から 70 万語まで拡張しました。その他で苦労したのは、同音異義語ですね。例えば「マツケン」で検索された時、表示すべきなのは松平健なのか、松山ケンイチなのか。お客様のニーズによってソートできるよう、チューニングを加えています。同音異義語だけでも数千個は登録しています。

本間／検索エンジンは、リリースがゴールではなく、スタート。新商品の追加や在庫状況、発売日の変更などがありますから、1 日に数回の検索インデックス更新を行っており、さらに頻度を高めるチューニング中です。

EC サイトでは、検索精度が低いと、リアルの店舗で考えれば、在庫はあるのにお店に並んでいないのと同じことが起こってしまいます。チューニングに関しては、運用のフェーズに入ってからも絶えず見直し続けていくことになるでしょう。

グループ各社への導入も視野に、ノウハウを蓄積していきた

い。

■今後の目標について教えてください。

北林／現在、要望として受けているものにサジェスト機能の実装があります。Solr/Lucene には検索キーワードを集積していく機能がありますから、いわゆる「もしかして機能」なども搭載していきたいですね。

将来的には、検索結果の一覧表示をパーソナライズしていく方向性になると思います。

本間／検索キーワードは、お客様の声とも言えるもの。ECサイトにとっては財産ともいいうべき重要なマーケティングデータです。それを有効に活用できるという点でも Solr/Lucene を採用したことのメリットは大きかったです。

都筑／今回、Solr/Lucene を採用することで、期待した以上の効果をあげることができました。今後は、セブン＆アイ・ホールディングスの他のサイトにも Solr/Lucene を採用していくことも検討しています。その意味でもセブンネットショッピングでの検証を重ね、活用のノウハウを蓄積していきたいですね。

北林／Solr/Lucene で、こんなことができないか、など、オンラインには、色々と無理難題をお願いすることになるかもしれません。新しい機能が実装された時のアナウンスやパッチの更新など、セキュリティ面でのサポートにも期待しています。



株式会社セブンインターネットラボ

代表者 鈴木康弘

業務内容 IT におけるシステム開発、新規ビジネスおよびデータ活用の研究

設立 2009 年 3 月 24 日

資本金 4 億 9,995 万円

株主 株式会社セブン＆アイ・ネットメディア
日本電気株式会社

株式会社セブンネットショッピング



株式会社セブンネットショッピング

代表者 鈴木康弘

業務内容 インターネットによる商品の販売、およびサービスの提供

設立 1999 年 8 月 24 日

資本金 30 億円

株主 株式会社セブン＆アイ・ネットメディア
ヤフー株式会社

株式会社トーハン

鈴木康弘

※記事内のデータは、取材日現在の値です。

(取材日：2011 年 12 月)