

プレゼンテーション資料共有システムの一提案

木村聡宏 梶原清彦
NTT ソフトウェア研究所
〒180-0012 東京都武蔵野市緑町 3-9-11
{kimura,kajihara}@slab.ntt.co.jp

概要

プレゼンテーション資料の作成にはコストがかかるが、社内では過去に作成された多量のプレゼンテーション資料を有効活用できていないのが現状である。本稿では、既存の全文検索システムを利用した、プレゼンテーション資料再利用支援システムを提案する。本システムでは、プレゼンテーション資料スライド中のテキストを検索対象としてスライド中の図検索を行い、検索したスライドの縮小イメージを検索結果表示画面に含めることによって、利用者は容易に図を再利用することができる。プレゼンテーション資料の作成において、図の作成には最もコストがかかるので、本システムの利用によりプレゼンテーション資料作成コストの削減が期待できる。

A Proposal of Presentation Contents Sharing System

Akihiro KIMURA Kiyohiko KAJIHARA
NTT Software Laboratories
3-9-11 Midori-Cho Musashino-Shi Tokyo 180-0012 Japan
{kimura, kajihara}@slab.ntt.co.jp

Abstract

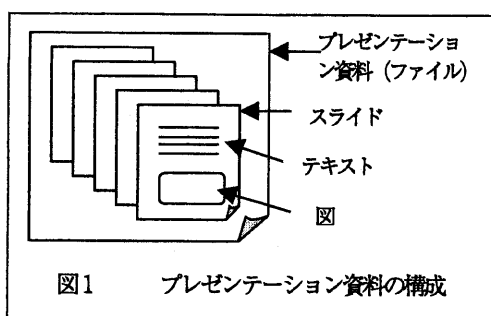
Because business presentation is usually made for each customer, it needs much time to make it. And more, there is a short time for preparing a presentation. Every company has their own presentation sotcks, but little people use them. If many people can use old presentation sotcks for making new presentaion, the company save the cost of making presentation.

In this paper, we introduce "presentation contents sharing system" which uses existing full-text serach engine. This system hepls users to get the graphics slides by presenting the full-text search result with reduced slide image. The users find reusable graphics, select the reduced slide image, and get the original presentation very easily.

1. 背景

プレゼンテーション資料（以下、資料と省略）の作成は人手に頼る作業であるため、コストの高いビジネスプロセスである。資料作成のコストを削減するためには、過去に作成された資料を有効活用することが効果的であると考えられ、我々の着眼点はここにある。

図1に示すように、資料は複数枚のスライドから構成され、スライドはさらにテキストと図によって構成されることが多い。テキストよりも図の方が、再利用のニーズが高く、再利用による資料作成コストの削減効果も高いと言える。この理由について次に説明する。



まず表1に、スライドの構成要素であるテキストと図について、再利用の観点から比較する。

表1 スライド構成要素の比較

	テキスト	図
作成コスト	低い	高い
再利用ニーズ	低い	高い
再利用可能性	低い	高い
再利用手段	全文検索システムの利用	自分/他人の資料/クリップアートの利用

テキストは作成にそれほど時間がかからないため、再利用ニーズはそれほど高くない。また、テキストは資料毎に書き換えなければならないことが多いため、再利用できる可能性もそれほど高くない。技術的には、既存の全文検索システムを利用することによって、テキストの再利用はある程度可能である。

一方図は作成に時間がかかるため、再利用ニーズは高い。また、図のパターンはある程

度決まっていて、似たような図が作成されることが多いため、再利用できる可能性も高い。しかしながら、現状では十分実用的な画像検索システムが提供されていないため、自分や他人が過去に作成した図、またクリップアートなど限られた範囲での再利用しかできていない。

そこで我々は、ニーズが高く、資料作成コストの削減にも有効であると考えられる図を中心とした資料再利用支援システムを検討したので本稿で提案する。

まず第2節で、イントラネットにおける資料再利用の現状について考察し、第3節で、我々が目標とするシステムについて述べる。次に第4節で、目標システムを実現するために考案した検索方式とその実現方式、実装について述べる。第5節で今後の課題について考察し、第6節で本稿をまとめる。

2. イン트라ネットにおける資料再利用の現状

ここ数年のイントラネットの普及により、組織内の情報共有環境は整備されつつある。イントラネットでは、WWWサーバと、情報検索システムあるいはデータベース管理システム等を統合することにより、多量の情報から必要な情報を検索することが可能となっている。

今回我々は、このような既存のイントラネットの利用を前提として、資料再利用支援システムについて検討した。2.1で、イントラネットで資料を再利用するために必要なソフトウェアについてまとめ、2.2で、イントラネットにおける資料再利用のニーズについて考察する。

2.1 ソフトウェア環境

イントラネットで資料を再利用するために必要な利用者側/サーバ側のソフトウェアについてまとめ、本稿で提案する資料再利用支援システムは、以下のソフトウェア環境を

前提として検討したものである。

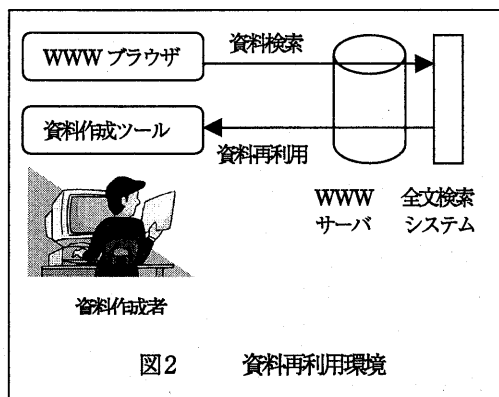
利用者側ソフトウェア：

- ・ WWW ブラウザ
- ・ 資料作成ツール

サーバ側ソフトウェア：

- ・ WWW サーバ
- ・ 全文検索システム

図 2 に示すように、利用者は資料作成ツール(例：Microsoft PowerPoint)を用いて資料を作成し、WWW ブラウザ(例：Microsoft Internet Explorer)を用いて再利用したい資料を検索・参照することができる(図 2)。



2.2 ユーザニーズ

社内組織のヒアリングにより調査した、イントラネットにおける資料再利用のユーザニーズをまとめる。

大ニーズ：

資料作成コストを削減したい

ニーズ 1：

再利用するための図を検索したい

ニーズ 2：

検索結果表示画面から、再利用するための図を速く簡単に発見したい

ニーズ 3：

発見した図を編集して再利用したい

利用者が資料を再利用する理由は、資料作成コストの削減であり、それが大ニーズとなる。大ニーズに応えるためには、ニーズ 1～

ニーズ 3 に応えることが必要である。

3. 目標システム

2.2 でまとめたユーザニーズに基づき、今回我々が目標としたシステムの要件をまとめる。

目標 1：

スライドに含まれる図を検索できること(ニーズ 1 に対応)

目標 2：

検索結果の表示画面に、検索した図のイメージを表示できること(ニーズ 2 に対応)

目標 3：

検索した図を利用者が編集して再利用できること(ニーズ 3 に対応)

4. 目標システムの実現

3 節で述べた目標システムを実現するために我々が考案した検索方式とその実現方式、実装について説明する。

4.1 考案した検索方式

目標システムを実現するためには、図を検索するための手段が必要となるが、今回我々は、既存の全文検索システムを利用した図検索方式を考案した。

図の検索手段としては、画像検索技術があるが、現在既に実用化されている画像検索システムでは、画像内容に関する検索用キーワードを手で付与しなければならないなどの制約があり、運用上のコストが高い。一方、キーワード付与を必要としない画像の内容検索技術・意味理解技術は、現在のところ研究段階であり、画像の特徴量として色や形状の抽出はある程度可能となっているが、内容や意味の理解は困難である。

そこで我々は、既にイントラネットに普及している全文検索システムを利用することにより、資料中のスライドに含まれるテキストを検索対象としてスライド中の図検索を実現する方式を考案した。この方式によって、キーワード付与を必要とせず、図の検索を実

現することが可能となる。また、スライド中には図の意味を説明するキーワードが記述されることが多いため、この方式は妥当であると考えられる。

4.2 実現方式

3 節で述べた目標システムを実現するためのアプローチの基本は、用途の異なる 2 種類の共有ファイル（編集用ファイル、参照用ファイル）を用意し、それぞれファイル形式と資料の所在を分離した上で、両者を関係付けることである（図 3 参照）。このアプローチによって、3 節で述べた目標システムを実現することが可能となる。

まず表 2 に、参照用ファイルと編集用ファイルの違いについてまとめる。

表 2 編集用ファイルと参照用ファイル

	編集用ファイル	参照用ファイル
ファイル形式	オリジナルファイル形式 (例:PPT)	HTML 形式
資料の所在	資料作成者のクライアント PC	資料共有サーバ

以下、実現方式のポイントについて説明する。

4.3 共有ファイルについて

以下、編集用ファイルと参照用ファイルについて説明する。

4.3.1 編集用ファイルについて

編集用ファイルは、資料作成者によって作成されたオリジナル資料である。ファイル単位は 1 資料、ファイル形式は資料作成に用いられた資料作成ツールに依存する。

利用者は、再利用したい図あるいはテキストを、編集用ファイルからコピー&ペーストして編集して利用することができる(目標 3 の実現)。

4.3.2 参照用ファイルについて

参照用ファイルは、編集用ファイルを HTML 変換した資料である。ファイル単位は 1 スライド、ファイル形式は HTML である。編集用ファイルから参照用ファイルへの HTML 変換は、編集用ファイル作成ツールの HTML 変換機能等により実行する。HTML 変換に求められる条件を次に示す。

- ・スライド全体を画像に変換するのではなく、スライド中のテキスト、さらに可能であれば、図中のテキストを抽出すること

参照用ファイル（スライド）は、スライド中の図検索を実現するために、スライド中のテキストを全文検索システムの検索対象として利用する。

参照用ファイルの単位はスライドであるので、スライド単位での資料検索が可能となる(目標 1 の部分的実現)。参照用ファイルの形式は HTML であるので、利用者は WWW ブラウザを用いることにより、検索したすべてのスライドを参照することができる。また、検索結果表示画面に検索したスライドの縮小イメージを埋め込むことによって、利用者は必要な資料を速く簡単に発見することができる(目標 2 の実現)。

4.4 共有ファイルの所在について

編集用ファイルと参照用ファイルが管理される所在について説明する。

4.4.1 編集用ファイルの所在について

編集用ファイルは資料作成者のクライアント PC の共有フォルダに置かれる。

4.4.2 参照用ファイルの所在について

参照用ファイルは資料共有サーバの共有フォルダに置かれる。

4.5 編集用ファイルと参照用ファイルの対応関係について

本システムでは、編集用ファイルと参照用ファイルの所在をルール化し、それぞれ以下に示すルールに従うこととした。このルールによって、図3に示すように、編集用ファイルと参照用ファイルの対応関係を定めることができる。

編集用ファイルの所在ルール：

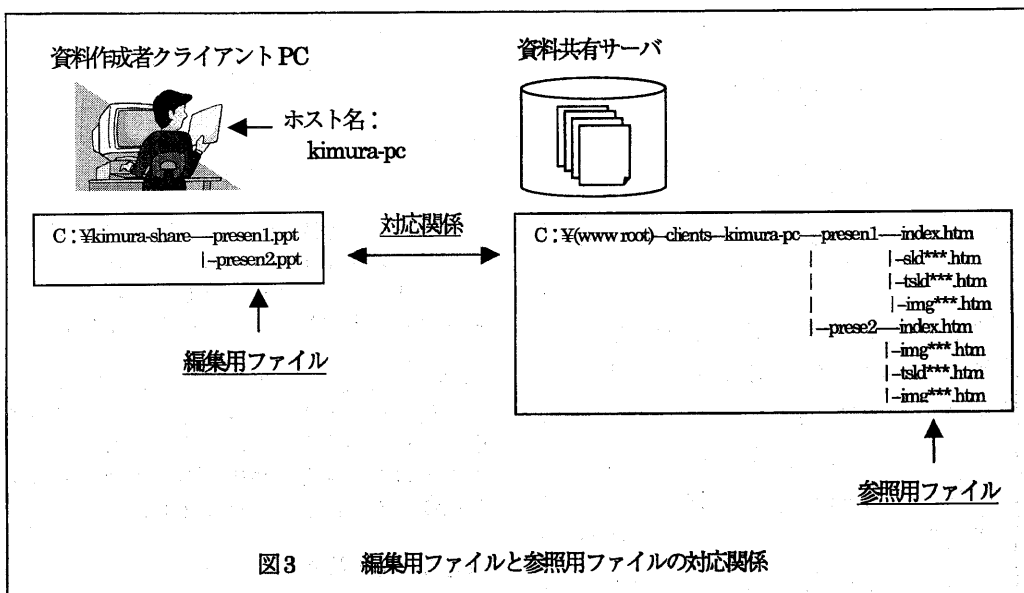
C:\¥共有フォルダ名¥編集用ファイル名

参照用ファイルの所在ルール：

C:\¥wwwroot¥clients¥クライアント PC ホスト名¥編集用ファイル名¥参照用ファイル

名

この所在ルールによって定まる編集用ファイルと参照用ファイルの対応関係を利用することによって、検索した参照用ファイル（スライド）と対応する編集用ファイルへのリンクを検索結果表示画面に埋め込むことが可能となり、利用者は、検索結果表示画面から、検索した参照用ファイル（スライド）と対応する編集用ファイルを容易に取得することができる（目標3の実現）。



4.6 実装

今回開発したシステムの独自機能は、検索結果を編集する次の2つの機能である。

機能1：

4.3.2で述べた、検索結果表示画面に検索した参照用ファイル（スライド）の縮小イメージを埋め込む機能

機能2：

4.5で述べた、編集用ファイルと参照用ファイルの対応関係に基づき、検索した参照用ファイル（スライド）と対応する編集用ファイルへのリンクを検索結果表示画面に埋め込む機能

機能1・機能2の開発にはVB (Visual Basic)スクリプトを用いた。開発したVBスクリプトは検索結果に埋め込まれ、検索要求元の利用者のWWWブラウザで実行される。

今回開発したシステムの利用者画面を図4、図5に示す。比較のため、従来の全文検索システムの利用者画面を図6に示す。

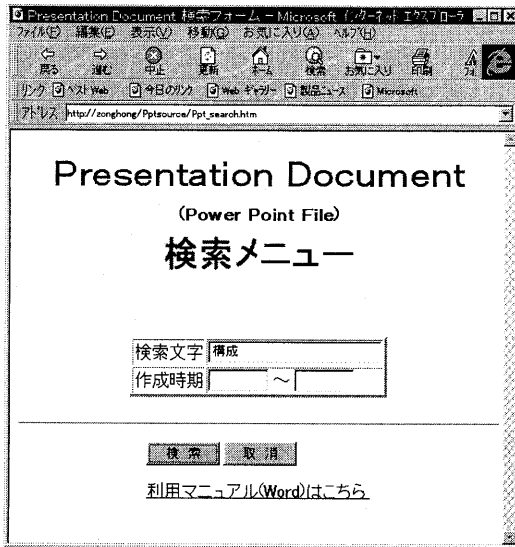


図4 検索条件入力画面

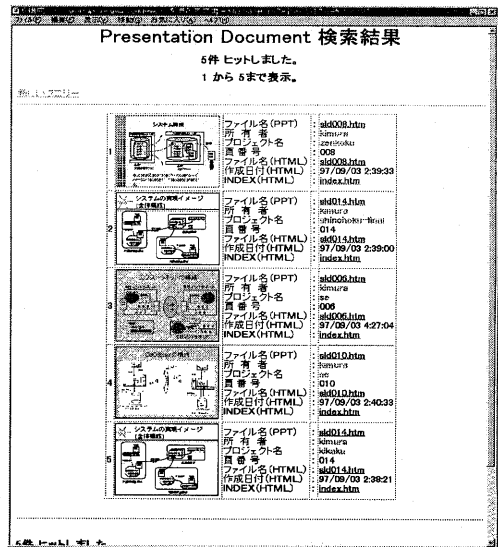


図5 検索結果画面

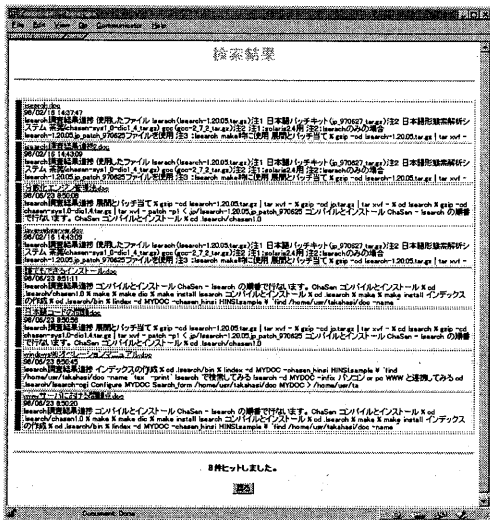


図6 従来の全文検索システムの検索結果画面

5. 今後の課題

以下、今後の課題について考察する。

課題1：システム評価

本システムは、現在一部の社内組織に導入

され試験的に利用されている。今後 WWW サーバのログが十分に蓄積されてから、ログ分析と利用者アンケートを行い、本システムを定量的・定性的に評価する予定である。

具体的には、利用者が再利用したいスライドを見つけるときに、検索したスライドの縮小イメージが検索結果表示画面に含まれることの有効性を、従来の全文検索システムと比較して評価する予定である。

課題2：検索精度評価

本システムの検索精度を、再現率と適合率 [1][2]の評価指標により評価する予定である。

本システムでは、スライド中のテキストを検索対象として、スライド単位での検索を行っているが、各スライドに含まれるテキスト量はそれほど多くないため、特に再現率による検索漏れの評価が重要であると考えている。

課題3：検索精度の向上

検索キーワードと資料・スライドの特徴を活用することによって、本システムの検索精度の向上が可能であると考えている。例えば、スライドの特徴を活用して、検索キーワードのスライド内における出現位地を評価するこ

とによって、検索結果のスライドをスコアリングすることが可能となる。今後、両者の特徴と関係を分析する予定である。

本稿中の社名・製品名はすべて商標・登録商標です。

課題4：エクストラネットへの適用

本システムは単一のイントラネットでの利用を想定しているが、今後、本システムをエクストラネット内でも利用できるように、資料共有サーバ間の連携機能を拡張する予定である。サーバ間連携機能には、エクストラネット用情報共有プラットフォームとして我々が提案している CoOrbiter[3][4]を適用する予定である。

6. まとめ

イントラネットの普及によって、社内の情報共有環境が整備されつつある。しかしながら、作成コストの高い資料に着目してみると、社内では過去に作成された多量の資料を有効活用できていないのが現状である。

そこで本稿では、イントラネットに普及している既存の全文検索システムを利用した、図を中心とした資料再利用支援システムを提案した。本システムでは、従来の全文検索システムよりも容易に図を再利用することが可能である。図の作成は、資料の作成においてコストの高いプロセスであるため、本システムの利用により資料作成コストの削減が期待できる。

7. 参考文献

- [1] 長尾編：自然言語処理、岩波書店、1996
- [2] P.インゲベルセン著、藤原監訳：情報検索研究、トッパン、1995
- [3] 木村、梶原、飯尾：分散WWWサーバを使った検索エンジンの隠蔽、情報処理学会ソフトウェア工学研究報告 115-8、pp.57-64、1997
- [4] 飯尾、木村、梶原、伊集院：Extranet用分散情報共有システム CoOrbiter の概要、情報処理学会第 56 回全国大会講演論文集 (3)、1998