

ファイルサーバー検索システムの
効果を考える

株式会社 鉄飛テクノロジー
<http://www.teppi.com/>

目 次

「ファイルサーバー検索システムの効果を考える」

■ 全文検索システムの導入効果ってそんなに小さいの？

- 削減される検索時間だけでは測れません
- 検索時間はむしろ増える、それでいいのです！

■ 検索ができると、整理のやり方が変わる

- 検索システムが可能にする整理法
- 切り口を変えた分類の例
「顧客＞案件」別フォルダ階層を、「期間＞案件」別に再編
- 大雑把な整理整頓の例
ファイルサーバー検索は、「フォルダ検索」にも使える

■ それでも「なくて平気」ですか？

■ ファイルサーバー検索システムの導入コスト・運用コスト トータルで抑えるには？



ファイルサーバー検索システムを開発・販売している私たちがしばしば聞かれる問いに、検索システムの導入効果はどれだけか？というのがあります。全文検索システムの導入には、それなりの費用がかかります。「費用に見合うだけの効果が見込めるのか？」が気になるのは当然です。定量的な効果測定は残念ながら難しいのですが、ファイルサーバー検索システムの導入には必ず大きな効果が伴うと私たちは確信しています。ここでは、その確信に至った理由をご説明します。

全文検索システムの導入効果ってそんなに小さいの？

私たちは 2007 年から 10 年以上にわたってファイルサーバー全文検索システムを開発&販売しています。発売当初は売るのが苦労した時期が続きました。「これ、便利だね！ 欲しいね！ でも、無くても死なないからねえ。会社に必要性を認めてもらうのは難しいなあ。」といった声が少なからずありました。当初の販売で苦戦した一番のポイントは、「効果」を定量的に測定することの難しさでした。

当時、同業他社のホームページを調べると、ファイル検索にかかっている時間を積算して、その合計に人件費単価をかけ合わせた数字を出すというのが一般的でした。ファイル検索 1 回あたり 1 分、毎日 2 回、年間 200 営業日、時給 3000 円で人員数 100 人だと…年間 $2/60 * 200 * 3000 * 100 = 2,000,000$ 円 というような計算です。こうして計算した検索システムによるコスト削減効果は、やけに小さいのです。100 人の組織にとって年間 200 万円の削減では、「無くても死なない」と言われてもしかたありません。どうしてこんなことになったのでしょうか。

削減される検索時間だけでは測れません

私達は、検索エンジンの効果は、削減される検索時間だけではないと考えます。この 10 年間、FileBlog を購入されたお客様は、そこを信じて買ってくださいっており、おかげさまで 200~300 社のお客様から 6 割~7 割以上の率でリピートをいただいています。

私たちは、検索エンジンの効果は、削減される検索時間だけではないと考えます。この 10 年間、FileBlog を購入されたお客様は、そこを信じて買ってくださいっており、おかげさまで 200~300 社のお客様から 6 割~7 割以上の率でリピートをいただいています。



私たちの製品の販売が伸びてきた背景には、Google に代表される検索エンジンを当たり前のも
のとして使い慣れた世代が、お客様の意思決定に中心的に関与するようになったという変化が
あります。検索エンジンのないインターネットなんて、考えることもできない現在、自社のファ
イルサーバーについても、「検索できて当たり前でしょ!？」という考えを持つユーザーが増え
たのです。そういうお客様と話していただんだんわかってきたのは、検索の効果は、「探す」た
めに従来費やしていた時間の単なる削減ではないということです。

検索時間はむしろ増える、それでいいのです！

たとえば Google の効果は、Google の無かった時代に必要だったインターネット上のリサー
チ作業の削減効果でしょうか。とんでもないことです。Google がなければ見つからなかった
ものが見つかるようになったことで、インターネットの利用時間は削減されたどころか、むし
ろ爆発的に増えています。検索エンジンの導入は、そこで探し物をする 1 回あたりの時間を削
減しますが、情報を検索するストレスが小さくなる結果、ユーザーは今までの何十倍も何百倍
もの情報を探すようになり、そこでの情報処理にかかる時間と労力は、むしろ増えたといえる
でしょう。それでも、ほとんどのユーザーは、以前には二度と戻りたくないと思っているはず
です。交通費と労力をつかって「リアルで」行っていた様々な体験が、PC の前に居ながらに
して可能になったことのインパクトは計り知れません。

ファイルサーバー検索エンジンの導入効果も、これと同様です。従来 Windows エクスプロー
ラで行っていたファイル検索作業の削減よりも、お客様に電話、FAX、メールで行っていた確
認作業や、同僚のデスクに出向いて聞きまわっていた労力や、書庫や倉庫や現場に出向いて調
べていた時間など、ファイルサーバーの外で「リアルに」行っていた情報収集活動の時間と労
力こそが削減の主たる対象なのです。

また、整理され蓄積された情報の再利用・活用が増えることや、知識労働者が処理する情報量
が増大することが新たなビジネスチャンスを生むことも、期待されます。たとえば、製品デー
タや案件事例をお客様にすぐに紹介できれば、案件受注につながることもあるでしょう。こ
うして増える売上も、検索システムの効果といえるでしょう。大量の情報にさまざまな切り口で
アクセスできる環境は、新しいアイデアや仮説を生み出しつづけることが求められる知識勞
働者にとって心強い武器になります。そこから始まるイノベーションが、長期的成功の鍵にな
るものです。



検索ができると、整理のやり方が変わる

検索システム導入のもう一つの効果として、私たちがよく目の当たりにするのは、検索システム導入の前後で、文書データの整理整頓方法・管理方法が変わり、文書ライフサイクルの運用全体が大幅に簡略化されることです。整理整頓の簡略化によるコスト削減効果は、しばしば検索コストの削減を大きく上回ります。

私たちは長いこと「整理整頓は大切なもの」と教えられて来ました。整理整頓を必須のものとし、そのために手間暇をいとわない組織文化が作られていることも多いでしょう。しかし、整理整頓の目的は何でしょうか？ 最初にきちんと手間をかけて整理整頓するのは、後で検索するときに探しやすくするためです。検索システムの能力向上によって、探し方が変われば、それに合わせて整理整頓のやり方も変わらなければなりません。

検索システムが可能にする整理法

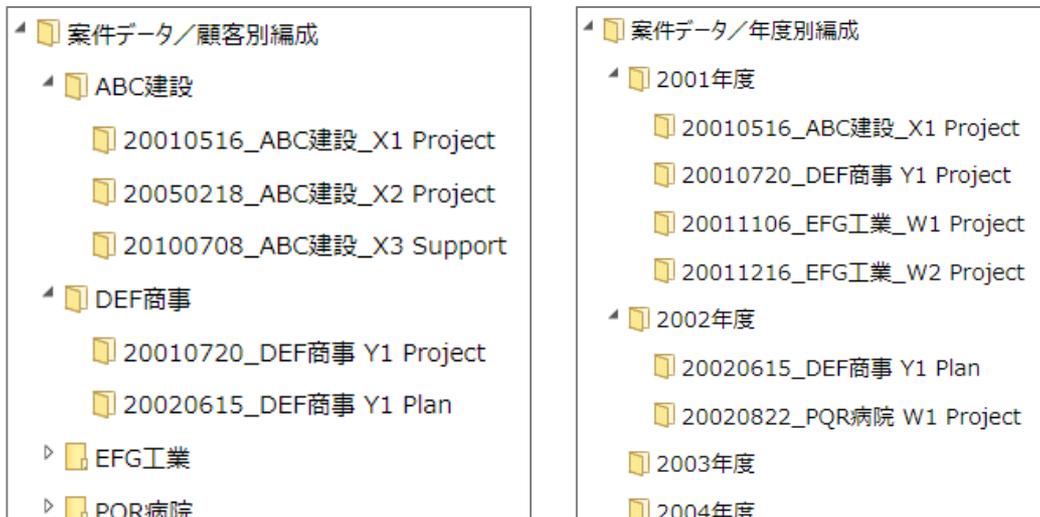
検索システムの導入時に採用して効果的だった整理法の代表として、下記の2例を紹介します

切り口を変えた分類が可能	検索時に、様々な切り口で瞬時に検索できるので 保管時の分類階層を変えることが可能です
大雑把な整理整頓でOK	精密な整理整頓をしなくても、大雑把な整理整頓で すませることができます

切り口を変えた分類の例 - 「顧客>案件」別フォルダ階層を、「期間>案件」別に再編

たとえば、従来の中小企業に多いのが、顧客別フォルダの下に案件別フォルダを作成して、提案書や成果物を管理するという、ファイルサーバのフォルダ階層です。このフォルダ階層のメリットは、それぞれの顧客別に、過去から現在に至るまでの案件の履歴を一覧することが簡単に（Windows エクスプローラだけで）できることです。

しかし、このフォルダ階層には弱点があります。それは、古くなって不要になったデータの削除が難しいことです。古い案件フォルダが複数の顧客フォルダの下に散在しているため、簡単には削除できません。



図：顧客別編成(左)から年度別編成(右)へ

新旧のデータを簡単に分離できるようにするには、年度別のフォルダ編成が有効です。年度別のフォルダ編成の従来のデメリットとして、特定の顧客に関する情報が複数年度に分散していると、簡単に一覧できないというものがありませんでしたが、検索システムがあれば、年度別フォルダ編成にしても、顧客名で検索すれば、瞬時に顧客別の切り口で一覧することが可能です。顧客軸で一覧したり、時間軸で一覧したり、視点を変えてデータを一覧できるという、検索エンジンの能力が、整理整頓方法の再編に生かされる、よい例だといえます。

大雑把な整理整頓の例 –ファイルサーバー検索は、「フォルダ検索」にも使える

ファイルサーバー検索システムの目的は、一般的に「ファイルを見つけること」と考えられていますが、私たちのFileBlogはしばしば「フォルダを見つける」ために利用されています。大量の文書を長期間保管したい動機には、「将来いつ必要になるかわからないが、『念のため』当分の間取っておきたい。」というものがありません。めったに再利用されない文書の保管においては、万が一必要になった時に文書を見つけ出すのは多少面倒であっても構わないので、初期登録の手間を最小限にとどめる必要があります。「大雑把な整理整頓」が合理的な代表的なケースです。

たとえば、デジタルカメラで撮影した画像ファイルなどを大量に扱う場合には、同時に何枚も撮影することが多いため、一つひとつのファイルにいちいち説明を付与するのは面倒です。紙媒体の文書をスキャナで読み込んだPDF文書が大量にある場合も同様です。



こんなとき、多数の画像ファイルやPDF文書のファイル名はデジカメやスキャナーが自動付与したまま、ひとつのフォルダ内に放り込んでしまい、将来の検索に必要なキーワードはフォルダ名やフォルダの属性に入力するという方法が有効です。

登録が楽になる代わりに、検索時にはシステムとユーザーの協働が必要です。検索システムによってフォルダが見つかったら、そこからは人力でフォルダ内のファイルを次々と閲覧して、内容を確認することになります。私たちのファイルサーバー検索システム FileBlog が、ファイルの内容をブラウザ上で閲覧する文書プレビュー機能に力を入れているのは、そのためです。

それでも「なくて平気」ですか？

検索システムの効果は、非生産的な「探し物」に費やされる時間を減らすとともに、「生産的な情報処理」に費やす時間を増やす触媒となることにあります。長い目で見れば、この差は企業の死活を分けるものです。それでも、なくて平気ですか？

ファイルサーバー検索システムの導入コスト・運用コスト。トータルで抑えるには？

ファイルサーバー検索システムの導入・運用において、検索システムのハード・ソフト以外に必要なコストは、ほぼ検索インデックスの構築・保守のコストです。

全文検索エンジンは、Google がインターネットに対して行っているのと同様に、全ファイルの内容を予め読み込んで索引(インデックス)構築することによって、検索を高速化しています。新しいファイルが追加されたり、既存のファイルが追加されたりするたびに、インデックスを更新する必要があり、大量文書を有するファイルサーバーに対していかに効率的にインデックスを保守できるかが、ファイルサーバー検索システムの大事なポイントです。

文書数が 1000 万件未満ぐらいの小規模環境であれば、1 サーバでほぼ問題なく運用でき、オープンソース製品の活用でまかなえるかもしれませんが、数千万・億単位のファイル数をインデクシングするのは、なかなか重い処理となります。たとえば毎週一回実施するはずのクローल処理の所要時間が 1 週間を越えてしまえば、運用は破綻しますから、ハードウェア増設や複数サーバの協調処理などによって処理能力を補う必要があります。スムーズな導入や安定した運用のためには、やはり商用製品を導入する意味があります。

鉄飛テクノロジーの FileBlog は、ファイルサーバー変更検知に基づくリアルタイムのインデックス更新と、週末・深夜にスケジュールされる全フォルダ一括スキャン(クロール)処理との、二本立てでインデックスを保守するため、1 台のハードウェアでも 1000 万文書を越える大規模に対応可能なうえ、複数台分散協調動作によって理論的には無制限の文書数に対応できる、ファイルサーバー全文検索パッケージシステムです。1 ファイルサーバーの特定のフォルダのみの小規模から導入できて、複数ファイルサーバーにまたがる数千万文書規模の大規模にも対応できる検索システムですから、当初の導入試験・使い勝手の評価段階から、最終的な全社展開まで、無理なく段階的に導入可能です。