



新しい検索手法の探求

<情報検索の最先端>

2015年度情報検索委員会
第1小委員会 第3WG

6月21日(火) 関東部会
6月28日(火) 関西部会



第1小委員会 第3WGメンバー

☆六坂	明彦	ルネサス エレクトロニクス株式会社
藤田	尚吾	凸版印刷株式会社
鈴木	憲	日油株式会社
能登	信博	株式会社リコー
澤井	博	住友ベークライト株式会社
鈴木	浩三郎	株式会社ミットヨ
小田	將文	TOA株式会社
生井	丈夫	テルモ株式会社

☆ リーダー



目次

- 1) 背景、目的、課題、方針
- 2) 検証対象ツールの紹介
- 3) 文章系情報検索ツール
- 4) 画像系情報検索ツール
- 5) 総括
- 6) 提言



1) 背景、目的、課題、方針

◆ 背景

◇近年、情報検索技術の進歩は凄まじく、セマンティック検索や人工知能の活用、更には画像データそのものを検索項目として用いるものなどが、一部の官庁や企業では既に採用され始めている。このように、情報検索の環境は、大きな変革期を迎えようとしている。

セマンティック検索とは：
「サーチャーの意図やキーワード(文章)の意味を理解した上で行われる情報検索」



1) 背景、目的、課題、方針

◆ 日本特許庁の人工知能に関する動向

- ・平成28年度「人工知能技術を活用した特許行政事務の高度化・効率化実証的研究事業」の企画提案の公募（5/10締切済）
http://www.jpo.go.jp/koubo/koubo/jinkou_chinou.htm
- ・外国特許文献へのFターム等付与に関する機械学習活用可能性調査事業についての一般競争入札公告（6/13締切済）
http://www.jpo.go.jp/koubo/choutatu/choutatu2/1605011_chousa.htm
- ・特許行政サービスの質向上：人工知能技術を活用した特許行政サービスの高度化・効率化の可能性の検討（5/9web掲載）
http://www.jpo.go.jp/shiryou/toushin/shingikai/pdf/newtokkyo_shiryou014/04.pdf



1) 背景、目的、課題、方針

◆ 目的

- ◇ 情報検索の"今"を把握し、そこで得られた内容を JIPA 会員へ広く展開。
- ◇ 最新の情報検索ツールを可能な限り試行し、より理想的な機能や手法をツールベンダへ提言。



1) 背景、目的、課題、方針

◆ 課題

◇ 検索者の知見と技量に基づいた【キーワード】や【分類コード】を用いて検索式を立てることを前提とする現行の情報検索に対し、

~~((touch+tap) adj3 (feel+robot)) or (deep adj (learning or study)) and (memory or store) + G06F13/00,F+B60Q7/06 + ...~~

新たな手法により、

作業効率の向上

より関連性の高い情報の取得

を目指している次世代検索システム(手法)の探求。





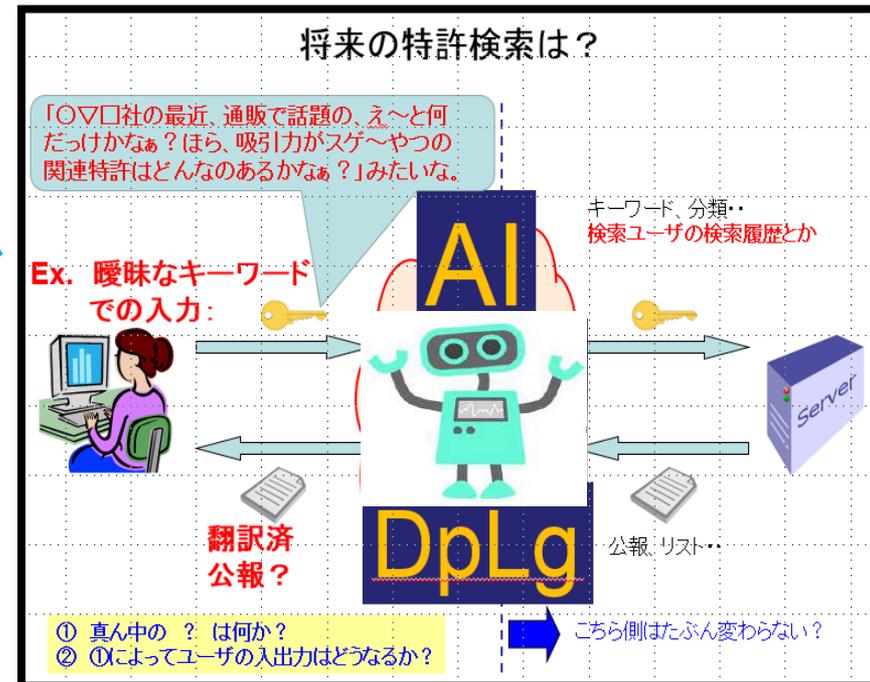
1) 背景、目的、課題、方針

◆ 方針

◇ 検索ツール(手法)の選定は、**国内外を問わない。**

◇ **【文章系情報検索】と【画像系情報検索】の2方向**
で試行、検証を進める。

◇ **検索ツールのアルゴリズム**
には、深く立ち入らない。
(あくまでもユーザ目線。)

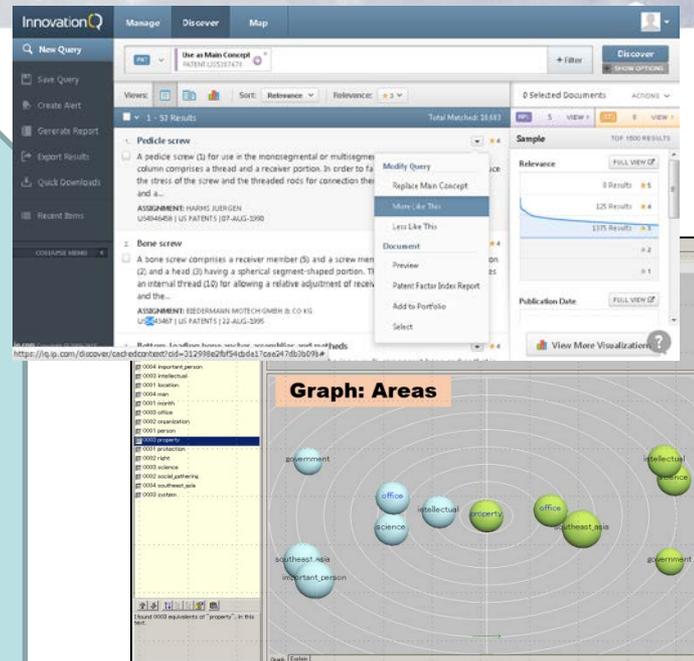




2) 検証対象ツールの紹介

◆ 文章系情報検索ツール

- 1) Smart Search (英語、日本語)
- 2) InnovationQ (英語)
- 3) Lit i View PATENT EXPLORER (日本語)
- 4) Invention Machine Goldfire(英語、日本語)
- 5) INNOGRAPHY(英語)
- 6) Google Patent Prior Art Finder(英語)
- 7) Tropes (英語)
- 8) R&D Navi (英語、日本語)



◆ 画像系情報検索ツール

- 9) Global Brand Database
- 10) 画像意匠公報検索支援ツール
- 11) Google画像検索
- 12) JST文献画像検索システム
- 13) Yahoo!ラボ VisualSeeker 他



文献画像検索システム
科学技術情報機構 (JST) & 日本機械工学会 (JAE) 共同開発

「図例1」：手元画像との類似検索
手元の画像ファイル（手書きの文字）をアップロードして、類似画像を検索することができます。
類似画像を検索することで、類似画像の検索結果を簡単に確認することができます。
この検索結果を元に、類似画像の検索結果を簡単に確認することができます。
この検索結果を元に、類似画像の検索結果を簡単に確認することができます。





3) 文章系情報検索ツール

3-1) Smart Search

3-2) InnovationQ

3-3) Lit i View PATENT EXPLORER



3-1) Smart Search

- ◆ 開発元 : Thomson Reuters
- ◆ Thomson Innovationの機能の一つ。
- ◆ 入力したキーワードから、自動的に検索語を広げ、さらにIPCを含めて検索を行い、検索結果をスコア順に表示する機能。
- ◆ 2015.9.27に正式リリース (Thomson Innovation Ver4.5にて)

<Thomson Reuters資料より>

スマートサーチ機能

・スマートサーチ機能は、簡単に入力した単語から、Thomson InnovationでDWPIコンテンツを利用し、自動的に検索語を広げ、外国語から英語、IPCを含めた検索を行い、検索結果を関連度スコア順に表示する機能です。 ※V4.3ではベータ版のご提供となります。

1 検索対象のコレクション: すべて

2 検索条件を入力: テンプレートを表示

Smart Search-トピック ↑ laser eye surgery BETA



3-1) Smart Search

<利点>

- ◆ キーワードを入れるだけでスコア順に表示されるので簡単かつ効率的なスクリーニングが可能。

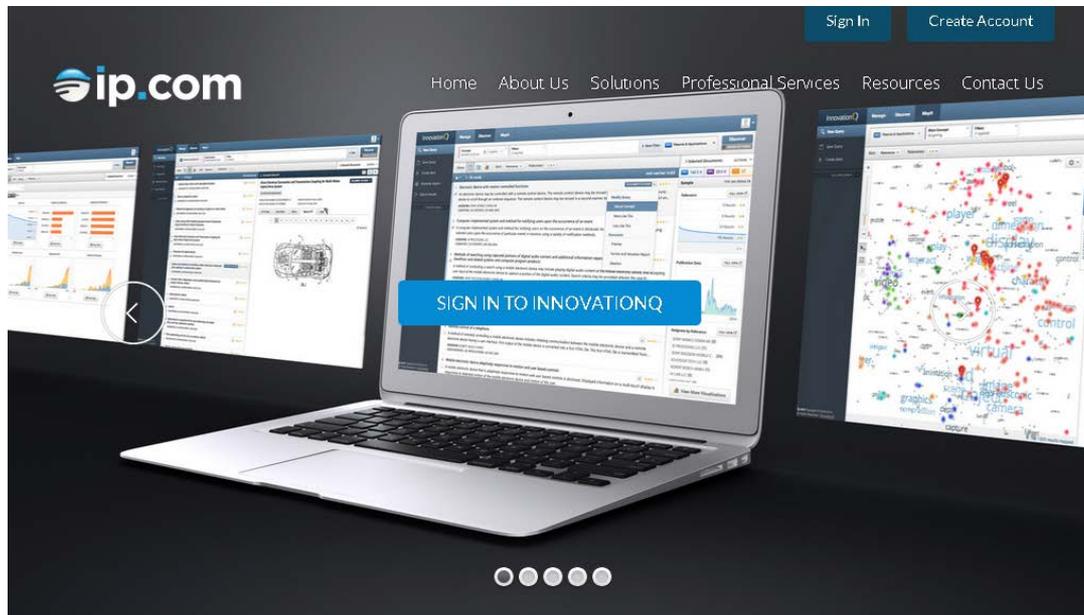
<所感>

- ◆ 制限された入力単語数の中で、適切なキーワードを選ぶことがポイントとなる。



3-2) InnovationQ

- ◆ 開発元 : IP.com
- ◆ 一種の概念検索。「Semantic Gist[®]」というサーチエンジンを使用してキーコンセプトを解析し、検索をする。
- ◆ USPTOの審査官が採用しているという情報あり。
- ◆ 特許のみならず非特許文献、ライセンス技術も検索可能。





3-2) InnovationQ

<特徴的な機能>

◆ Semantic search (意味解析・検索機能)

- 文献番号やテキストを入力することで、類似のコンセプトを有する文献を検索する機能。
- テキストの入力をするだけで基本的な検索が可能。
(番号検索も可能。尚、JP番号による検索も可能となっている。)
- 入力するテキストは英語で、入力できる文章量に特に制限はない。
ただし、6 words以上の文章が推奨されている。

◆ Concept Modifier (重み付け変更と再検索)

- 検索対象のコンセプトに、

近づけたい文献を指定するコマンド=>"More like this"
遠ざけたい文献を指定するコマンド=>"Less like this"

◆ MapIt

- 検索文章を解析した結果から、関連性が高いと判断したキーワードをビジュアル化する機能。



3-2) InnovationQ

<利点>

- ◆ 検索集合の内容を確認した上で、その集合をさらにモディファイできる点や、入力した文章を解析して**キーコンセプト**を創出する点は、他の検索システムにはない優れた点である。
- ◆ 客観的に示し仕切れていないが、既存の概念検索システムに比べて、その検索結果の精度は高い印象を受けた。
- ◆ 解析機能を使う場合も、検索集合のエクスポート／インポート等の作業がなく、簡便である。

<所感>

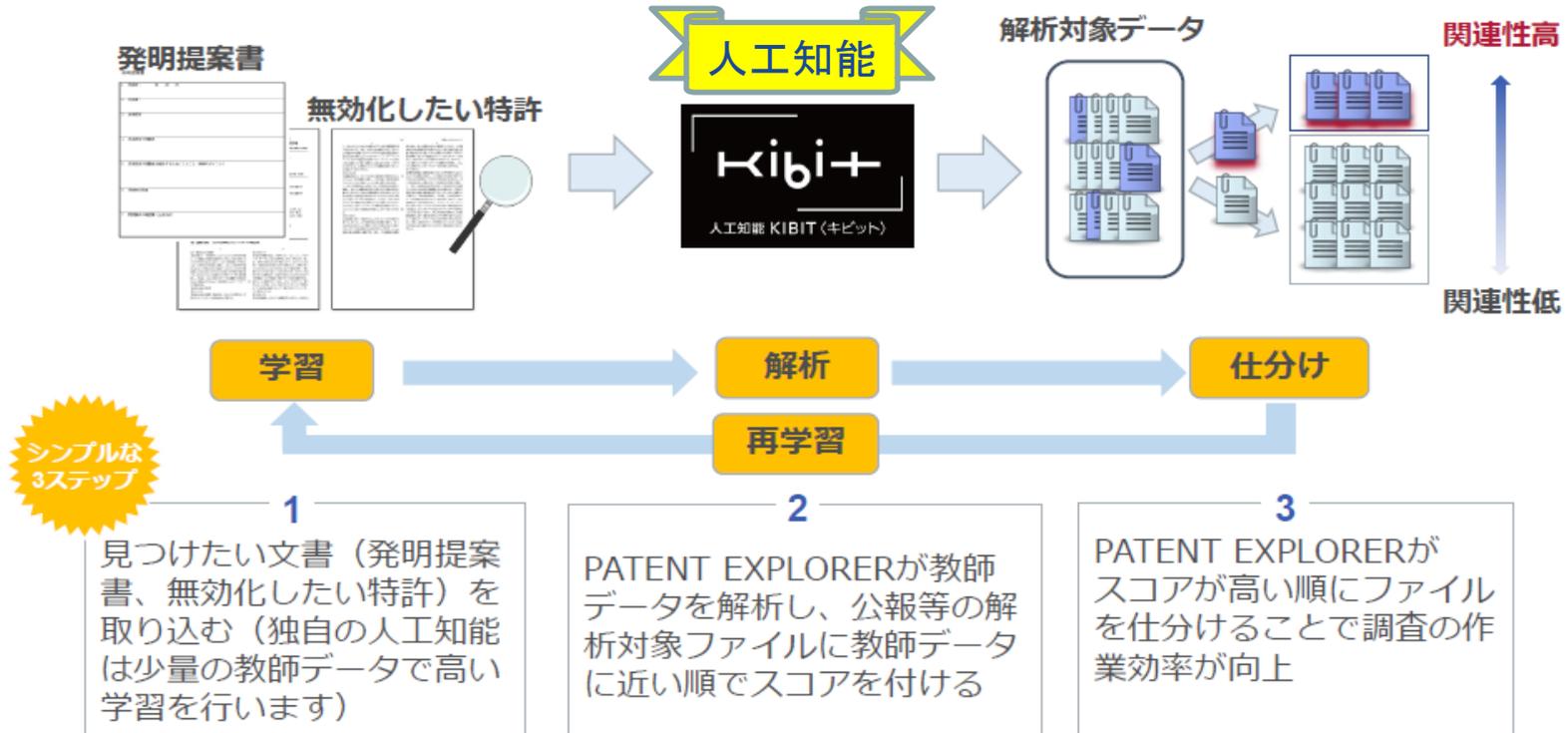
- ◆ 多機能であるので、どのような用途に特に力を発揮するのか未だ図りかねるところはあるが、その一つとしては、開発者とのブレインストーミングなどには有用であろう。



3-3) Lit i View PATENT EXPLORER

- ◆ 開発元 : 株式会社UBIC
- ◆ 人工知能を活用した調査支援ツール
 - 教師付き機械学習(※)によるスコアリング
- ◆ 2015.11より有償トライアル版リリース

※事前に正解をツールに教えておき、その正解に近いものを抽出させる。



(引用元:UBIC社紹介資料)





3-3) Lit i View PATENT EXPLORER

<利点>

- ◆ 使用感としては概念検索とほぼ同じ。
 - 検索クエリはフリーテキスト(word, text)、公報データ自体でもOK。
- ◆ 同義語、関連用語などの辞書データは不要。
- ◆ (ツールの仕様上厳密な比較は不可能だが)
検索精度は既存の概念検索よりもよりの良い印象を受けた。



4) 画像系情報検索ツール

4-1) Global Brand Database (WIPO)

– 商標データベース

4-2) Graphic Image Park (INPIT)

– 画像意匠公報検索支援ツール

4-3) Google画像検索

4-4) 文献画像検索システム(JST)

– 試験運用版



4-1) Global Brand Database (WIPO) 商標

ここに画像ファイルを
ドラッグ & ドロップ

4種類のFilter

WIPO WORLD INTELLECTUAL PROPERTY ORGANIZATION

Home Reference Global Brand Database

Global Brand Database

Data from the Republic of Korea available Over 3,000,000 records added

Moldovan data available Over 35,000 records added

German data available Over 1,800,000 records added

PHM data available Over 1,250,000 records added

Lao data available Over 33,000 records added

Tonga data available Over 3,000 records added

Japan data available Over 1,700,000 records added

SEARCH BY Brand Names Numbers Dates Class Country

Text = e.g. wipo OR ompl, "ntel", ompl-

Image Class = e.g. 05.07.13, apple AND tree

Goods (All) = e.g. footwear, comput*

FILTER BY Source Image Status Origin App. Date Expiration

Pick an image

browse or drag an image here

Pick a strategy

Shape Color Texture Composite

Pick an image type

Verbal 2,648,986

Nonverbal 1,520,556

Combined 6,774,872

Unknown 277,370

1 - 30 / 24,668,419 TMview

Brand	Source	Status	Score	Origin	Holder	Number	App. Date	Image Class	Nice Class	Image
ROKI	EM TM	Pending	1	EM		014901862	2015-12-13		25, 30, 32	
RAYPACK	EM TM	Pending	1	EM		014901854	2015-12-13		9, 35, 42	
POWERBALL	EM TM	Pending	1	EM		014901898	2015-12-13		9	
SEDI HISHAM	EM TM	Pending	1	EM		014901847	2015-12-13		29, 30, 31	
No Verbal Elements	EM TM	Pending	1	EM		014901839	2015-12-13		9, 25, 35	
LAKI	EM TM	Pending	1	EM		014901871	2015-12-13		25, 30, 32	
KUBI ICE TEA	EM TM	Pending	1	EM		014901888	2015-12-13		28, 30, 32	
appful	EM TM	Pending	1	EM		014906077	2015-12-13		9, 42	
GRILL ROCKER	EM TM	Pending	1	EM		014906085	2015-12-13		11, 21, 25	
praguesirens	EM TM								41	
TORO SPORT	EM TM								25, 26, 28	
GIANNINI ROMA	EM TM								12, 35, 37	

WIPOが提供する商標データベース。
画像による検索が可能





4-1) Global Brand Database (WIPO) 検索例



Shape

線の分布

Color

色の分布

Texture

線のタイプ

Composite

線と色の分布





4-2) Graphic Image Park (INPIT) 意匠分野

Graphic Image Park
画像意匠公報検索支援ツール

ヘルプデスク (受付時間: 平日 9:00-17:00)
☎ 0120-633-163
✉ E-mailでのお問合せはこちら

ガイドライン ヘルプ JPO INPIT J-PlatPat

画像を入力 ファイルから選択

Drag&Drop

モードを選択 標準 形 色 標準+90°回転
 単一部品 複数部品

絞り込み条件を選択

年月日 指定しない 出願日 登録日 公報発行日
例) 20140101 から 例) 20150101

意匠に係る物品

意匠分類 日本意匠分類 (特許庁HP)
例) H76? H772?

ご利用になる前に 必ずお読みください

- このツールは、画像意匠分類 (W) が付与された登録意匠の公報と、その公報に掲載された、機器の操作画像等の画像を蓄積しています。
- ソート結果として表示されるサムネイルイメージは、機械的な照合によって入力画像に近いと評価された順に表示されるものであり、意匠法上の「類似」する意匠が必ず上位に表示されるとは限りません。
- このツールに利用者がアップロードした入力画像は、システム内に蓄積されることはありません。



ガイドライン

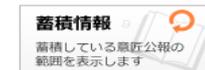
ご利用の際の留意事項やお役立ちメモを記載しております。

お知らせ

2015年12月07日
新たに、意匠登録第1539253号(11月30日公報発行)までの13件の公報を蓄積しました。



アンケート
のお願い



蓄積情報
蓄積している意匠公報の
範囲を表示します



▶ プライバシーポリシー ▶ アンケート



独立行政法人 工業所有権情報・研修館
National Center for Industrial Property
Information and Training

Copyright(c); 2015 JPO and INPIT

- このツールは、**画像意匠分類 (W) が付与された登録意匠の公報**と、その公報に掲載された、機器の操作画像等の画像を蓄積しています。



~世界から期待され、世界をリードするJIPA~

Confidential



4-2) Graphic Image Park (INPIT) 検索例1

入力画像



モード: 標準



モード: 色

モード: 形





4-3) Google 画像検索 — カラー画像 —

※ 一般的な画像検索

Google

すべて **画像** ニュース ショッピング 地図 もっと見る▼ 検索ツール

約 100 件 (0.68 秒)



画像サイズ:
1600 × 1201

この画像の他のサイズは見つかりません。

この画像の最良の推測結果: [東京 ゲート ブリッジ](#)

東京ゲートブリッジ - Wikipedia

<https://ja.wikipedia.org/wiki/東京ゲートブリッジ> ▼

東京ゲートブリッジ(とうきょうゲートブリッジ、Tokyo Gate Bridge)は、東京港第三航路(東京東航路)を跨ぎ中央防波堤外側埋立地と江東区若洲を結ぶ橋梁である。東京3影影計画に位置付けられた港湾施設(幹線臨港道路)・東京港臨海道路の1期事業区間 ...

東京ゲートブリッジ観光レジャー|東京都港湾局公式ホームページ

www.kouwan.metro.tokyo.jp/kanko/gatebridge/ ▼

トピックス 東京ゲートブリッジのライトアップについて 下記のとおり特別ライブアップを実施いたします。なお、現在の予定のため変更になることもあります。ご了承ください。実施日、内容、カラー、10月1日、ピンクリボン月間、ピンク(10月の設定色 秋景色)、10月16 ...

類似の画像

画像を報告



一致した画像を含むページ

東京ゲートブリッジと富士山|空の風景]



524select.com > photography > 夕日・夕焼けの写真 ▼

800 × 450 - 2012/12/01 - 東京ゲートブリッジと富士山の撮影ポイントからの写真です。夕焼けの東京ゲートブリッジと富士山は都内ならでは美しい夕焼け撮影ポイントです。

東京ゲートブリッジと夕焼け|Memorial]



524select.com/archives/1343 ▼

800 × 450 - 2011/09/13 - 2011年9月13日の東京湾の夕景です。東京ゲートブ

写真によって特定できるかは異なる

Google

すべて **画像** ニュース ショッピング 地図 もっと見る▼ 検索ツール

約 3 件 (0.42 秒)



画像サイズ:
1600 × 1199

この画像の他のサイズは見つかりません。

一致した画像を含むページ

The World's Best Photos of 東京ゲートブリッジ - Flickr Hive Mind



flickrhivemind.net/Tags/東京ゲートブリッジ ▼

240 × 159 - 1680mmf3545za 2012 a77 airplane architecture bay bridge cloudy sky construction crane f28 gate japan koto night park rinkaiohashi sd15 sea sky s1ta99v sony tokyo tokyobay tokyogatebridge underconstruction urbanrenewal wakasu ...

東京ゲート - Who is talking about 東京ゲート on FLICKR



whotalking.com/flickr/東京ゲート ▼

240 × 159 - Check who's talking about 東京ゲート right now on FLICKR.

東京ゲートブリッジ - WhoTalking.com



whotalking.com/flickr/東京ゲートブリッジ ▼

240 × 159 - Check who's talking about 東京ゲートブリッジ right now on FLICKR.

類似の画像

色々な橋



ゲートブリッジのみ

カラー写真の検索精度は非常に高い
画像(写真)が何であることを特定できれば特定したもので検索



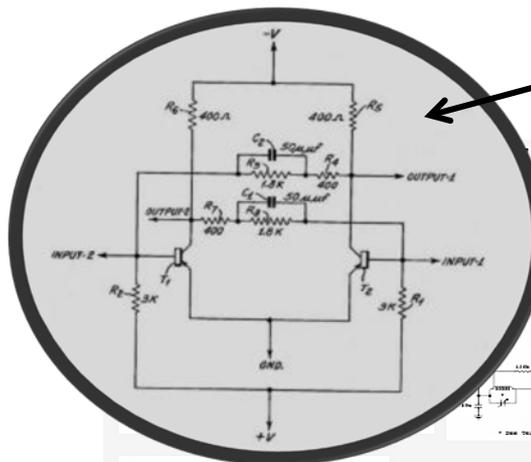


4-3) Google 画像検索 - カラー画像 -

※ 一般的な画像検索

特許図面での検証

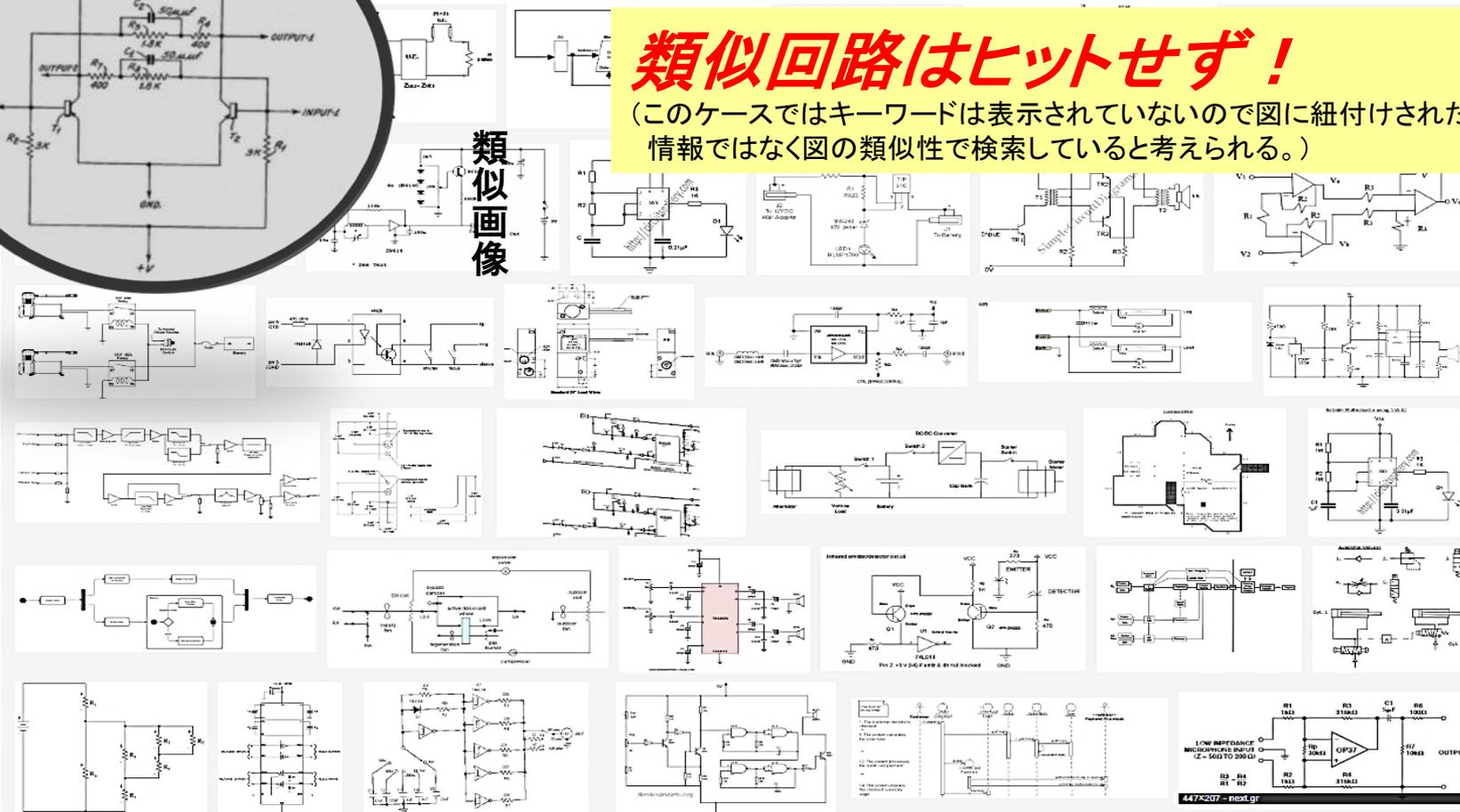
Googleの画像検索で、この図を検索項目として入力。



類似回路はヒットせず!

(このケースではキーワードは表示されていないので図に紐付けされた情報ではなく図の類似性で検索していると考えられる。)

類似画像





4-4) 文献画像検索システム(JST)試験運用版

※ 運用元(JST・JAE)の都合により、システムの運用を2016年4月で終了

文献画像検索システム 科学技術振興機構 (JST) & 日本地震工学会 (JAE)

従来、論文検索はキーワードで、論文に付与されているタグや、全文中の文字を検索するのが一般的でした。
この度、JSTはJAEの協力も得て、論文に掲載されている画像から論文を検索する。更に画像も画像につけたタグでは無く、類似画像検索を使って文字を介せずに検索するシステムを構築しました。このようなシステム構築は我が国でも最先端の取組であり、今後活用例を増やして行くため、活用例を紹介して行きます。

JST文献画像検索システムへはこちらから

【活用例 1 ; 手元画像との類似検索】

- 手元の画像ファイル（手描きグラフのスクリーンデータでも可）の類似した画像を検索できる。
- 画像を直接使うので、曲線の特徴や色といったキーワードで指定しにくい絞り込みも可能。
- これらの画像を開くと掲載論文に飛ぶことができるので、このようなデータを得た条件を詳しく調べることができる。
- そこから、何か共通する条件が見つかるかもしれません。

上の図は、手描きの減衰曲線をスキャンして類似画像を検索した結果である。応答スペクトルの他、対数曲線やローカットフィルタの図がヒットしている。

(お願い)
本取組は、新たな取り組みなので、まだ十分な活用例が生まれておりません。
今後、皆様方の協力を得て、活用例を紹介しつつ、利用を高めていきたいと考えておりますので、良い活用例を見つけて皆さんに紹介しても良いと言われる方は、fs-admin@jst.go.jpまで連絡を頂けましたら幸いです。





5) 総括

文章系情報検索ツール

- ◆ 検索式検討の**負荷削減**の期待大
 - 簡単な単語、提案書・公報自体を検索キーにできる。
 - 特許分類や同義語・関連用語の**検討そのものが不要**。
- ◆ 機能的に出願前調査、無効資料調査向き
 - 近い特許文献をピンポイントで拾ってくるのが得意な印象。
 - 網羅性重視のクリアランス調査には不向きか。
- ◆ 高機能化(=**ブラックボックス化**)の代償
 - なんとなく良い検索結果は出ているが、なぜ？この結果が出たのかはよくわからない。
 - ⇒ 原因追求ができないので、トライアンドエラーは増加
 - コンピュータが評価した結果であなたの上司は納得するか？

導入への
ハードル



5) 総括

画像系情報検索ツール

- ◆ 写真やシンプルなイラストなどを使った類似画像検索は比較的高いレベル
 - 特にカラー(色情報)のものは、ヒットする割合が高い傾向。
- ◆ 商標、意匠などの画像検索システムも提供され始めているが、まだ十分な抽出精度を有しているとは言い難い
- ◆ **画像データ(特許図面)を用いての特許文献検索は時期尚早**
 - 特許図面は、商標、意匠のように形(模様)や色情報を持ち合わせてはならず、大半が「機能」や「メソッド」を表現している。そのため、従来手法(画像認識)では実現が難しい。



6) 提言

文章系情報検索ツール

- ◆ 更なる検索作業量のミニマム化
 - 検索範囲(母集合)の簡易設定。
- ◆ 検索結果の再現性の保証化
 - 検索者が検索式を立てる必要が無い検索システムでの再現性の保証方法の提案。
(その文献は、「なぜ漏れたのか？」または「なぜヒットしたのか？」を検証するときに必要。)
- ◆ 多言語対応
 - 特許公報の言語を意識する必要がないセマンティック検索ツールの開発。



6) 提言

画像系情報検索ツール

- ◆ 白黒線画(フリーハンド画像)に対応した高精度画像検索技術開発。
- ◆ 図面版セマンティック検索(図面版概念検索)の研究促進。
 - 特許図面の形状認識と、図面上の表記文字及び引き出し記号の認識(意味理解)より、**検索図面の機能的な意味を把握した上での特許文献検索システム。**
- ◆ ナレッジグラフのアーキテクチャを活用した特許図面検索システム。
- ◆ 特許/論文図面のディープラーニング。