

Title	コンテキストネットワーキングを実現するCatalogue system
Sub Title	
Author	金子, 晋丈(Kaneko, Kunitake)
Publisher	慶應義塾大学デジタルメディア・コンテンツ統合研究センター
Publication year	2018
Jtitle	慶應義塾大学DMC紀要 (DMC review Keio University). Vol.5, No.1 (2018. 3) ,p.16- 26
JaLC DOI	
Abstract	
Notes	特集 DMC研究センターシンポジウム第7回 デジタル知の文化的普及と深化に向けて : コンテキストネットワーキングの分散型ミュージアムへの展開 開催日時 : 2017年11月24日 (金) 14:00 ~ 17:30 開催場所 : 慶應義塾大学日吉キャンパス西別館1 講演録
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO32002001-00000005-0016

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

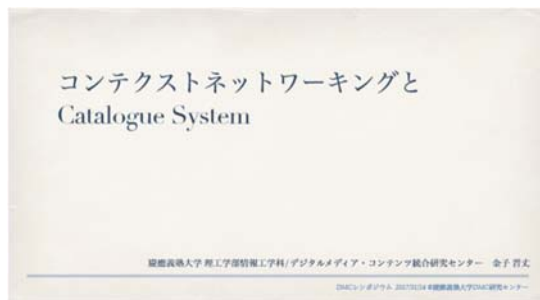
The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

コンテキストネットワークング を実現する Catalogue System

金子 晋丈

(慶應義塾大学理工学部専任講師・
DMC 研究センター研究員)

今日のタイトルは、コンテキストネットワークングとカタログシステムということでお話をしたいと思います。



まずコンテキストネットワークングについてです。コンテキストというものについて先ほど渡部先生の方からとても分かりやすい発表をいただいて、何も私からは説明を付け加えることが何もないと思いながら拝聴しておりました。ですが、まず、コンテキストネットワークングについて私なりの理解をまとめさせていただきまして、その後コンテンツをネットワーク化すること、そしてそのコンテンツをネットワーク化してできたコンテンツネットワークをどのように設計していけばいいのか、技術的な話をさせていただこうと思っております。私たちは実際にコンテンツネットワークとしてカタログシステムというものを作成しており、本日は、カタログシステムの構成とそれがもたらす世界、また、これからの技術的な挑戦についてお話しできればと思っております。

先ほど、きれいなゴッホの絵や自画像を見た後ですが、福沢諭吉先生の胸像を例に話します。まず、コンテンツは基本的には無数のコンテキストの中に存在しているというのが私たちの考えている立ち位置です。

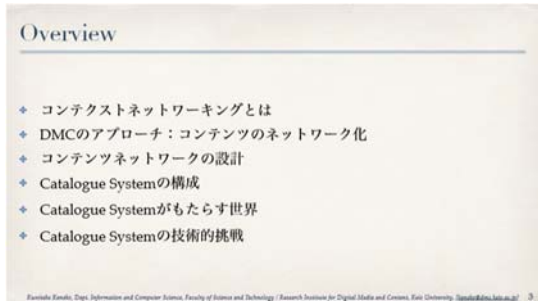


例えばここに福沢諭吉の胸像があります。これは三田キャンパスの中にある胸像を写真撮影したものです。これをコンテキストで見ると、三田キャンパスにあるので三田キャンパスの彫刻の中にも位置付けられますし、福沢諭吉ということであれば1万円札や、福沢諭吉がアメリカで撮った写真、咸臨丸の話、また、この胸像は誰が作ったか、誰が撮った写真なのか、といったさまざまなコンテキストをこの1枚のデジタルファイルに与えることができます。

このつながりをどのように使っていか、ということが、コンテキストネットワークングになるわけですが、例えば福沢諭吉の胸像をスタートポイントにして、このいろいろな矢印で表されるコンテキストを見渡し、あらゆるコンテキストを知ること、このコンテンツに対する理解が進みます。

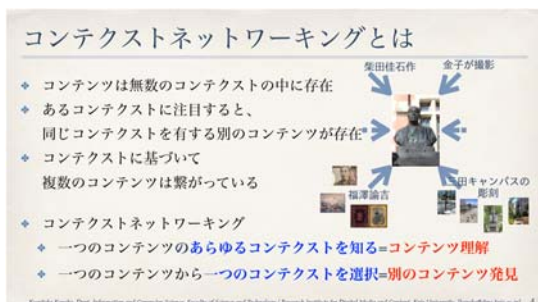
これはまさに先ほど渡部先生がおっしゃっていた内容だと思います。また、1つのコンテキストをたどり、選択し、先に進んでいくことによって、別のコンテンツの発見につながる。これを繰り返し実現していくのがコンテキストネットワークングであると

考えております。



DMCはこのコンテキストネットワーキングの実践に当たり、コンテンツをコンテキストでネットワーク化することがそれを解決する技術的な手法だと考えております。すなわち、ここで先ほど出てきたコンテンツをコンテキストでうまくつなげてあげることによって、コンテンツ間をつなげるネットワークをつくってあげようということになります。

従って、コンテンツとコンテンツをつなげるのはコンテキストであり、もっと広い視野で見ると、複数のコンテンツ間にさまざまなコンテキストでネットワークが構成されていくことになります。皆さんは今、コンテキストとコンテンツというものを1個のペアとしてとらえているかもしれませんが、それを重層的に組み合わせることで、コンテンツネットワークを目指そうとしているのです。



ネットワークというと皆様の頭の中に浮かぶのは、例えば電話網、もしくはインターネット、コンピューターをつなげるネットワークが頭に浮かんでくると思います。我々がやろうとしているのは、コンテンツとコンテンツをつなげるネットワークを新たに作っていくことです。それをコンテンツネットワークと呼んでいるわけです。



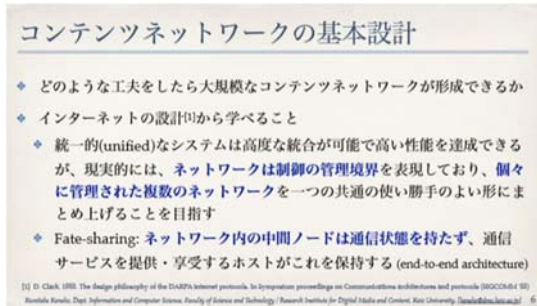
では、どのようなコンテンツネットワークをつくれれば、コンテキストネットワーキングの世界を最大限に楽しむことができるのでしょうか。それは、コンテンツの数を増やす、コンテキストの数を増やすということにほかなりません。

なぜならコンテキストが少ない、コンテンツが少ないと「ああ、そうだよね」と、完結しやすいからです。そこに新しい発見は何も得られませんし、新しい理解も得られません。

従って、いろいろな人がいろいろな価値観を提示する。それを同じ場所に持ち込むことによって、コンテキストネットワーキングは非常に豊かになっていくと考えることができます。

そうしますと、技術的には、いかに大規模なコンテンツネットワークを構成していくかが肝になってくるというのは明白であろうと思います。ここがコンテンツネットワ

ークを技術的に構成していくための、基本設計のスタートポイントにあたるわけです。



コンテンツネットワークという言葉自体が皆様にとっては初めてかもしれません。多くの人にとって初めてだろうと思います。さらにそれを大規模にするというのは、どのようなことなだろうと、思われるでしょう。ここで過去の事例から、大規模なネットワークをつくるために何をしなければいけないのかといったことについて話してみたいと思います。

私をご紹介する事例は、インターネットの設計です。皆さんはおそらくインターネットを当たり前のように使い、その設計がどういうものか、どういう考え方にもとづいてインターネットというものがこの世に存在することになったのか、そういったことについてはあまりご存じないかもしれません。



そこで今日は、『The design philosophy of the DARPA Internet protocols』という論文

を持参しました。

IT系の方々は聞いたことがあることかもしれません。これはデービッド・クラークという MIT の計算機学者が、1988 年に発表した論文です。内容は、約 15 年前に TCP/IP が開発されて今に至るが、実はその技術的な背景が誤解されていたため、ここでまとめておこうと思う、といった主旨の論文です。

ページ数はそれほど多くないものの、非常に深みのある、今また読み返してみても面白い論文です。私は、大学生のときに、この論文と出会いました。

大学生のときに読んで、分からないなりに納得しましたし、今日の講演に向けて久しぶりに読み返してみたところ、味わい深く、さまざまな言葉が心に突き刺さり楽しい思いをしました。この論文は、インターネットのデザインの中で、大規模にネットワークをつなげていくためには、何をしなければいけないのかといったことを示唆しています。

ゴールとして、まずファンダメンタルなゴールと、セカンドレベルのゴールの 2 つに分けています。ファンダメンタルなゴールの中にはこのような表現があります。『ユニファイドなシステムは高度な統合が可能で高い性能を達成できるが、(インターネットは)現実的には、ネットワークは制御の管理境界を表現しており、個々に管理された複数のネットワークを 1 つの共通の使い勝手のよい形にまとめ上げることを目指す。これによって大規模なネットワーク空間をつくる』ということをおっしゃっています。

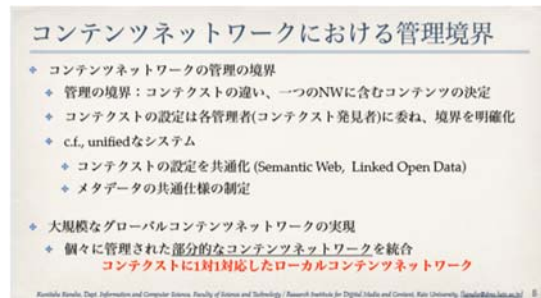
すなわち大規模なネットワークを津々浦々まで 1 つの方式で統一し、これをみん

なでやるんだというような方法を取るのではないのです。個々個別の組織が持つ状況やネットワークに求められる機能、それはそのまま維持した状況で、差異がある状態でそれらをつなげていく。各ネットワークの違いを了承した状況でつなげていき、つなげたときに1つの共通の使い勝手のよい形(ア・コモンユーティリティー)を実現するというのがファンダメンタルなゴールとして挙げられております。

もう1つ、セカンドレベルのゴールから抜き出してきたのは、フェートシェアリングという考え方です。これはエンド・ツー・エンドのアーキテクチャーとインターネットが言われるもとなった考え方です。インターネットの基本的な考え方では、クライアント端末とサーバーが通信に関する状態を、どこまでデータを送ったか/届いたかといった状態として管理しサービスを提供しています。インターネットの真ん中にはルーターやいろいろなスイッチがあるのですが、そこは一切のサービスに関する通信状態を持たない。端末が壊れたらサービス終了。でもこのクライアントとサーバーが生きているのであればサービスは継続すべきである。それ以外のほかのところは状態を持たないようにしよう、という考えです。こういった内容が、実はインターネットの設計に書かれています。私はこれがインターネットをここまで拡大させることになった、とても重要な設計指針であると考えております。

さて、先ほどやや難しい言葉が出てきました。制御の管理境界ということですが、制御の管理境界を簡単に言うと、LAN です。自分のLAN、家のLAN、ホームネットワー

ク。家のブロードバンドルーターがあって、プリンターとかコンピューターをつなぐと思いますが、それは自分自身で管理できます、ほかの人は管理しません、という考え方です。



では、この考え方をコンテンツネットワークに適用した場合はどのようなのでしょうか。我々はコンテキストネットワークキングをコンテンツネットワークで実現しようとしていますので、コンテキストの違いが管理の境界になるのではないのでしょうか。それはほかの人が踏み込めるものではなく、その人自身が管理するととらえることができます。

別の表現をすると、LAN と同等のローカルコンテンツネットワークに存在するコンテンツは自分で決定することができるということです。それが、管理の境界が明確ということかと思います。

すなわち、コンテキストの設定とは、コンテンツネットワークにおいて、各管理者、いわゆるコンテキストを発見した人に委ねて境界を明確化することが、先ほどのインターネットの設計指針から得られる、大規模ネットワークを構築するための重要な策ではないかと考えるわけです。

ここで参考までに、デービッド・クラークの論文において、大規模なネットワークを

つくるまでのオルタナティブアプローチとして出てきているユニファイドなシステムとコンテンツネットワークの関係についてお話しします。ユニファイドなコンテンツネットワークシステムでは、あるコンテキストをこのように定義し、その定義にマッチするものを、このように整理しなさい、というような一種のルール決めを行い、それに完全に従うことになるかと思えます。

具体的な例を挙げると、メタデータの共通基準をつくり、それにもとづいてデータを管理しましょうという考え方は、ユニファイドなシステムを大規模につくろうとしているにほかならないと思うわけです。

また、セマンティックウェブ、もしくはセマンティックウェブから派生して出てきたリンクドオープンデータは、自由な記述が特徴ですが、一種の検索エンジンのような中央集中型のものがウェブ空間上クロールしなければその情報にたどり着くことはできないのです。

すなわち大規模で、グローバルなコンテンツネットワークを形成しようとするときに、どのようなアプローチをとるかと言いますと、個々に管理された部分的なコンテンツネットワークをひとつひとつ構築することになります。部分的なというのは、コンテキストに1対1対応したローカルなコンテンツネットワークを積み重ねていき、大規模なグローバルなコンテンツネットワークをつくるということです。これが、大規模ネットワークをつくるための鍵になると考えられます。

さて、最近ウェブサーフィンやネットサーフィンといった言葉をまったく聞かなくなりましたが、ウェブはコンテンツにアク

セスできるメカニズムです。では、ウェブはコンテンツネットワークとして機能するかを考えてみたいと思います。

WWWはコンテンツネットワークとして機能するか？

- 自由なコンテンツネットワークの構築ができない
 - 他人のコンテンツを自分のコンテキストに取り入れられない
 - リンクがコンテンツ(htmlファイル)内に記述されている
- ネットワーク境界が明確でない
 - 検索エンジンが勝手に同一文字列によってネットワーク境界を構成
 - キーワードは言語の曖昧性を有している
- ネットワーク内の中間ノードがサービス提供に必要な状態を持っている
 - プロローガー(Google, Facebook, etc.)がないと、サービスを受けない

Key A: url 1
Key B: url 2
Key C: url 3

Keiichi Arita, Dept. Information and Computer Science, Faculty of Science and Technology, Research Institute for Digital Media and Content, Keio University, Yokohama, Japan, arita@ipc.keio.ac.jp

まず、HTML ファイルについてです。ウェブというものは HTML ファイルがあり、そこにリンク情報があります。始点となるコンテンツにアクセス権がないと、つまり、自分が編集できる状態でないとリンクを記述できないという問題があります。すなわち HTML ファイルに書くとリンクを張れるが、HTML ではないファイルから、他の誰かのウェブページへのリンクを加えたいと思っても加えられない。すると自由なコンテンツネットワークが実現できません。

次は、ネットワーク境界が明確ではないことが挙げられます。具体的に言うと、皆さん検索エンジンを使っていると思いますが、検索エンジンというのは同一な文字列を用いて、その文字列でグループ化をしているにすぎません。その文字列がネットワーク境界を構成しているわけで、私たち人間が選別して構成しているわけではありません。さらに、文字を使っているというところが大きな課題となっております。例えば、キーワードは言語でできているため、文字を使った曖昧性が生じるわけです。入力した人、もしくはアクセスする人も、その言語の使い方に違いがあるとうまくネットワーク境

界を定めることができないといえます。

最後に、ネットワーク内の中間ノードがサービス提供に必要な状態を持っているということがインターネット、WWW のもう 1 つの大きな課題になっていると思います。

実際に例を書いてみました。利用者＝ウェブブラウザです。下にウェブサーバーがあり、「Google」、「Facebook」、私はこれらを勝手にブローカーと呼びました。この例から、利用者はまずブローカーに行き検索をして得られた結果でアクセスをすることがインターネットアクセスの現在の構図になっています。ブローカーは何を持っているのかというと、キーワード A を含む URL は何々ですよという対応付けのリストを持っています。

そのリストをなくしてサーバーにはアクセスできないということが、ウェブサービスの課題です。すなわち、ブローカーがこれを変えることによって違うところにナビゲートされてしまう。本当は欲しいデータはここにあるのに、テーブルが異なっていると他の情報源に行かざるを得ないということが WWW では発生します。

これがデービッド・クラークの書いた論文の 2 個目のネットワークの中のノードが状態を持っているというところに反している例と考えることができます。

ブローカーの例 (Google)

- ◆ 検索：キーワードとコンテンツをつなげる
 - ◆ 言語化の壁
 - ◆ コンテキストを適切に言語化できない
 - ◆ 作る側、探す側のキーワード選択の差異
 - ◆ 網羅的閲覧性の欠如
 - ◆ コンテンツに潜在するコンテキストに気づけない
 - ◆ キーワード推薦で補完しているつもり
- ◆ 検索技術
 - ◆ テレビや雑誌、友人等からキーワード(言語化されたコンテキスト)を手することを前提

Kanako Kuraki, Dept. Informatics and Computer Science, Faculty of Science and Technology / Research Institute for Digital Media and Content, Keio University, Kanako@itc.keio.ac.jp 10

ブローカーについて、具体的に「Google」の例と「Facebook」の例を出してみました。

「Google」は、何と何をつなげているコンテンツネットワークかといいますと、キーワード(文字列)がコンテンツをつなげています。先ほどお話したように、キーワードを使っている以上言語の壁というのがそこには立ちはだかります。コンテキストというものは、適切に言語化できない限りそれをキーワードで探し出すことはできません。

そして、作る側、探す側でキーワード選択の差異が発生すると、文字列で識別しているので、その差異を埋めることができないというのが言語識別の壁です。

さらに、網羅的閲覧性の欠如がキーワード検索にはあります。キーワード検索はキーワードを入れるとコンテンツが出てくる検索です。コンテンツは複数のコンテキストを持っていますが、それらのうち自分の知らないコンテキストは、それに見合ったキーワードを入れない限り、永遠に出てこないというのが網羅的な閲覧性の欠如です。

「Google」はある日から、キーワード推薦を実装してきました。キーワードを 1 つ入れる、2 つ入れると、候補となる 3 つ目のキーワードが列挙されてくるというものです。

これは、一種、網羅的閲覧性の欠如を回避するための技術と言えるかもしれませんが、一方でそれはよく使われるキーワードを順番に出しているに過ぎないため、結果的には存在しているコンテキストを埋もれさせてしまうシステムになっているのかもしれない。

一般に「Google」のような検索技術は、最初にどこからか、コンテキストが与えら

れなければならず、これに対して、何も手助けをしてくれません。テレビや雑誌、友人等からキーワードが与えられる前提のサービスです。

すなわち言語化されたコンテキストを誰かから入手しないと、「Google」検索サービスというのは使えないことになります。

ブローカーの例 (Facebook)

- ◆ 人と人/コンテンツをつなげる
- ◆ コンテンツネットワークの構築
 - ◆ タイムライン
 - ◆ タイムラインに情報を並べることでキーワード入手を不要に
- ◆ コンテンツに関するコンテキストの網羅的閲覧性の欠如
- ◆ コンテンツに潜在するコンテキストに気づけない
- ◆ 出てくるのは「いいね」した人の一覧
- ◆ ネットワークが与える付加価値
- ◆ どの友人が「いいね」したか

Kanako Kinoshita, Dept. Information and Computer Science, Faculty of Science and Technology / Research Institute for Digital Media and Content, Keio University, Kanako@ipc.fsc.keio.ac.jp 11

一方、「Facebook」は、一種のコンテンツネットワークであると言えます。

これは人と人、もしくは人とコンテンツをつなぐコンテンツネットワークでしょう。非常に面白い仕組みを持っており、コンテンツネットワークはタイムライン、つまり時系列で並ぶという特徴があります。タイムラインに情報を並べることで選択が可能になり、従って、言語にしなくていいといったコンテキストの言語化を省いていると言ってもいいかもしれません。

一方で、コンテンツに関するコンテキストの網羅的な閲覧性ができているかというと、推薦されたコンテンツは上がってきますが、そのコンテンツにどんなコンテキストがあるのかを提示してはくれません。

このコンテンツに誰々さんがいいね！しましたという情報が出てくるだけで、基本が人です。ソーシャルネットワークと言われるように人を中心としたネットワークなので、いいね！した人の一覧が出てきます。

ネットワークを考えるときにその付加価値についても考えますが、これは先ほどのデービッド・クラークから得られる示唆とはやや離れますが、どの友人がいいね！しているから見よう、といった行動が「Facebook」の大きな売りになっております。その付加価値はここにとどまっているということになります。

Catalogue System

- ◆ コンテンツネットワークを実現するシステム
- ◆ 意味を共有するデジタルデータの集合としてCatalogue(グラフ)を作成
 - 1 コンテキスト = 1 Catalogue = 1 ローカルコンテンツネットワーク
 - ◆ コンテキストの設定は自由(意味の統一・標準化はしない)
 - ◆ 誰でもいくつでも作成可能
 - ◆ 他人のデジタルデータを含めてよい
 - ◆ 集合だけでなく、順序性を表す矢印の記述も可能
- ◆ Catalogueは共有されデジタルデータがCatalogueのハブとなる
 - ◆ グラフのエッジは双方向にたどれる
- ◆ デジタルデータ(コンテンツ)そのものは別管理
- ◆ 前提：すべてのコンテンツ・カタログが重複しないIDを有する

Kanako Kinoshita, Dept. Information and Computer Science, Faculty of Science and Technology / Research Institute for Digital Media and Content, Keio University, Kanako@ipc.fsc.keio.ac.jp 12

さて、我々が考えているコンテンツネットワークを実際に具現化したものが、このカタログシステムというものです。カタログシステムはコンテンツネットワークを実現しているもので、意味を共有するデジタルデータの集合としてカタログを作成します。

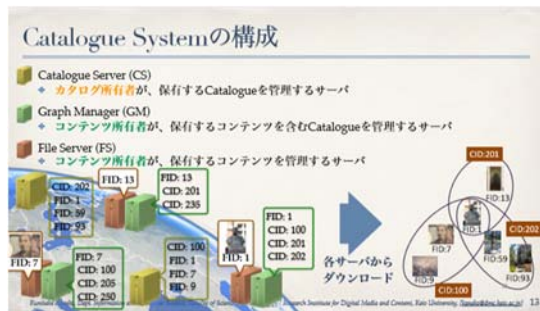
すなわち、カタログシステムでは、1 コンテキストが 1 カタログで、1 つのコンテキストを共有するコンテンツ群が 1 つのカタログの中に入ってきます。それで 1 つのローカルコンテンツネットワークを形成するという考え方です。

誰でもいくつでもローカルコンテンツネットワークを作成可能で、他人のデジタルデータをこのカタログの中に入れておき、そのカタログの中で順序を示すような記号を加えても構いませんよ、というのが、カタログの考え方です。

そして、このカタログをグローバルに共

有する。共有してどのように使うのかとい
いますと、このカタログの中に入っている
デジタルデータが、複数のカタログのハブ
になります。1つのデジタルデータが複数
のカタログのハブになることで、コンテ
ンツネットワークを形成することがカタログ
システムの特徴になっています。

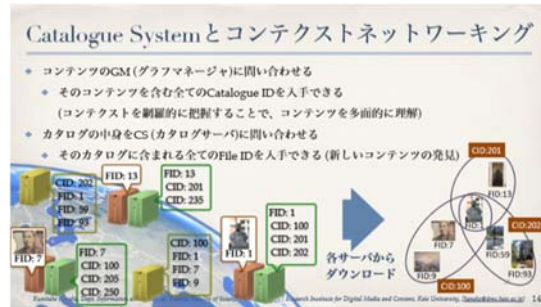
カタログシステムでは、このようなつな
がりだけを管理するので、コンテンツその
ものは別管理したり、デジタルデータをカ
タログのハブにするため、すべてのコンテ
ンツカタログが重複しない ID を有するこ
とが前提のシステムとなっています。



さらに具体的にシステムの中身を見てみ
ますと、カタログシステムは大きく3つの
サーバーで構成されます。カタログサー
バー、グラフマネージャー、ファイルサー
バーです。皆様にとって馴染み深いものはおそ
らくファイルサーバーです。

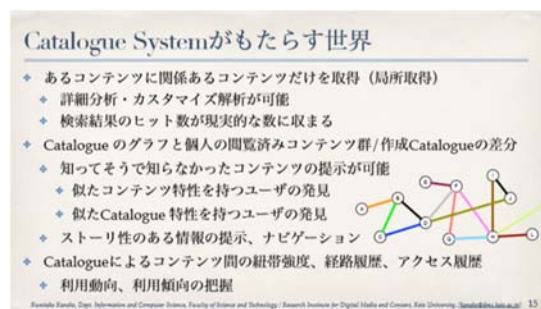
先ほどカタログを発見した人、つまり新
しいコンテンツを発見した人は、自分の
カタログサーバーに保有するカタログを保
存していきます。

グラフマネージャーは、コンテンツ所有
者が持つサーバーで、先ほどのハブを実現
します。デジタルデータがハブで、そのデ
ータにどのようにカタログがスポークとし
て存在するかを示すサーバーです。



この仕組みを図示しますと上記のようにな
ります。ユーザー側ですべて取得し、組み上
げてみるとこのようなコンテンツネットワ
ークが出来上がります。

さて、この仕組みを使ってコンテキスト
ネットワークがどのように実現される
のかと言いますと、このサーバーから問い
合わせてきて出来上がったここを見たとき
に、ファイル ID 1 番のこのコンテンツに対
して、カタログが3つある、つまり3つの
コンテキストが登録されているというこ
とが分かります。それを実現するのがグラ
フマネージャーになります。



1つのカタログに注目するということは、
つまり1つのコンテキストに注目するこ
とになります。そして、別のファイルを発見
するのが、カタログサーバーの果たす役割に
なります。

このようなカタログシステムがどのよう
な世界をもたらすかと言いますと、可能性
としては上記のようなことが挙げられます。

ここに書いてある内容は、とても奥深いものが並んでおりまして、簡単に説明できるようなものではありませんが、代表的なものとしては、1つのあるコンテンツに関係のあるコンテンツを高速に取得する、局所取得が可能になります。

「Google」が行っているような、ウェブ空間を全部クローリングしてデータをかき集めてくる方式に比べてみますと、自分に関係するデータだけを持ってきて、その部分だけで詳細の分析、カスタマイズされた解析が可能になります。

一方、現在の「Google」は例えばキーワードを入れた結果、数百万件の検索結果が出てきます。皆さん、あの膨大な量の検索結果は嬉しいものでしょうか。おそらくは、全く嬉しくないですね。きっと最初に出てきた5個程度しか見ないと思います。つまり数百万件の結果を出す価値はありません。

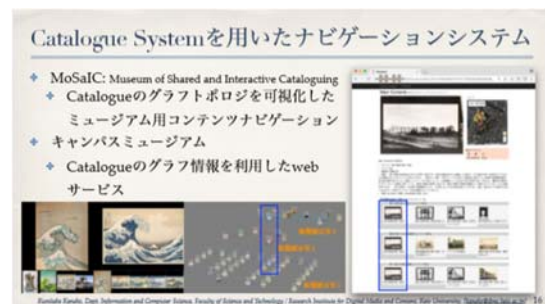
検索結果のヒット数が、現実的な数に常に収まっている必要があるとするならば、ある程度の絞り込みができるシステムが、良いシステムなのではないかと思います。

次に、パーソナライズにつながっていくという観点です。カタログからグラフ構造を持ってきて、その中の何個を自分が知っているか、自分がすでにアクセスしたかというものを比較することによって、自分が見ていないコンテンツを浮き彫りにする。その関係あるコンテンツで、見ていないコンテンツを浮き彫りにすることができるのがこのパーソナライズ、グラフを使ったパーソナライズのサービスです。

ただ単に閲覧したコンテンツを見るというだけではなくて、閲覧したコンテンツを軸にグローバルのコンテンツネットワーク

との差分を取ったり、自分がコンテンツ間に引いているグラフを確認することで、自分の物事の理解ができるようになります。グラフは、いわゆる自分が発見したコンテキストを表しているもので、サーバー側にある誰かのカタログとその構造が似ていれば、同じものの理解をしている可能性が高いということになります。すると、似ているユーザーを発見することもできますから、また新しいサービスが可能になるのではないかと思います。

そして、グラフ構造、コンテンツのカタログで構成されるグラフは、図のようになりますが、あるコンテンツがどの程度強い関係を持っているのか、ということも計算できるようにになっていきます。



DMC ではこのカタログシステムを用いて、ナビゲーションシステムを2つ作っております。1つはモザイクのビジュアルライズです。もう1つは、キャンパスミュージアムというウェブサービスです。こちらはコンテンツネットワークのグラフの形、トポロジーと私たちは言うておりますが、トポロジーを表現して可視化し、それをクリックすることで新しいものを発見していく。トポロジーを俯瞰的に見ながらサービスをするアプリケーションです。

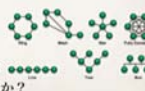
キャンパスミュージアムは、バックエン

ドにこの構造を持っているのですが、可視化はせずに、例えばこのコンテンツを軸に複数のカタログを提示するようなウェブサービスに仕立て上げております。

最後になりますが、カタログシステムの技術的挑戦をいくつかご紹介したいと思います。

Catalogue Systemの技術的挑戦

- ◆ グラフ取得のオーバーヘッドの削減 (取得時間の短縮)
 - ◆ 世界中に分散したサーバへのアクセス
 - ◆ グラフキャッシュ、グラフレプリカの導入
- ◆ 超大規模なコンテンツネットワークの処理負荷の軽減
 - ◆ 有名コンテンツは数億個のCatalogueに含まれると推測
 - ◆ 分散グラフシステム技術の活用
- ◆ Catalogueで用いるグラフ表現の模索、定型化?
 - ◆ 多様な情報をグラフ化できればコンテキストが豊かに
 - ◆ 時間情報、位置情報などをグラフで扱うにはどうするか?



Keio University, Dept. Information and Computer Science, Faculty of Science and Technology / Research Institute for Digital Media and Content, Keio University, Sanda@ipc.keio.ac.jp 17

先ほどよりお伝えしておりますとおり、コンテンツネットワークといっても、1つのネットワークがあるのではなく、コンテンツとそれをつなぐリンクのようなイメージです。グラフの話だとエッジと呼びますが、エッジは、世界中のサーバーにばらばらと保存されています。いろいろなサーバーから情報をかき集めてこないコンテンツネットワークは作ることができません。

これを作るためには、グラフの情報を世界中からかき集めてこなければならないのですが、そのオーバーヘッドはかなり大きなものが必要になります。

これを解決するために、グラフキャッシュやグラフレプリカというものを設置し、よりユーザーの近くで、もしくは少ないクエリーでその情報を取得することを研究としています。

また、超大規模なコンテンツネットワークの処理負荷が課題です。例えば、ループルのモナリザといった著名なコンテンツは、

皆さんが関連付けをしたがります。そのような作品になりますと、規模として数兆のカタログをどのように取得するのか、それをどのように処理するのか、という課題が生じます。

逆に、コンテキストを把握するといった際には、ループルのモナリザなどは恐らく数億のコンテキストとつながっているため、実は全てのコンテキストを把握することはもうほぼ不可能に近いということになります。

次に、グラフ表現をどうするのかという課題です。今は、時間情報、位置情報などもグラフで扱い、それをコンテキストとして表すことができればもっと面白いカタログシステムになるのではないかと考えております。このあたりはまだまだ挑戦しなくてはいけない分野と思っております。

まとめ

- ◆ コンテキストネットワーク
 - ◆ 一つのコンテンツのあらゆるコンテキストを知る=コンテンツ理解
 - ◆ 一つのコンテンツから一つのコンテキストを選択=別のコンテンツ発見
- ◆ 世界中のコンテンツをあらゆるコンテキストで繋ぐコンテンツネットワークが豊かなコンテキストネットワークを実現
- ◆ Catalogue System
 - ◆ 1 コンテキスト = 1 Catalogue = 1 ローカルコンテンツネットワーク
 - ◆ Catalogue Systemによるコンテキストネットワークの実現

Keio University, Dept. Information and Computer Science, Faculty of Science and Technology / Research Institute for Digital Media and Content, Keio University, Sanda@ipc.keio.ac.jp 18

最後に、コンテンツネットワークというのは、1つのコンテンツが持つあらゆるコンテキストを知ること。コンテキストを知ることによってコンテンツを理解しようという試みです。そして、あるコンテンツから、1つのコンテキストを選択することで、別のコンテンツを発見しようという試みがコンテキストネットワークだと考えています。これらを実現するためのシステムとしてカタログシステムを作っています。

本日、インターネットの設計指針からコンテンツネットワークを作り、維持していくためには、何をしなくてはいけないのか。それらを踏まえた上で、現在はカタログシステムを設計しているということをご説明させていただきました。ありがとうございます。

金子 晋丈 (かねこ くにたけ)

慶應義塾大学理工学部専任講師・DMC 研究センター研究員。専門はアプリケーション指南ネットワーク。特に、デジタルデータの利活用を促すデジタルデータのネットワーク化について研究を行っている。2001年東京大学卒業。2006年同大学院情報理工学系研究科博士課程終了、博士(情報理工学)。同大学院新領域創成科学研究科での特任助手を経て、2006年9月より慶應義塾大学デジタルメディア・コンテンツ総合研究機構、特別研究助手。2007年、同機構特別研究講師。2012年4月より現職、デジタルメディア・コンテンツ総合研究センター研究員を兼任。