

# VMS, TMS 新機能紹介

～データとテキストを融合した次世代の分析ソリューション～

株式会社 数理システム

岩本 圭介

2

## 内容

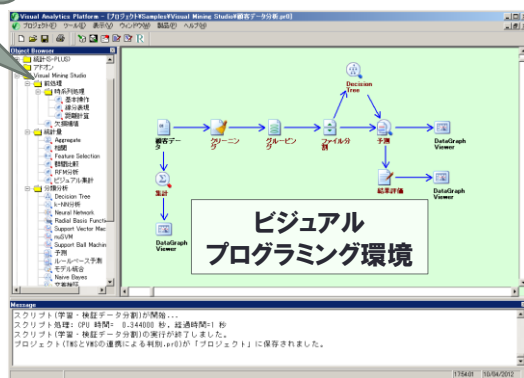
- 1. Visual Mining Studio (VMS)**
  - ・製品概要 と 次期バージョン での新機能
- 2. Text Mining Studio (TMS)**
  - ・製品概要 と 次期バージョン での新機能
- 3. データマイニング と テキストマイニング の連携**
  - ・連携の仕組み
  - ・デモンストレーション
- 4. リリーススケジュール**

## Visual Mining Studio

手軽な試行錯誤環境  
豊富なマイニング手法

…を兼ね備えた  
汎用データマイニングツール

- **完全自社開発**
  - ユーザ要望、提案を数多く取り込んだ**実用性の高い製品**
- 『より**手軽に**、より**正確に**、より**高速に**』



Visual Mining Studio

Visual Mining Studio 新機能

## マイニングプロセス高速化

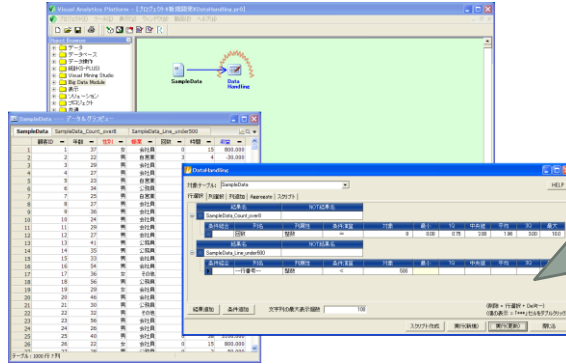
- 並列実行の適用対象を拡張
  - スクリプト関数レベル（ユーザ定義関数、VAPScript 固有関数）で複数コアでの並列実行を行うことが可能に
- 適用の場面
  - 大量データの分割処理 → 分割して処理を行った後に統合
  - モデルのパラメタチューニング → パラメータのみを変更し、学習過程を一括実行



- 新規スクリプト関数 `procedure_in_parallel`, `function_in_parallel` によって実現

# 高速データハンドリング

- マウス操作でインタラクティブにデータを操作・加工・可視化



- 行と列の選択 (フィルタリング)
- 計算式の適用
- 集計
- スクリプト実行

次期バージョンで

## 大幅高速化

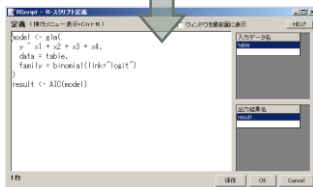
# R連携機能の強化(1)

- スクリプト記述なしに、GUI から R の機能を実行可能

既存機能



Rスクリプトアイコン

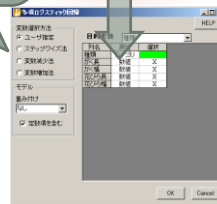


Rスクリプト記述

新機能



R処理アイコン



R分析パラメタ指定

Rの実行が  
より手軽に

## R連携機能の強化（2）

- R の処理に対しても、VAP が提供する
  - ビジュアルプログラミング環境
  - 並列実行の仕組み
 を提供
- R との連携を今後強化予定
  - R は、特にアカデミック分野では統計解析ツールとしての地位を確立
  - 金融工学・時系列解析・機械学習・バイオ等様々な分野への広がり



R + ビジュアルプログラミング  
R + ビッグデータ  
R の高速化

## 新規スクリプト関数群

- `sys_info()` 関数に `messageHandler` の追加
- `read_file()` 関数に “`fill_missing_by_min_val`” の追加
- 累積確率計算関数を追加

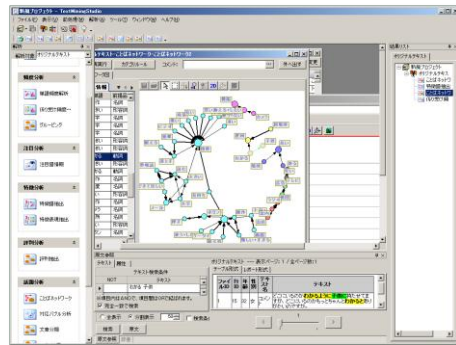
<code>prob_unif</code>	一様分布	<code>prob_beta</code>	ベータ分布	<code>prob_F</code>	F分布
<code>prob_norm</code>	正規分布	<code>prob_chisq</code>	χ二乗分布	<code>prob_exp</code>	指数分布
<code>prob_gamma</code>	ガンマ分布	<code>prob_t</code>	T分布	<code>prob_pois</code>	ポアソン分布

- 新関数 `is_na` 欠損か否かの判定
  - 実数は `HUGE_VAL`、整数は `LONG_MIN` を欠損表現とする
- 関数 `quantile` で、キー列指定(キー毎の計算)が可能に
- ユーザ定義並列実行関数を新設
  - `procedure_in_parallel`, `dump_procedure_def`
  - `load_procedure_def`, `function_in_parallel`

# Text Mining Studio

## テキストデータから有益な情報を抽出するための テキストマイニングツール

- **誰にでも**高度なテキストマイニングを
  - マウスの操作のみで分析が可能
  - 豊富な分析機能と強力なグラフ機能
- より**自由度**の高い分析を
  - カテゴリ機能等「意味」に着目した分析
  - データマイニングツール **Visual Mining Studio** とのシームレスな連携



Text Mining Studio

# TMS アドオンモジュール

## TextCutter

新機種の  
話題

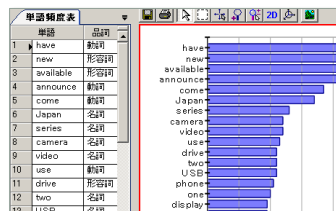
新機種XXX-01に乗り換えました。機能的にかなりイイ感じです。

飲食店の  
話題

機種変のあとで  
高校自体の友人と落ち合っ  
て焼肉屋へ。満足でした。

- 雑多なテキストを  
トピック毎に**分割**するツール
- **分析精度向上**のための  
前処理に威力を発揮
- 当社完全オリジナルの技術、  
**類似製品はありません！**

## 英語アドオン



- **TMS** 上で **英文** の  
テキストマイニングが可能に
- **係り受け**の抽出、  
**否定表現**の抽出にも対応
- 日本語分析と全く同様の使用感

## グラフ機能の全面刷新（1）

- ・グラフ表示機能をリニューアル
- ・ワンクリックでグラフの表示形態を変更可能に

皆様のご希望に  
お応えいたしました

The screenshot illustrates the new graph transformation capabilities. At the top, a toolbar contains icons for various chart types. A red dashed box highlights the '縦・横切替えボタン' (Vertical/Horizontal Switch Button), '折れ線グラフボタン' (Line Graph Button), and '円グラフボタン' (Pie Chart Button). Below, three windows demonstrate the results: a horizontal bar chart, a vertical bar chart, a line graph, and a pie chart. A large double-headed arrow at the bottom indicates the dynamic switching capability.

**動的な切り替えが可能に！**

## グラフ機能の全面刷新（2）

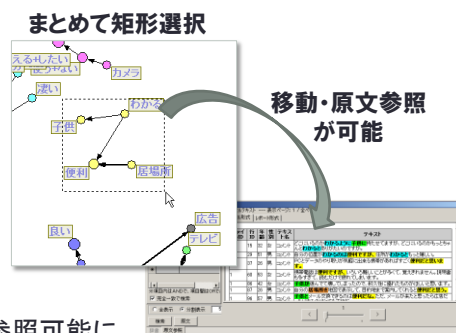
- ・複数種類グラフの同時作成をサポート
- ・グラフ上の任意の場所へのコメント挿入が可能

The screenshot shows the ability to create multiple graphs simultaneously and insert comments. On the left, a window displays '棒グラフ+円グラフの表示' (Display of Bar Chart + Pie Chart). On the right, a window shows 'コメントの挿入' (Comment Insertion). A 'テキスト追加' (Add Text) button is shown above the bar chart, with a red box indicating '注釈を挿入できます' (You can insert a note). A callout bubble explains that comments or titles can be entered and confirmed directly on the graph.

**コメントや  
タイトルなどは  
グラフ画面上  
その場で入力・  
確認が可能**

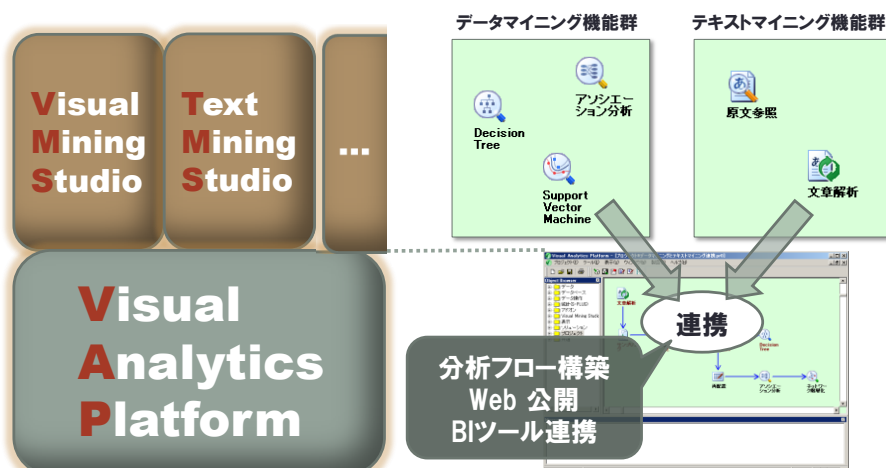
## 様々なユーザビリティ向上

- ネットワーク図の操作性向上
  - 矩形選択で選択した領域をまとめて移動・原文参照・ズーム
  - Ctrl+C によるクリップボードへの図のコピー



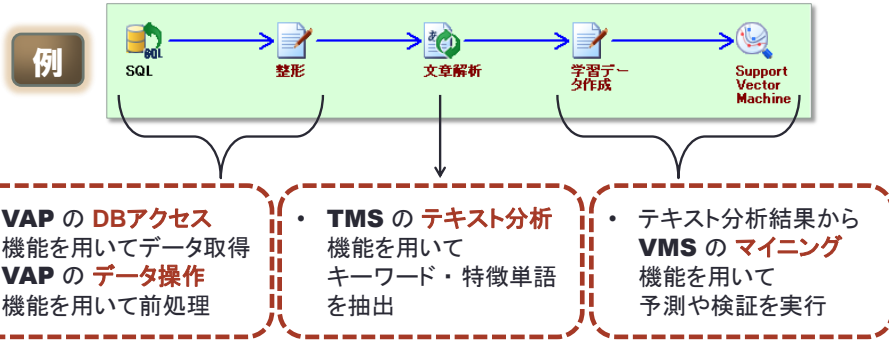
- 原文参照改良
  - ショートカットキーでいつでも原文参照可能に
- 日本語解析エンジン 3.1 と 3.2 の動的切替え

## プラットフォームVAP と VMS, TMS



## データ／テキストマイニング連携

- **VAP**上で **TMS, VMS** の機能を連携させて利用することが可能

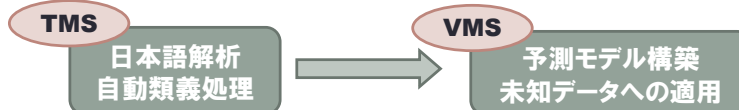


オリジナル・カスタムなテキスト分析フローを構築

## VMS, TMS連携による分析ストーリー

### 適用例

- テキストを用いた判別・予測問題
  - 問合せ内容からの 重要度ランク付け や ジャンル振分け、
  - ネット上の書き込み内容からの ポジネガ分類 や プロフィール推定



- 数値情報 と テキスト情報 との融合
  - 自由回答部分 と 選択回答部分 が存在するアンケート分析





## リリーススケジュール

**Visual Mining Studio 新バージョン V7.3**

**Text Mining Studio 新バージョン V4.2**

**2013年3月 リリース**

ご期待ください。

## 後継バージョン

### Visual Mining Studio

- ・データ可視化機能の強化
  - ・データマイニング + **VAP** BIツール との連携
- ・R連携機能の強化
- ・新しいマイニング機能の導入

### Text Mining Studio

- ・**VAP**, **VMS** との連携機能の強化
  - ・**VAP**上でより利用しやすい形に **TMS**機能を再構築
- ・特定の適用分野に特化した、お客様にとってより即戦力となるテキスト分析ソリューションの開発