

押し寄せる膨大な「今」を瞬時に賢く分析する

～フロー型ビッグデータを分散オンライン機械学習で分析～

どんな研究

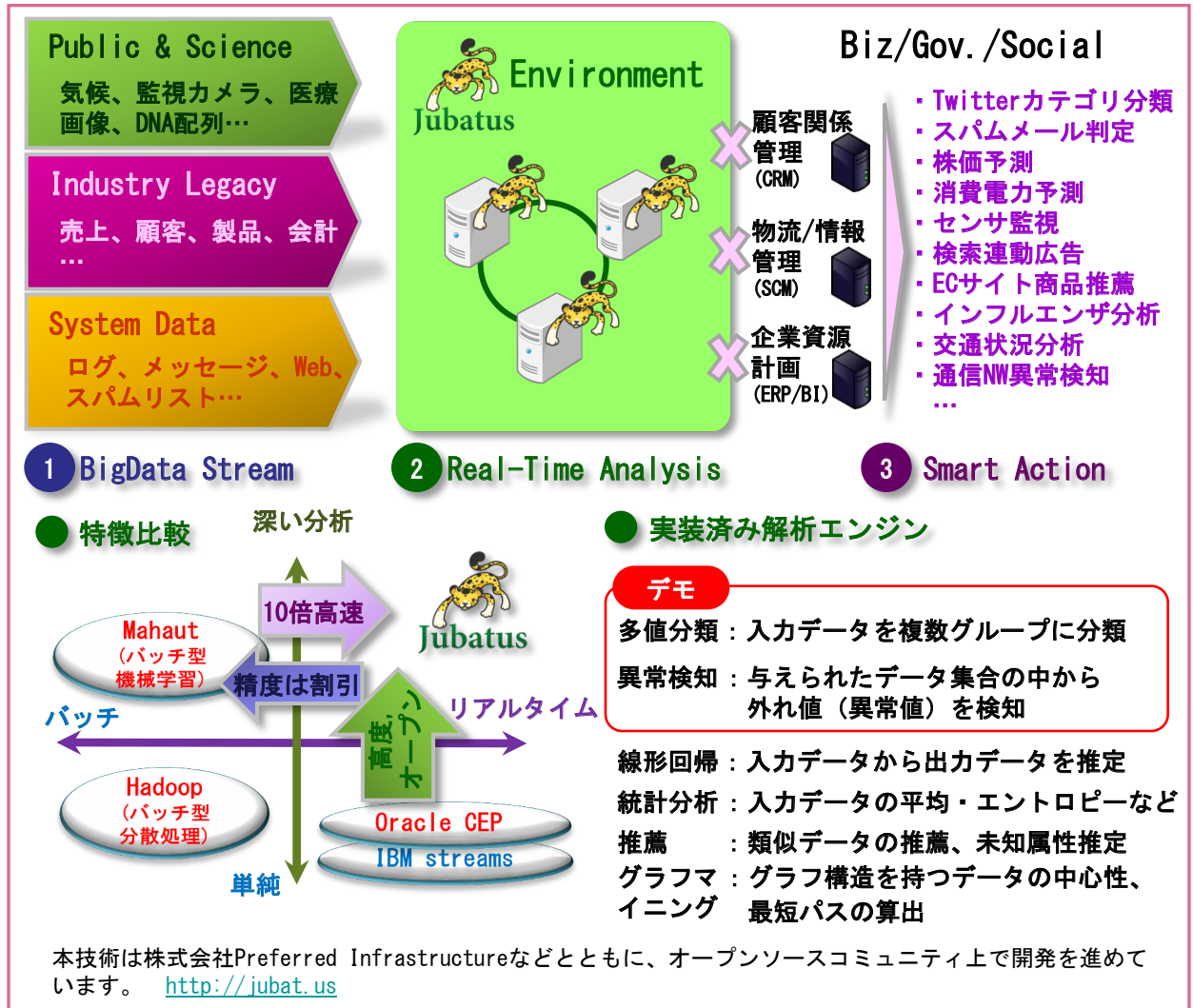
ビッグデータ分析の新たなニーズである、(1)広範囲に散在する多様な情報源のデータを統合し、(2)素早いリアルタイム分析でリスクを最小限に抑えて、(3)事前に全貌を把握しきれない多様な要求にも柔軟に対応できる処理基盤技術を開発しました。

どこが凄い

3つの要件「大量」「速い」「深い」処理はトレードオフの関係です。すべて両立することは本質的に困難ですが、これを克服し、毎秒16メガバイトのデータ流量に対し、自然言語解析と複数カテゴリへの自動分類を、遅滞なく高速実行します。またOSSとして広く公開しています。

めざす未来

世界規模でのオープンイノベーションとマーケット創造を推進し、リアルタイムマーケティングやスマートな社会インフラマネージメントへの適用拡大を図ります。



関連文献

- [1] 岡野原大輔, 海野裕也, 熊崎宏樹, 小田哲, “大規模リアルタイム解析エンジンJubatusの創り方,” 情報処理学会デジタルプラクティス, Vol. 4, No. 1, pp. 20-28, 2013.
- [2] 小田哲, 中山心太, 上西康太, 木下真吾, “Jubatus: Big Dataのリアルタイム処理を可能にする分散処理技術,” 信学技報, Vol. 111, No. 409, IN2011-126, pp. 35-40, 2012.
- [3] H. Makino, “Jubatus: Scalable Distributed Processing Framework for Realtime Analysis of Big Data,” in Proc. XLDB2012, 2012.

連絡先

堀川 桂太郎 (Keitaro Horikawa) NTT ソフトウェアイノベーションセンタ 分散処理基盤技術プロジェクト
E-mail: horikawa.keitaro[at]lab.ntt.co.jp ({at}の部分>@に置き換えてください)