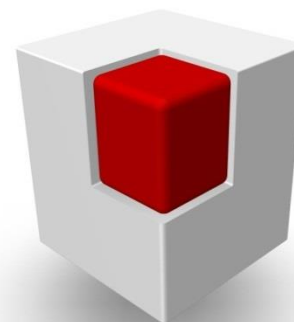
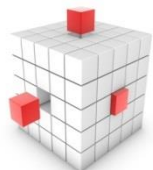
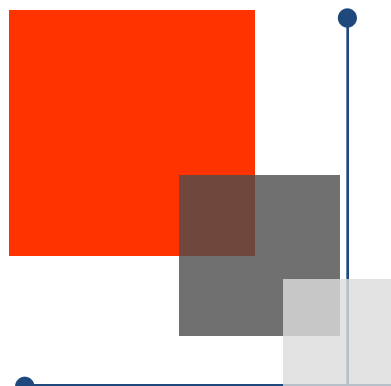


構造化データと非構造化情報の 統合ソリューション

～サーチアプリケーションSMART/InSightとTMSの連携～

SMART/InSight® G2 Open
Generation2

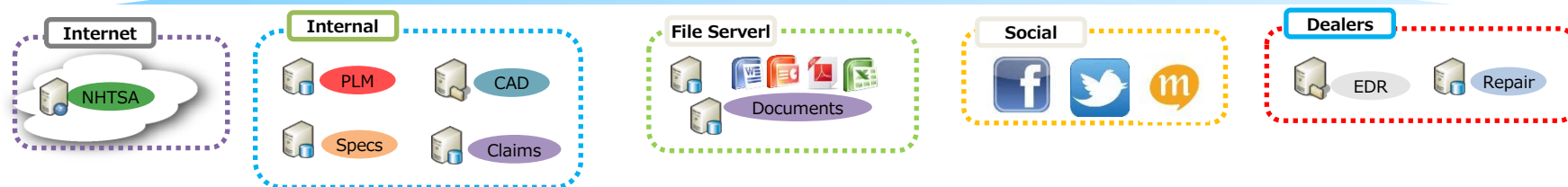


平成25年11月22日（金）
SMARTInSight Corporation
万代 豊

SMART/InSight G2 Openとは…

SMART/InSight G2 Open
Generation 2

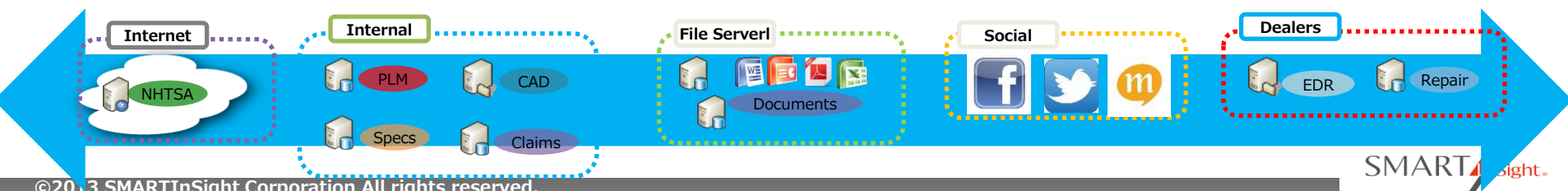
構造・非構造の非常に多彩なデータ
を対象にアドホックな可視化・分析
シナリオに対応する為に従来の
RDBMSでは無くサーチエンジンを
活用した統合型アプリケーション基盤

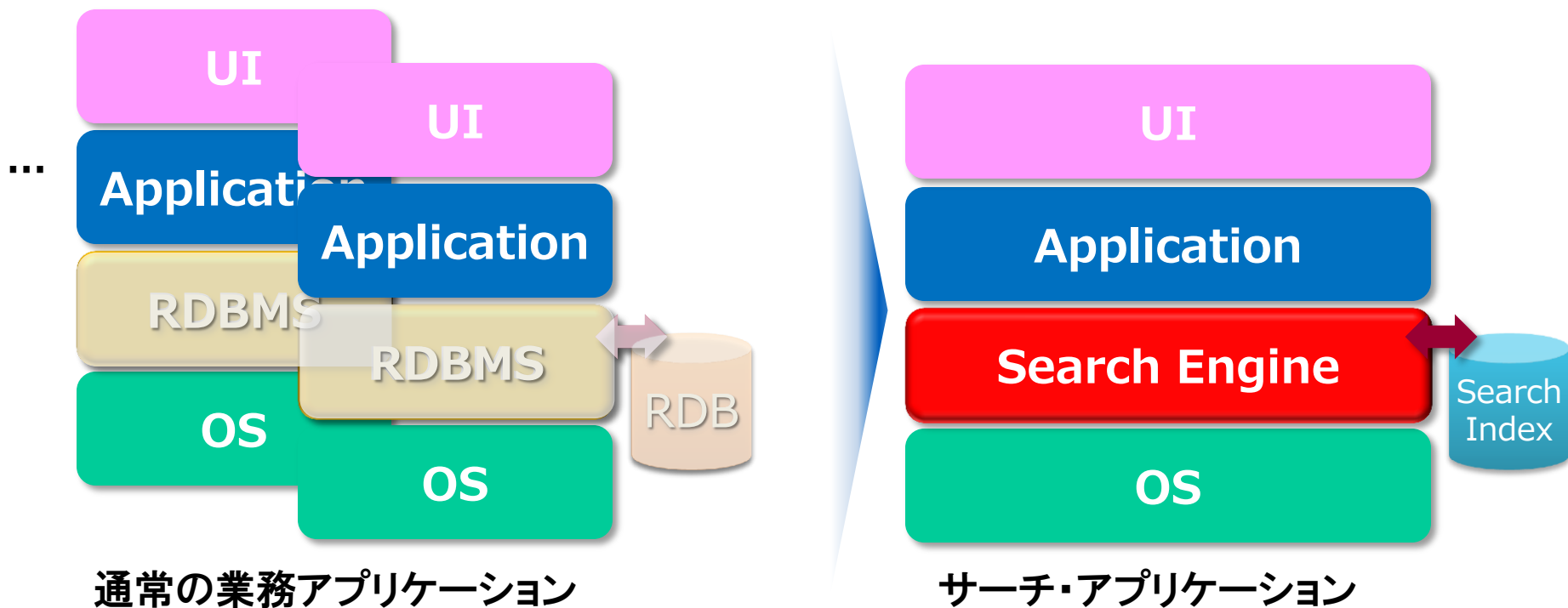


RDBMSではなく
サーチ基盤を活用するので…..

SMART/InSight G2 Open
Generation 2

- BI/DWHの様な分析軸-分析用スキーマの設定が一切不要
- RDBMSベースの分析システムに比較して圧倒的な性能
- 構造・非構造情報を連携した可視化分析システムの構築が柔軟





- 構造化データ(Relational/Structured)→非構造化(Non-Relational/Unstructured)データの扱いと
- より大規模なデータへの対応を
- 圧倒的なスケーラビリティとコストパフォーマンスで提供！

Search-Based Applicationのイメージ

SMART/InSight® G2 Open
Generation2

通常のサーチェクスペリエンス

NTT DATA 株式会社 NTTデータ 数理システム
Global IT Innovator — 数理科学+コンピュータサイエンス —

お問合わせ | アクセス案内 | 採用情報 | サイトマップ | English

Googleカスタム検索

基盤技術 | ソリューション | 活用事例 | イベント・セミナー情報 | 製品 | 会社案内

テキストマイニングツール

Text Mining Studio

真の情報を引き出す。多角的に分析する。
やりたい分析が思いのまま。
テキストマイニングは新しい局面を迎えました。

お知らせ & セミナー

お知らせ イベント・セミナー

- 2013/09/10 サーバメンテナンスのため、9月14日～16日にかけて数分程度の接続が断続発生します。
- 2013/08/23 事務所移転に伴うサーバ移動のため、8月30日(金)17時半頃からサーバを停止します。
移転作業中は、電話、ウェブページ、メール等が不通となります。ご了承ください。
8月31日(土)19時頃から接続が回復する予定です。
- 2013/05/29 Text Mining Studio が日本統計学会から「ソフトウェア賞」を受賞しました。
- 2013/05/25 株式会社NTTデータ数理システムのホームページリニューアルが完了しました。
- 2013/04/22 ビッグデータ分析ツール Big Data Module Ver 1.1 リリース致しました。
- 2013/04/06 平成24年度データ解析コンペティション(伊藤健定部門)に出場し最優秀者を受賞しました。
(12/6)
- 2013/04/02 汎用シミュレーションシステム S³ Simulation System Ver 3.1 リリース致しました。
- 2013/02/15 Visual Mining Studio 7.3 リリース致しました。
- 2013/02/15 Text Mining Studio 4.2 リリース致しました。

数理システムユーザーフォーラム

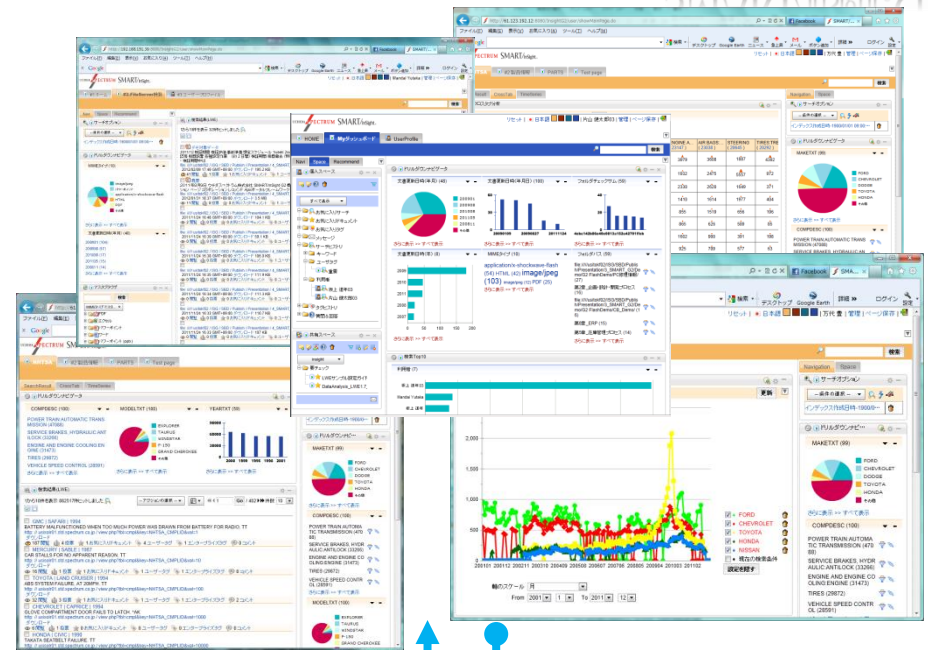
製品情報

株式会社NTTデータ数理システム
Q&A
株式会社NTTデータ数理システム
Q&A

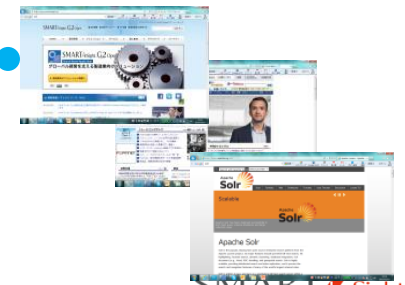
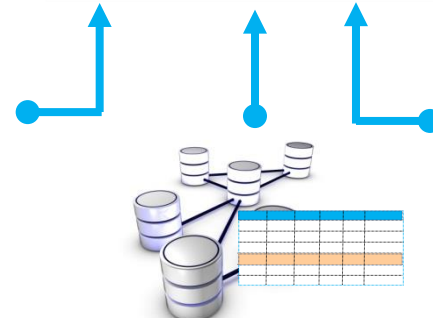
活用事例
株式会社NTTデータ数理システム
技術が応用される場所
活用事例をご紹介

製品情報

NUOPT
汎用統計計画パッケージ
Visual Mining Studio
汎用データマイニングシステム
Text Mining Studio
テキストマイニングツール
S-PLUS
統計解析ツール
Big Data Module
ビッグデータ分析ツール
S³ Simulation System
汎用シミュレーションシステム



SMART/InSight® G2 Open
Generation2



SMART/InSight®

DEMO

自動車メーカ品質保証部門

- ◆ メーカー: TOYOTA, 年度: 2010, 車種: PRIUS, 部品: ブレーキ、油圧系を指定、697件がヒット
- ◆ 個別クレーム帳票、詳細が表示

The screenshot displays the NHTSA Safercar.gov website. The main window shows the 'Complaints - Search Results' page. The search criteria are: Report Date: June 7, 2011 at 12:45 AM, Search Type: VEHICLE, Year: 2010, Make: TOYOTA, Model: PRIUS. The results show 697 records displayed. The search results table lists the following information:

Make	Model	Year	Component
TOYOTA	PRIUS	2010	SERVICE BRAKES, HYDRAULIC

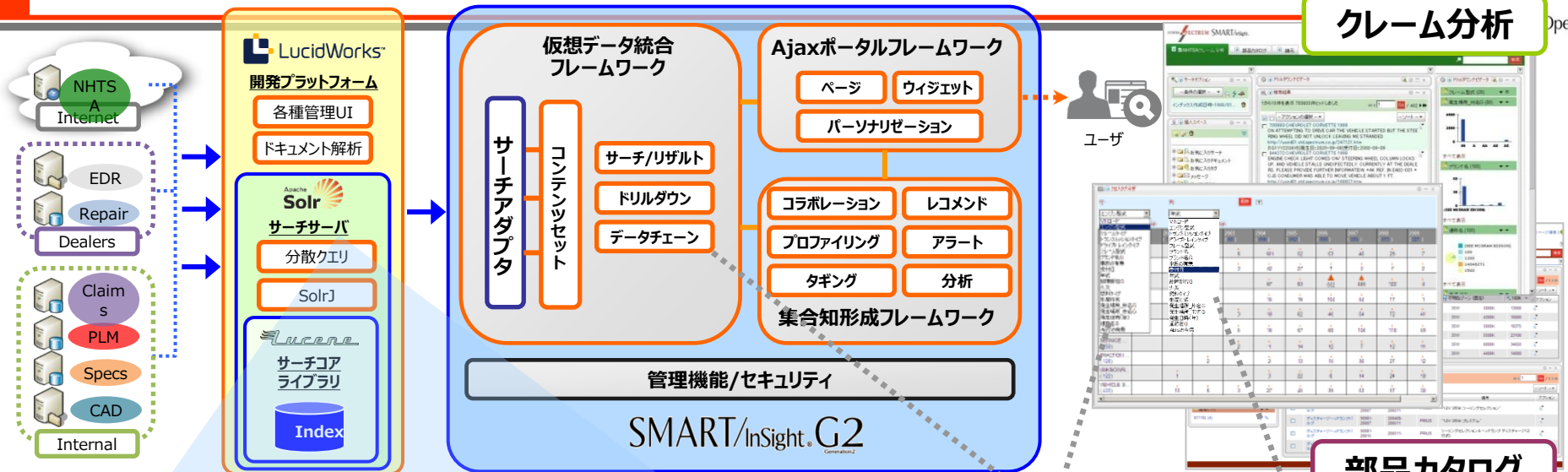
The 'Document Search' button is visible. A 'Complaint Summary' window is overlaid on the right, showing the following details:

Make	Model	Year	Complaint Number
TOYOTA	PRIUS	2010	10402158

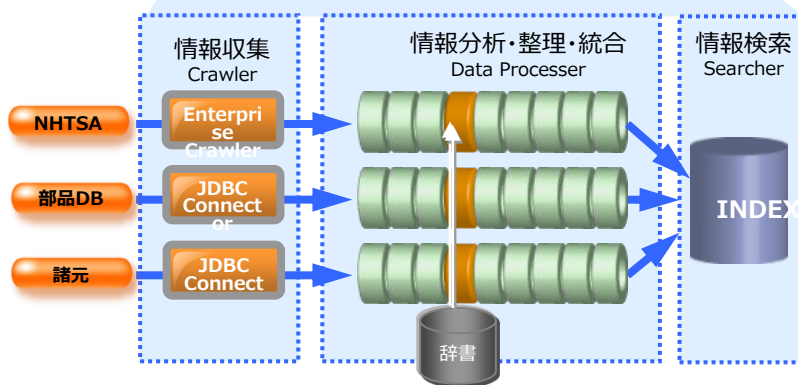
The summary text reads: 'TL*THE CONTACT OWNS A 2010 TOYOTA PRIUS. THE CONTACT STATED WHILE DRIVING SPEEDS IN INCLEMENT WEATHER AND TRYING TO DECELERATE THE VEHICLE WOULD STOP MOVING AND IT CRASHED INTO THE REAR OF ANOTHER VEHICLE. THERE WERE NO INJURIES FROM EITHER VEHICLE. A POLICE REPORT WAS FILED. THE VEHICLE WAS NOT BECAUSE IT WASN'T A MAJOR CRASH. THE MANUFACTURER WAS THEN CONTACTED AND ADVISED HIM THAT THEY WERE UNAWARE OF ANY FAILURES IN REGARDS TO THE BRAK FAILURE MILEAGE WAS APPROXIMATELY 65,000.'

The bottom of the screen shows the Windows taskbar with the time 13:52 and date 2011/06/07.

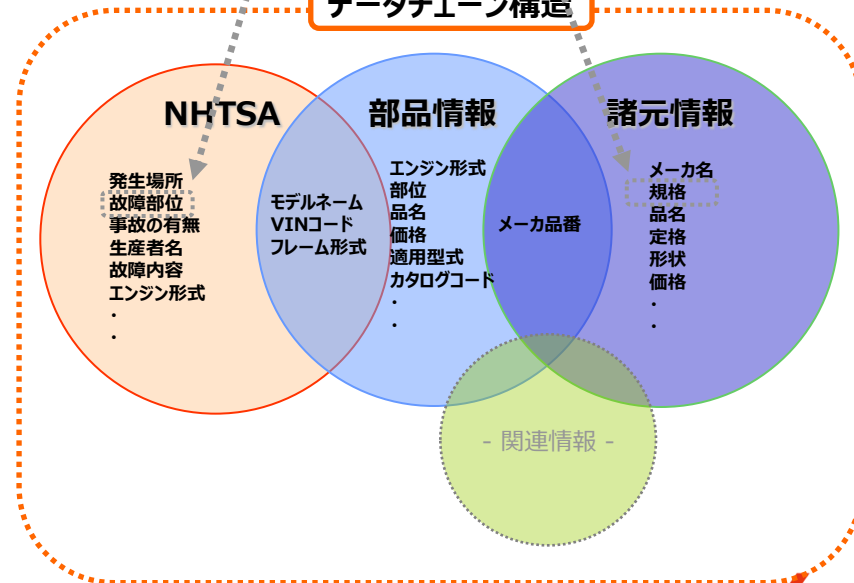
クレーム分析



部品カタログ



データチェーン構造



情報の取り込み

- Webサイト
- DataBase

情報の付加価値化（処理の中核）

- VINコードから
- フレーム形式生成
- エンジン形式生成
- 年式の生成
- モデルの特定

索引（インデックス）生成

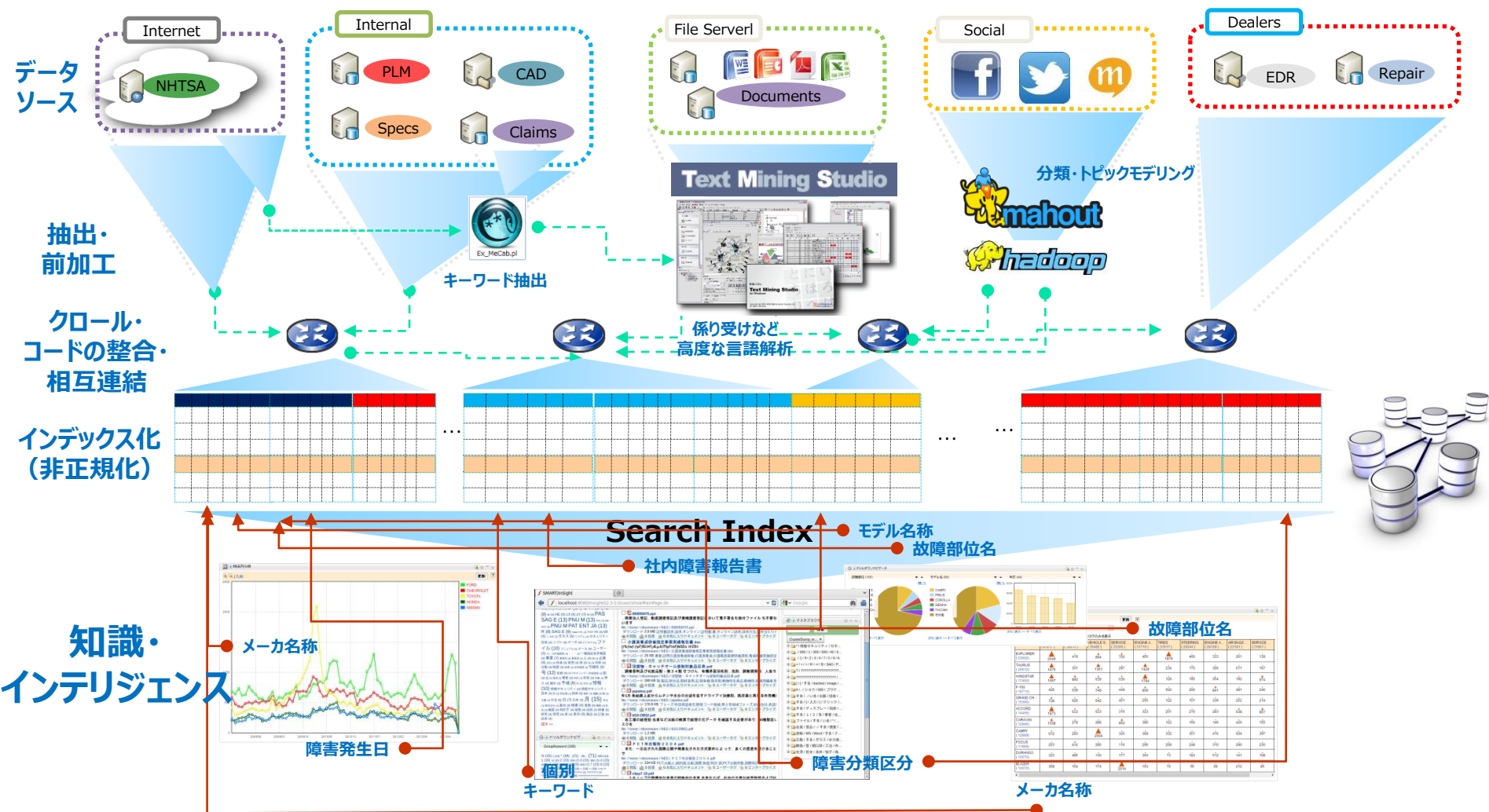
- フィールドのマッピング

--- 検索精度、適合性を上げるために、ロジックを組み込み ---



対象データソースからの情報をサーチインデックスに非正規結合、 単一レコードとして格納

SMART/InSight G2 Open



あらゆるフィールド値に着目した行レコード集約が瞬時に可能に

DBとの性能比較実証実験

600万件の取引明細抽出

~Star Schema Benchmark~

SMART/InSight G2 Open

```
mysql> select LO_ORDERKEY,LO_LINENUMBER,C_NAME,P_NAME,LO_REVENUE  
-> from LINEORDER_6M  
-> inner join CUSTOMER on LO_CUSTKEY = C_CUSTKEY  
-> inner join PRODUCT on LO_PARTKEY=P_PARTKEY  
-> where P_NAME='beige blue' and C_NATION='CHINA'  
-> ;
```

600万件の取引明細を対象に

顧客マスタと商品マスタを結合した上で、

商品名が'beige blue'で仕向け先が'CHINA'の取引明細を抽出

LO_ORDERKEY	LO_LINENUMBER	C_NAME	P_NAME	LO_REVENUE
12644	4	Customer#000023149	beige blue	926801
239335	1	Customer#000014530	beige blue	3722668
431780	2	Customer#000013600	beige blue	3131955
460931	2	Customer#000026239	beige blue	3636219
614720	2	Customer#000019799	beige blue	4347601
711654	3	Customer#000017135	beige blue	433353
870819	2	Customer#000006763	beige blue	6601407
894151	1	Customer#000006232	beige blue	2454298
986759	5	Customer#000017917	beige blue	2453035
1071559	7	Customer#000016060	beige blue	734073
1165701	1	Customer#000017677	beige blue	1380523
1189860	2	Customer#000017885	beige blue	5866943
1200578	2	Customer#000008498	beige blue	1665570
1268577	3	Customer#000027832	beige blue	4922555
5487395	6	Customer#000019534	beige blue	5935410
5548161	3	Customer#000000124	beige blue	6353645
5661792	3	Customer#000025930	beige blue	5390733
5672324	1	Customer#000026134	beige blue	3479029
5703041	2	Customer#000018044	beige blue	1498056
5895681	5	Customer#000019423	beige blue	5414495
5955623	3	Customer#000011167	beige blue	5426740

79 rows in set (41.42 sec)

79件の明細を抽出する応答
時間は41.42秒

同じ条件でのSMART/InSightの絞り込みは… 瞬時！

SMART/InSight G2 Open

商品名='beige blue' かつ
仕向地='CHINA'
の絞り込み

The screenshot displays the SMART/InSight G2 Open interface with several panels:

- Search Options (検索オプション):** Shows filters for SSB_P_NAME (beige blue), SSB_C_NATION (CHINA), and a keyword search.
- Drill-down Navigator (ドリルダウンナビゲータ):** Displays three charts: SSB_P_COLOR (35) with a pie chart showing yellow, steel, puff, drab, smoke, and others; SSB_P_CATEGORY (20) with a bar chart showing MFGR categories; and SSB_LO_SHIPMODE (7) with a pie chart showing TRUCK, REG AIR, RAIL, FOB, SHIP, and MAIL.
- Multiple View Navigator (多次元ナビゲータ):** Shows a hierarchical tree structure for SSB_C_REGION, SSB_C_NATION, and SSB_C_C, with a focus on ASIA (79) and CHINA (79).
- Search Results (検索結果):** Displays a list of results, including details for Customer#000023149, Customer#000014530, and Customer#000019508, all for the product 'beige blue'.
- Drill-down Navigator (ドリルダウンナビゲータ):** Shows a chart for SSB_P_NAME (1) with a single data point for 'beige blue'.
- Drill-down Navigator (ドリルダウンナビゲータ):** Shows a chart for SSB_C_REGION (1) with a single data point for 'ASIA'.

LWSから返されたResponse Headerの内容

```
-<response>
-  <lst name="responseHeader">
    <int name="status">0</int>
    <int name="QTime">137</int>
  +<lst name="params"></lst>
</lst>
-<result name="response" numFound="79" start="0">
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
  +<doc></doc>
</result>
+<lst name="facet_counts"></lst>
+<lst name="highlighting"></lst>
</response>
```

クエリ応答時間は137msec
(0.137sec)

79件のヒット

検索結果としての表示用
ドキュメント10件

フルテーブルスキャンの発生しないサーチの
性能はMySQLと比較して300倍以上
高速である事を確認

自動クラスタリング・ラベリングの実行例

~NHTSAクレーム情報~

UCHIDA SPECTRUM SMART/InSight. リセット | 日本語 | Mandai Yutaka | 管理 | ページ保存 |

mahout NHTSA parts

グループ検索結果

1から10件を表示 2722件ヒットしました

2,722件のエアバッグ障害、搭乗者負傷のトピックのクレームがある事が判明

2,722件は全てエアバッグ障害、搭乗者負傷のトピックのクレームのみ。

クラスタリング・トピックモデリング等の技術によりラベル抽出を行う事で、大量のドキュメントから意識すらしていないキーワードさえも発見する事が可能に

検索

MAHOUT_keyword-AIRBAG / PASSENG...

MAHOUT_keyword (1)

AIRBAG / PASSENGER / THE / AIR / LIGHT / BAG / SEAT / ON / AIRBAGS (2722)

MIMEタイプ (1)

マスタブラウザ

検索

キーワード - MAH...

AIRBAG / PASSENGER / THE / AIR / LIGH...

BRAKE / THE / TO / AND / I / ON / PEDAL / ...

CONTACT / THE / WAS / STATED / FAILU...

DOOR / THE / SIDE / WINDOW / AND / DRI...

FUEL / GAUGE / TANK / GAS / THE / AND / ...

I / THE / CAR / IT / TO / AND / MY / ON / W...

I / THE / THIS / IT / TO / IS / AND / HAVE / ...

RUSTED / FRAME / THE / BROKEN / TANK...

STEERING / POWER / THE / AND / TO / C...

THE / AND / TO / A / CAR / I / ON / IN / OF

THE / CONSUMER / DOOR / STATED / WA...

THE / CONTACT / WAS / FAILURE / VEHIC...

AS / VEHICLE / FAILU...

CONTACT / VEHICLE / ...

TO / OF / ON / LEF...

AND / A / IT / OF / ...

THE / WINDSHIELD / DASH / TOP / FROM

特定の1件まで2クリックで到達 ～トピックグループとキーワードに沿って直観を支援～

SMART/InSight G2 Open

The screenshot displays the SMART/InSight G2 Open web application interface. The browser address bar shows `localhost:8080/InsightG2.5.0.0/user/showMainPage.do`. The top navigation bar includes tabs for #1 ホーム, #2 検索, #3 ユーザープロフィール, #4 グループプロフィール, #5 スペースプロフィール, #6 監査, and a search bar with the text "NEG".

The main content area is divided into several panels:

- entkeyword (20)**: A list of keywords and topics, including "スマートフォン (1) セキュリティ対策 (1) ベースライン対策 (1) マニュアル (1) 対策 (1) 対象 (1) 情報 (1) 情報漏えい (1) 情報漏えい対策 (1) 搬媒体 (1) 暗号 (1) 暗号アルゴリズム (1) 暗号化 (1) 暗号製品 (1) 本書 (1) 業務 (1) 私的利用 (1) 秘密情報 (1) 紛失 (1) 編 (1)".
- グループ検索結果**: A search result for "000026760.pdf" with a download size of 2.1 MB. It includes a description: "情報漏えいを防ぐための モバイルデバイス等 設定マニュアル ～安心・安全のための暗号利用法～".
- 検索キーワードを知らなくても的確に目標にリーチ!!**: A yellow arrow points to the search result, indicating that the system can find the target even without knowing the search keywords.
- 重要キーワード群を確認、選択しながら...**: A yellow arrow points to the search result, indicating that the system can confirm and select important keyword groups.
- 想定したトピックグループを指定して...**: A yellow arrow points to the search result, indicating that the system can specify the intended topic group.
- 検索オプション**: A panel with a search bar and a dropdown menu for "条件の選択".
- マスタブラウザ**: A panel with a search bar and a dropdown menu for "ClusterDump_m...".

The bottom of the interface shows a JavaScript error message: `javascript:void(0)`.

- リレーショナルデータベースでは無く**サーチエンジン**をデータ格納・アクセスに活用（サーチアプリケーション）
- あらゆるデータソースを非正規結合して集約する為、データソース間の**「隠れた関係」の抽出が容易に**
 - リレーショナルデータベース(BI/OLAP)でも可能な設計だが….
 - 応答性能はBI/OLAPの数10秒～数時間のオーダーに対しサブミリ秒ベース
 - データベースに比べると全フィールドがインデックス対象
- データソースはRDB、ECM、ファイルサーバ、Web、ソーシャル等、**あらゆる構造・非構造情報**を対象
- 「Text Mining Studio」等、**外部連携が容易**
 - 分析結果をフィールドとしてレコードに追加、ドキュメントメタタグとして登録など

宜しくご検討の程、お願い致します。

SMART/InSight® G2 Open
Generation2

SMART/InSight® G2 Open
Generation2

Discover+ Your Business

エンタープライズサーチは、無意識に諦めて
いた発想の壁を取り払う所に価値がある。