

文書を可視化する高速表示ソリューション “DATA-COSMOS”

“DATA-COSMOS”: High-speed Viewing Solution to Make Documents Visible

永 田 紀 光

要 約 現在、世界では多数の文書管理システムが存在する。日々更新される文書を時系列に沿って系統立てて管理するドキュメントシステムである。多くの利用者の細かい要望に応えるために改良が加えられ、今後も有望なツールとして期待できるソリューションである。近年ではセキュリティ機能が必須となり、内部統制需要の影響により、大規模向け市場の拡大が続いている。また、文書管理システムは広義ではナレッジ管理システムとしての側面もあるため、企業の知的財産を有効活用するツールとして期待されている。

CMS (Content Management System) やポータル等のアプリケーションと連携して広がりを見せる一方、増大する文書数により、利用者が目的の文書にたどり着くまでの時間が長くなっているのも事実である。そのため、対象コンテンツの視認性や表示応答速度の向上、全文検索などの追加機能が今日の文書管理ソリューションには必要不可欠な要素となっている。

本稿では、目的の文書ひいては情報を得るまでの時間が増加している現状の課題に対し、文書管理ソリューション “DATA-COSMOS” を取り上げ、その特長を示すとともに、文書の可視化、視認性及び高速表示についての技術的な手法を述べる。

Abstract Currently, there are a lot of document management systems in the world. It is one that is updated every day and is managed systematically in chronological order. It keeps being improved to meet the delicate demand of many users, and it will be the solution which we can expect as a promising tool in future. Recently, it has been asked for the security capabilities, and the demand for it continues to expand for the large-scale system market according to the internally controlled demand. Also, because the document management system contains an aspect of the knowledge administration system in the broad sense, it is expected as a tool to make effective use of the intellectual property of the company.

While the document management system spreads in cooperation with the application such as the CMS or the portal, it is true that a user takes long time to get a desired document due to the increase in the number of documents. As a result, the visibility of object contents and improvement of the response speed by indication, and additional functions such as the full text search become the essential elements in today's document management solution.

In this report, for the present issues that a user takes long time to get desired document and therefore information, we show the document management solution “DATA-COSMOS”, and features, as well as and discuss the technical methods about the visualization of documents and the high-speed indication.

1. は じ め に

文書管理ソリューションは、文書情報を対象とした業務の基盤となるシステムである。企業は文書を保存し、できる限りその文書を有効活用することに努めてきた。文書管理ソリューションは以前から企業生産性を高めるためのツールの一つであり、近年では管理対象のファイル

フォーマットも多様化している。また、ネットワーク上に文書を分散配置できるようになり、他のアプリケーションとの連携も強化され、合わせてセキュリティ機能も強化されていった。その市場規模は、2009年から2013年まで、数量/金額とも約7～15%ずつ増加すると予測される。

今後、文書管理ソリューションは、監査に対する証拠を開示できる仕組みとしての需要、情報共有のための企業ポータルとしての需要、検索エンジンとの連携により個々の社員が求めている情報へのアクセス時間を短縮する機能としての需要が拡大し、全社規模の導入が増加すると予測される。

文書管理ソリューションは、今後も有望なツールとして期待できるソリューションであると同時に、一方では膨大な文書管理の整備の運用・構築が適切に行われていない現状や既存業務フローとの連携における課題等も把握する必要がある。

本稿では、文書管理ソリューション「DATA-COSMOS」の視認性、高速表示の仕組みを中心に解説しながら、上記のような課題に対する一つの解決策を提案する。

2. DATA-COSMOS 概要

文書管理システムには、セキュリティ、改定履歴、権限管理、データ管理、アクセス記録といった機能要素がある。DATA-COSMOSは、これらの機能要素を満たした文書管理ソリューションであるが、以下に述べる動作の特長からナレッジ管理・情報共有ソリューションの意味合いも強い。

動作の特長として、すべてのデータを一元的に扱うための統一中間ファイルフォーマットへ変換する方式を採用している点が挙げられる。これは、開発元であるナレッジオートメーション株式会社^[1]独自のファイルフォーマットであり、高速表示を実現する核となる技術である。利用者は統一中間ファイルに対してアクセスし、表示・編集・印刷・検索等の業務を遂行する。また、統一中間フォーマットに対してアクセスすることにより、異なるファイル形式のデータであってもWindowsエクスプローラ風の簡単な操作と統一された表示を実現している。元データは、運用形態を変更することなく扱うことができるため、運用への影響は最小限に抑えることができる。図1に動作概要を示す。また、本章にて本ソリューションの特徴を解説する。

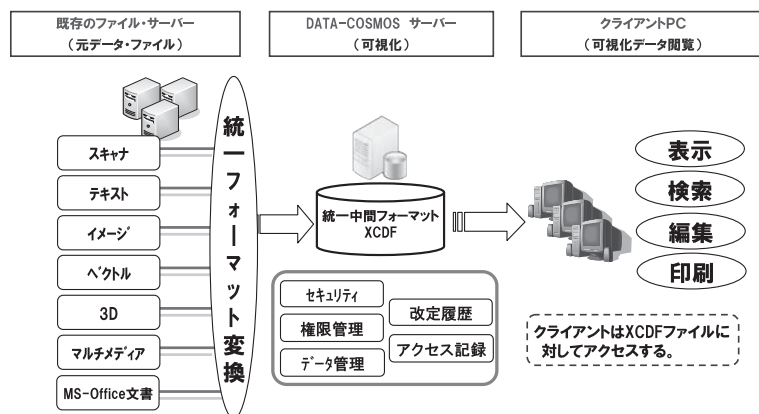


図1 DATA-COSMOS 動作概要

2.1 視認性

文書管理システムという観点において、利用者にとって統一感のある表示で、使いやすく、かつ目で見て目的の文書を素早く探し出せること、すなわち視認性が良いことは非常に重要な要素である。コンテンツの視認性を向上させるために、本ソリューションは以下の方式を採用している。

2.1.1 サムネイル画像

近年の検索エンジン等にも導入されている手法であるが、コンテンツのサムネイル画像^{*1}を使用する方法である。登録されたデータのサムネイル画像を一覧で見ることができ、異なるフォーマットであっても統一感のある表示で視認性を高めている。複数ページにわたる文書の場合には、各ページごとのサムネイルビューが表示されるため、ファイルの視認性だけでなく、各コンテンツ内の視認性も高めている。サムネイル画像は統一中間ファイルへ変換される際に同時に作成され、データ容量も数 KB 程度の大きさなので、表示の高速性を保つ一因となっている。

2.1.2 統一されたファイル操作

異なるアプリケーションのコンテンツを一つのファイルフォーマットに統一することで、同一ウィンドウ内で統一感のある操作・表示を実現している。ウィンドウについては、ツリー表示、サムネイル表示、ファイル属性表示、ミックス表示などの多彩なビューがあり、利用者の用途に応じて選択できる。ファイルの移動や削除、中身の閲覧など、ほとんどの操作が Windows エクスプローラ風の「ドラッグ」や「ダブルクリック」のできるので、利用者に対して教育が必要ないというメリットがある。図2に表示例を示す。



図2 統一感のある表示

2.2 情報漏洩の防止

内部統制・J-SOX 法に対応するための業務プロセスの見直しや、契約社員・派遣社員等の雇用形態の多様化により、企業にとって情報漏洩問題を無視することはできず、それを防ぐこと

は非常に重要となってきた。以前まで海外・国内の関係会社・協力会社へのデータの受け渡しは、メール添付などで安易に行われていたが、今では情報漏洩のリスクとして考えられるようになっていく。

情報漏洩という観点では、通常の Web クライアントがコンテンツを表示する際、一時ファイルに自動的に保存されてしまうため、情報がクライアント端末に残ってしまうというリスクがあったが、DATA-COSMOS では、Web クライアント版（Web ビューワ）と専用ビューワ版のどちらのインターフェイスであっても、キャッシュを自動的に削除するため、そのリスクも最小限に抑えることができる。

アクセス制御に関しては、ユーザ ID 毎に表示のみ可、印刷可等の 7 段階の独自のアクセス制御を適用することができる。また、Active Directory 等の既存の認証システムと連携し、認証・認可情報をそのまま利用することができるため、既存システムの変更を最小限に抑えて連携することができる。アクセス権の設定項目について表 1 に示す。

表 1 アクセス権の設定項目

No.	権限	内容
1	権限なし	ファイル・フォルダが表示されない
2	ビュー	表示のみ可能
3	印刷	印刷まで可能
4	フル持ち出し	全ての情報をシステム外へ持ち出し可能
5	属性編集	検索用キーワードの変更まで可能
6	追加	データ内容の変更まで可能
7	フルコントロール	削除も含む全ての変更が可能

2.3 コスト削減

利用者が多くのアプリケーションを使用するためには、そのアプリケーションの種類分のソフトウェアがクライアント PC にインストールされている必要がある。通常利用者はアプリケーションを起動し、対象のコンテンツを表示・編集することができるようになり、日々の業務を遂行する。このように、利用者の数の分だけソフトウェアライセンス費用がかかることが企業の大きな負担となっている。

表示するだけの使用に限れば、DATA-COSMOS 専用ビューワによって、多くのアプリケーションで作成されたコンテンツも一種類の統一されたファイルとして表示されるため、DATA-COSMOS のビューワライセンスのみで足りる。CAD や Microsoft Office 製品などのアプリケーション毎のソフトウェアライセンスが必要なくなることで、結果的にコスト削減につながる効果がある。Web ブラウザを使った Web ビューワの場合には、ビューワ自体も無償になる。また、ActiveX コンポーネントを許可する必要があるが、アクセス時に自動的にインストールされるため、利用者毎のインストール作業も必要ない。

3. 高速表示の実現方式

前章では、文書管理ソリューションにおけるコンテンツの操作性、応答性、視認性の重要性を述べた。利用者は特に応答性、高速性について重要視し、例えば表示するために時間がかかるシステムでは、そのシステムは利用されなくなる傾向が高く、文書管理システムとしての価値が著しく低下する。

応答性、高速性を高めるためにはいろいろなアプローチがある。ネットワークやストレージ等の最新の高速インフラを利用したり、データ自体を可逆圧縮し軽量化を図るといった方法、コンテンツすべてをメモリに展開するのではなく、オンデマンド転送による必要最小限の部分のみをクライアントへ転送し利用する技術などである。

本章では、DATA-COSMOS の特長である高速表示について、その要素の一つとなる XCDF ファイルフォーマットの長所とオンデマンド転送を解説する。

3.1 データの軽量化

DATA-COSMOS は、全てのデータフォーマットを一元的に扱う独自ファイルフォーマット「XCDF (eXtensible Contents Data Format)」へ変換して利用する。XCDF は、高速表示を実現する核となる技術である。テキスト (文字)、ラスター (画像)、ベクトル (線画)、3次元 (立体)、マルチメディア (動画・音声) の 5 大フォーマット分野をカバーし、XCDF 独自の圧縮技術により、データ自体の軽量化が図られている。コンテンツにも依存するが、大規模ラスター (画像) データ、大規模ベクトル (CAD) データなどは通常の圧縮方法に比べて 3 倍～10 倍まで圧縮される。

XCDF 変換時に合わせて作成されるサムネイル画像も、数 KB 程度の小さいデータであるため、高速表示の一因となっている。例えば、AutoCAD のような CAD データ (ベクトルデータ) は、属性情報等を取り除き、表示するための情報のみを使い、相似形状を一つにして圧縮する技術を用いている。結果として CAD データの高圧縮化が図られている。圧縮しても、元データの表示品質は維持し、高品質を保つ点も重要である。PDF のように変換後、寸法が異なる、線が消える、線が滲むということはない。図 3 に圧縮例を示す。

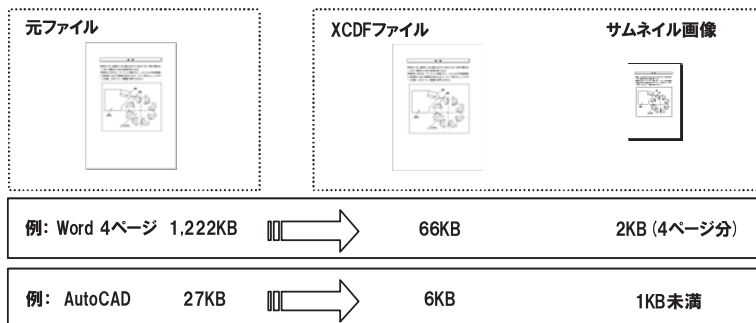


図 3 圧縮例

3.2 オンデマンド転送

ラスターデータの場合は画面表示に必要な最低限の情報のみを転送する。必要な解像度のデータを送る方式の採用で高密度写真の全体表示の場合も高速化される。クライアントの表示要求から最適なデータ精度に圧縮して送る方式により、ファイルサイズに左右されない一定の高速表示を実現する。

Microsoft Word や PDF などの複数ページを持つ文書データを通常のアプリケーションで表示する場合は、ファイル全体をダウンロードした後、アプリケーションを起動する。そのため、数十 MB、数百ページにわたるファイルは、表示するために非常に時間がかかる。DATA-

COSMOS の場合、XCDF 変換により文書データは、各ページに分けて管理される。数十 MB、数百ページのファイルであっても表示する最初の 1 ページ分だけ転送されるため、非常に高速にコンテンツを表示することができる。別ページを選択すれば、そのページ分の情報が転送・表示されるため、ページ単位のデータをオンデマンド転送していると言える。

また、ページ単位の転送のみならず、クライアント側の表示精度に応じた画像データをオンデマンドで圧縮し転送する方式により、巨大な画像の部分拡大時には、その拡大部分だけを抽出転送するなどの「完全部分転送」方式を採用している。サーバには、オリジナルの精度のデータを置いておきながら、瞬時にクライアントが求める精度の画像を生成できる独特のデータ形式にも変換しておく。これも必要最小限のデータをオンデマンド転送するための仕組みである。図 4 に表示例を示す。



図 4 オンデマンドによる転送表示

4. 文書管理システムとしての機能

文書管理システムを構成する要素は、登録、閲覧、編集、検索、保全、認証がお互いに連携されている状態で構成されていることが重要であり、それはソリューション全体の価値を高めることに直結する。いかに表示が高速であったとしても、他要素が適切に連携し、正しく動作しなければ、有効なソリューションとは言えない。

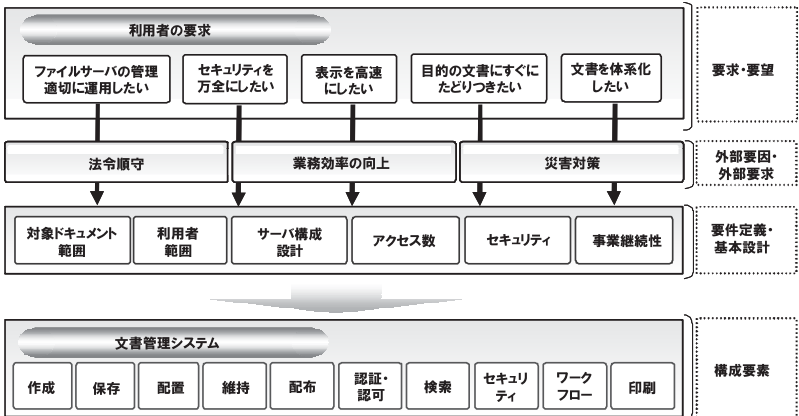


図 5 利用者の要求と文書管理システム

また、それらの要素と企業を取り巻く環境の変化から生じる利用者側の要望をいかにマッチングさせ、使いやすい機能としてソリューションに取り入れるかということが文書管理システム全体の価値を決定づけるポイントになる。図5に利用者の要求と文書管理の関連図を示す。

利用者の要望を汲み取り、フィードバックした機能も含め、それらの要素が付加価値としてどのように構成されているのか、また、どのような仕組みであるのかという観点から、本章では文書管理ソリューションとしての DATA-COSMOS の他の機能を解説する。

4.1 編集機能

DATA-COSMOS は、統一されたインターフェイスとして XCDF ファイルフォーマットへ変換するため、フォーマットに依存しない統一のオペレーションで自由に編集することができる。また、元ファイルを直接編集する機能もある。編集機能として以下が挙げられる。

1) イメージ編集

XCDF ファイルに対して、一つのレイヤーを追加し、コメントの追加、新規オブジェクトの貼り付け、ハイパーリンクや URL のジャンプ設定ができる。

2) ページ編集

異なるコンテンツの特定のページだけを組み合わせると一つの新しい文書を作成することができる。

3) 元ファイル編集

元ファイルに関連づけられたアプリケーションを起動し、元ファイル自体を編集することができる。元ファイルを編集、保存すると自動的に XCDF ファイルへ変換し、元ファイルと XCDF ファイルの相違を防ぐ。

4.2 版管理、アクセス管理

文書管理という観点において、ファイルに対してバージョンをつけて管理する版管理、監査証跡として、ファイルアクセスに関する因果関係を追跡したアクセスログを管理するアクセス管理の機能が含まれている。どちらも新しい技術ではないが、内部統制に対応するために一段と重要度が高まり、第三者への提示が正確に行える、文書管理ソリューションには必要不可欠な機能として実装されている。

4.3 検索機能

コンテンツ量の増大、およびファイルサーバ統合によるサーバあたりのデータ総容量増大に伴い、検索機能は必須であると同時に利用者の生産性向上ツールとして注目されている。通常の文書管理ソリューションが独自に実装した検索機能では、応答性、正確性、拡張性に欠けており、また近年ではファイルのメタ情報のみの検索では利用者の要望を満たすことができず、キーワード全文検索機能の重要度が飛躍的に高まっている。

DATA-COSMOS は独自のファイル属性検索「簡易検索」「複合検索」機能が実装されているが、インデックスを利用した他社の全文検索エンジンと連携して検索結果を出力する機能もあるため、それらを使い分けることで検索効率の向上を目指している。

全文検索エンジンを利用した検索結果は、通常構成において、検索結果タイトル、抜粋等のテキストベースの簡素な作りとなっているが、DATA-COSMOS と連携した検索結果は、検索

キーワードのハイライト、検索でヒットしたページのサムネイル画像、ヒットページ数を表示し、視認性を高め、結果的に利用者が目的の情報にたどり着くまでの時間を短縮することが期待できる。図6に他社検索エンジンと連携させた検索結果表示例を示す。

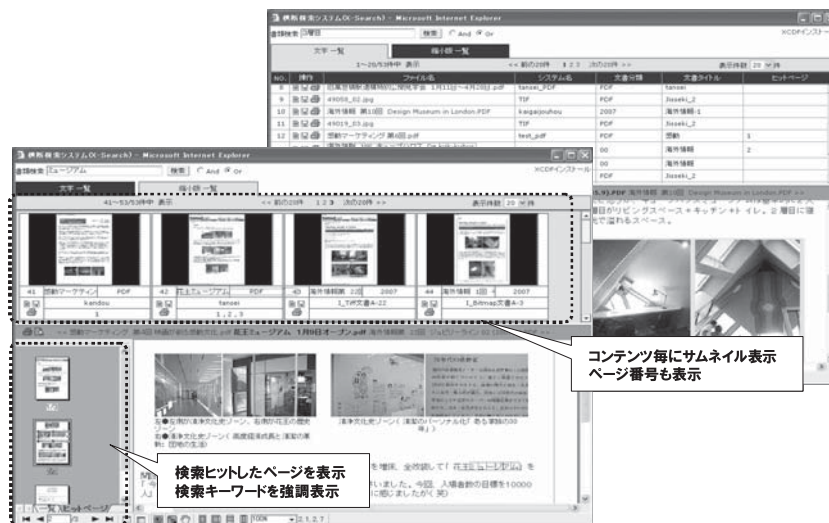


図6 他社の全文検索エンジンと連携させた画面例

DATA-COSMOSがXCDF変換する際に、XCDFファイルとは別に、コンテンツの文字列部分を抽出したテキストファイルを作成する。全文検索エンジンはそのテキストファイルをインデックスすることで対応する仕組みとなっている。そのため、検索結果出力をカスタマイズする必要があるが、基本的にどの全文検索エンジンを利用しても容易にDATA-COSMOSと連携することができるという高い汎用性がある。

4.4 印刷

統一された表示は、ビューだけでなく印刷にも適用されている。CADなどの図面ではA1用紙、MS-Wordなどの文書ではA4用紙のように異なるサイズの用紙が必要な場合においても、連続印刷や振り分け設定により、適切なプリンタ毎に用紙や給紙向きを設定でき、大量に印刷することができる。それぞれのアプリケーションを起動して個別に印刷する必要はなく、ファイルフォーマットが異なるコンテンツであっても一連のジョブとして連続印刷が可能となる。

5. 分析・用途と今後の取り組み

前章まで、本ソリューション「DATA-COSMOS」の特長を中心に解説してきたが、ここで改めて他ソリューションとの違いを比較・分析した。表2に文書管理システム比較表を示す。実際に利用者への提案を考慮すると、本ソリューションが、すべての機能について優れているわけではないが、以下のニーズがある場合には、極めて有効なソリューションである。

- 1) 表示高速性（圧縮・転送技術）
- 2) CAD対応・Webブラウザ対応

- 3) 既存システムとの連携・統合
- 4) ライセンス低減によるコスト削減

表 2 文書管理システム比較

	DATA-COSMOS	文書変換型 他文書管理	他文書管理	PDF応用文書管理
対象データ	○	○	△	△
・Office	○	○	○	○
・各種画像	○	○	○	○
・CAD	○	△	×	△
	2D、3D各種CADに対応	AutoCADの2Dのみ	原則CAD対応不可	プリンタ出力結果のみ対象
・動画音声	○	×	×	×
高速化(圧縮・分割伝送)	○	○	×	×
・ページ単位伝送	○	○	×	△
	指定ページのみを伝送	指定ページのみを伝送	ファイル一括のみ	設定によって可能
・画像可変圧縮	○	×	×	×
	表示側精度に応じて可変	変換時の設定で固定	変換時の設定で固定	変換時の設定で固定
・CADデータ圧縮	○	×	×	×
	精度を保持した圧縮	非圧縮	非圧縮	非圧縮
Web対応	○	○	×	△
・表示	○	○	×	○
	ActiveXプラグインで対応	ActiveXプラグインで対応	Web非対応	ActiveXプラグインで対応
・部分転送	○	○	×	×
	要求部分のみを転送	ページ単位転送	Web非対応	ファイル一括転送
検 索	○	△	△	(各システム個別)
・階層ツリー	○	○	○	—
・キーワード検索	○	○	○	—
・サムネイル	○	○	○	—
・全文検索	○	×	×	—
セキュリティ	○	○	△	(各システム個別)
・アクセス権限管理	○	○	○	—
・アクセス履歴管理	○	×	×	—
・非ファイル転送表示	○	○	×	不可

その中で、それらの特長をもつ利用シーンとしては、以下となる。

- 1) 編集する必要がない文書が多量に存在する。(表示のみで運用されている)
- 2) 利用するファイル形式の種類の多い。
- 3) 利用するファイル容量が大きい。

上記を踏まえたうえで、既存で導入されている文書管理システムを置き換えるのではなく、追加機能としての導入が可能である。もちろん文書管理としての機能(履歴管理、アクセス管理等)も備えている点も考慮すべき重要なポイントとなる。

インテグレータである株式会社ネットマークスは、ネットワークをはじめとする他ソリューションとの連携により、高い付加価値のある独自ソリューションを提供している。今後の取り組みを以下に挙げる。

- 1) WAFS (wide area file services) 製品との連携によるリモートアクセス高速性を高めたソリューション
- 2) 認証ソリューションとの連携によるアクセス管理ソリューション
- 3) 検索エンジン製品との連携による高速全文検索ソリューション
- 4) ストレージ製品との組み合わせによるストレージ管理ソリューション

6. お わ り に

本稿では、文書管理ソリューションDATA-COSMOSについて、その特長を中心に解説した。しかし、「表示は高速だが、XCDF 変換する時間を要する」、「変換した文書のみ文書管理の対象となる」といったDATA-COSMOSの課題も認識しなければならない。

DATA-COSMOSを含め、文書管理というツール・ソリューションが多数存在するということ自体が、利用者にとって唯一の完全なソリューションに成りえていないことを表している。

それだけ文書管理・ナレッジ管理には大きな課題があり、逆に考えると大きな改善を望める市場であることが期待できる。

近年では、文書管理ソリューションは広義でCMS（コンテンツマネジメントシステム）やSaaS（Software as a Service）とも密接に関わり、文書管理としての位置付けに変化が見受けられるようになった。今後は、それらを含めた課題に対応したシステム構築が必要になり、アクセスの急増やシステムのスケーラビリティを考慮しつつ、セキュリティの強化が必要となるであろう。そのためにDATA-COSMOSの形態も更なる発展のために変え続ける必要があり、システム提案、設計、構築ベンダの付加価値を高めていくことも重要であると考え。

* 1 画像や印刷物ページなどを表示する際に視認性を高めるために縮小させた見本のこと。

参考文献 [1] ナレッジオートメーション株式会社 製品情報,
<http://www.ka-j.com/product.html>

執筆者紹介 永田 紀光 (Norimitsu Nagata)

1998年(株)ネットマークス入社。ファイバーチャネルを基盤とするストレージシステムの設計・構築に取り組む。2003年よりGoogle 検索アプライアンス、DATA-COSMOS等の新製品の技術担当に従事。現在、DC技術統括部 第二技術部に所属。

