

# 情報空間があまりだす社会構造の変化

山本 強

北海道大学大学院情報科学研究科

yamamoto@ist.hokudai.ac.jp

(C) T. Yamamoto, Hokkaido Univ.

1

# Agenda of the day

- ICTの技術動向から考える情報空間の構造変化
- 網羅的情報空間のインパクト
  - 網羅的情報空間から見える社会と事件
- Intelligence と Information
  - 情報の価値構造の変化
  - 情報労働の産業革命

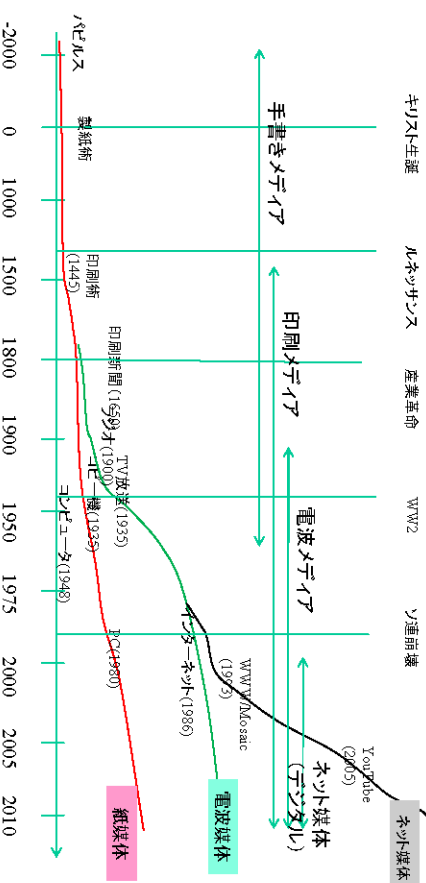
# メディアの変革と情報流量の増加

キリスト生誕

ルネッサンス 産業革命

WW2

ソ連崩壊



(C) T. Yamamoto, Hokkaido Univ.

3

# メディアのデジタル化

- 伝送・蓄積メディアと表現メディアの分離
  - 高速デジタルネットワークがすべての表現メディアを伝送可能(伝送メディアの透明化)
- 劣化しない複製(デジタルコピー)
  - 原本とコピーが同価値
- コンピュータによる機械的メディア処理が可能
  - データベース化、検索、合成が可能

## デジタルメディアの時代

- デジタルメディアはアナログメディアの置き換え、高機能化と単純化しては行けない
  - 伝送・蓄積のモデルがアナログメディアとは根本的に違う
- 意図した操作(改ざん、修正)を見抜けない場合がある
  - メディアの機能、目的、リスクが根本的に違う

## 新しいデジタル価値の創造 Googleの発明

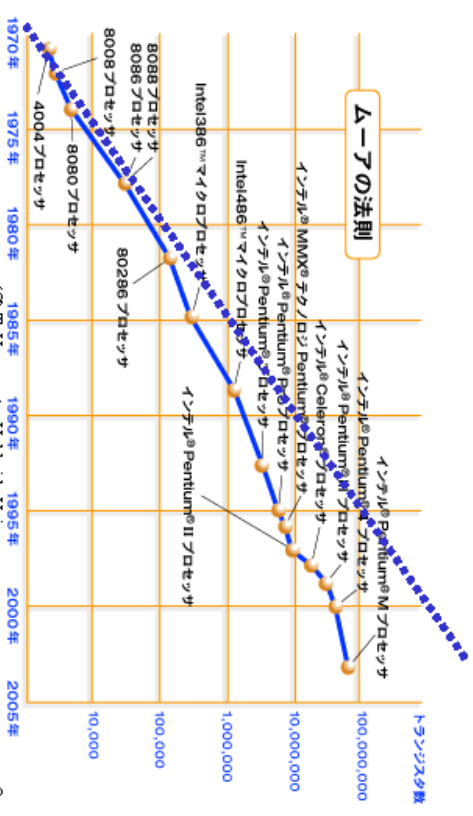
- ホームページの参照関係から言外の意味を抽出する技術
  - リンクカウント(Link Count)
    - 重要な情報はより多く他から参照される
    - 論文の参考文献引用と同じ理屈
  - ホームページの重要度を定量化することが可能
    - 重要度によるホームページの広告価値算定
    - 検索キーワードを不動産化
- インターネット上の全情報を獲得し、再編成することで、新しい情報価値を創造した

## 情報を量的に把握する

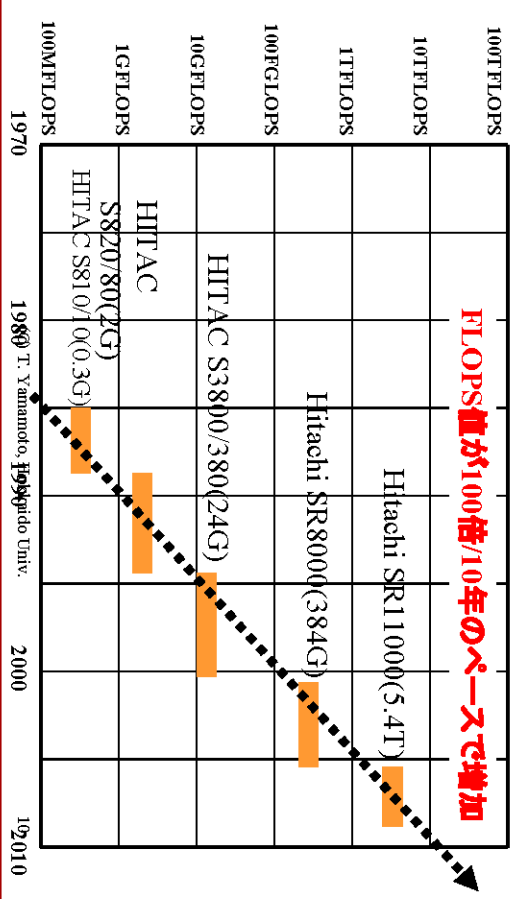
- 日本全土の1m×1mの衛星写真の容量は
  - 国土面積: 377,899Km<sup>2</sup>
  - 1Km<sup>2</sup>当たり1Mpixelとして377,899 M Pixel (0.38Tpixel)
  - 1Pixel 3byteとして0.38T × 3 = **1.14TByte**
  - JPEG圧縮して55GB程度(このノートPCに格納可能)**
- 人間が一生に発声する音声データ全て
  - 活動時間 1日16時間として 3600\*16\*365\*80=1.68G秒
  - 9.6Kbpsで音声を全て記録したとして 1.68Gs\*9.6Kbps = 16.1Tbit = **2TByte (無口な人はこの100分の1以下)**
  - 2TBのフルサーバーは約100万円程度で買える**

## 情報流量を加速するIT基盤技術

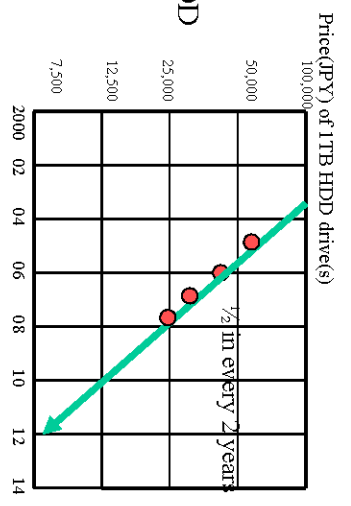
ムーアの法則 LSIの集積度は1.5年で2倍のペースで増加する



### 北大情報基盤センターのスーパーコンピュータシステム



### 1TB相当のHDDの市場価格の動向

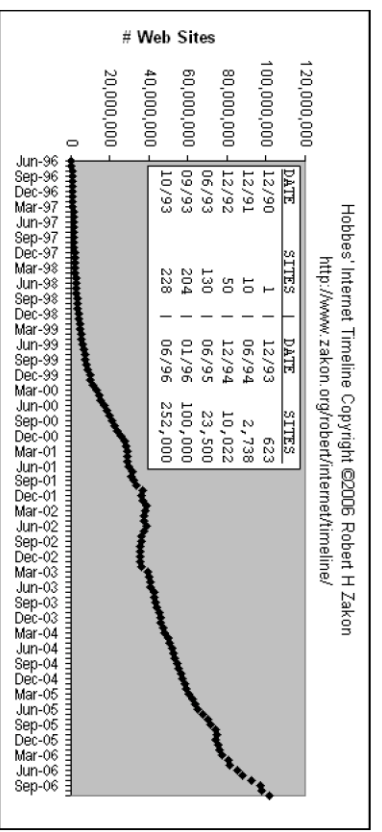


- Expecting price down of HDD drive unit
  - As of 2005, 1TB@60000JPY
  - As of 2006, 1TB@45000JPY
  - As of 2007, 1TB@30000JPY
  - As of 2008, 1TB@25000JPY
  - ...
  - As of 2010, 1TB@10000JPY

Source: [http://pc.watch.impress.co.jp/docs/2007/01/09/pa\\_cphdd\\_hdd.htm](http://pc.watch.impress.co.jp/docs/2007/01/09/pa_cphdd_hdd.htm)

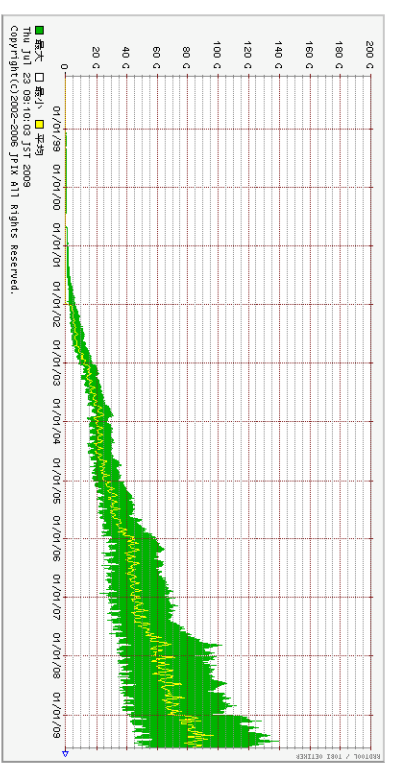
(C) T. Yamamoto, Hokkaido Univ.

### インターネット空間の情報爆発



Source: <http://www.zakon.org/robert/internet/timeline/>

### インターネット情報流量の爆発



Source: <http://www.jpix.ad.jp/graphs/TOTAL.In.minmax256.gif>

## デジタル情報通信環境の歴史と予測

- 同じコストで
  - 10年前の100倍の計算速度、探索速度が買える
  - 10年前の100倍の記憶容量が買える
  - 5年前の100倍のネットワーク速度が買える
- 5年後には同じコストで
  - 今の10倍の計算速度、探索速度が買える
  - 今の10倍の記憶容量が買える
  - 今の10倍のDLAN速度、QoSが保証されたインターネットと高速移動体通信が買える

(C) T. Yamamoto, Hokkaido Univ.

13

## 網羅的情報の時代

- 成功している網羅的情報サービス
    - サーチエンジン
      - Google, Yahoo
    - 巨大データベース
      - 新聞記事データベース、POS端末売上集計データベース
    - 地図データベース(GIS)
      - Google Map, Google Earth
    - 画像、映像データベースサービス
      - Flickr, YouTube
- 人間が全貌を理解できない巨大データベースだが、この中に人間の行動や社会動向が埋め込まれている

(C) T. Yamamoto, Hokkaido Univ.

14

## 網羅された情報空間を探索

- 情報は網羅されているから価値がある
  - 情報システムの処理能力はムーアの法的拡大傾向にある
  - 情報処理に関する常識は3年程度で見直される
- インターネットの情報空間拡大で巨大データ空間を個人がアクセス可能な時代
- 網羅的で大規模なデータベースをコンテンツに転化させる仕掛けの時代

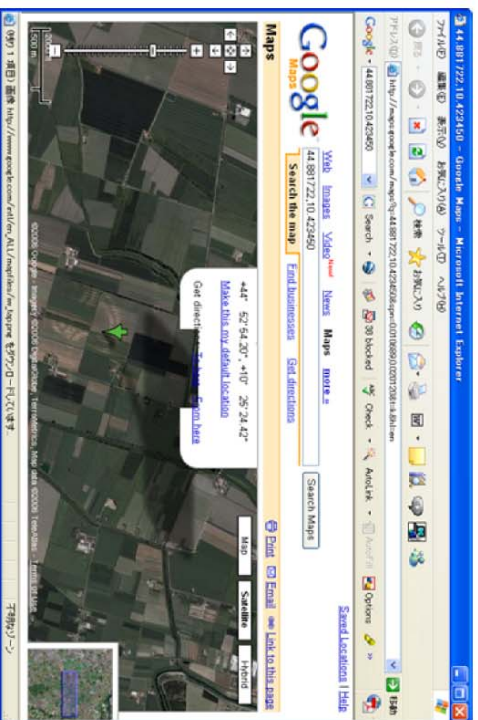
(C) T. Yamamoto, Hokkaido Univ.

15

## 網羅された情報空間からの発見

- 意図して作られた情報からは、意図された情報しか出てこない
  - Yahoo型のSearch Engine
- 網羅された情報は意図されていない情報を含んでいる
  - Altavista, Google型の巡回収集型Search Engine
- 網羅されたデータを後から調べると発見がある
  - Google Earthからの発見

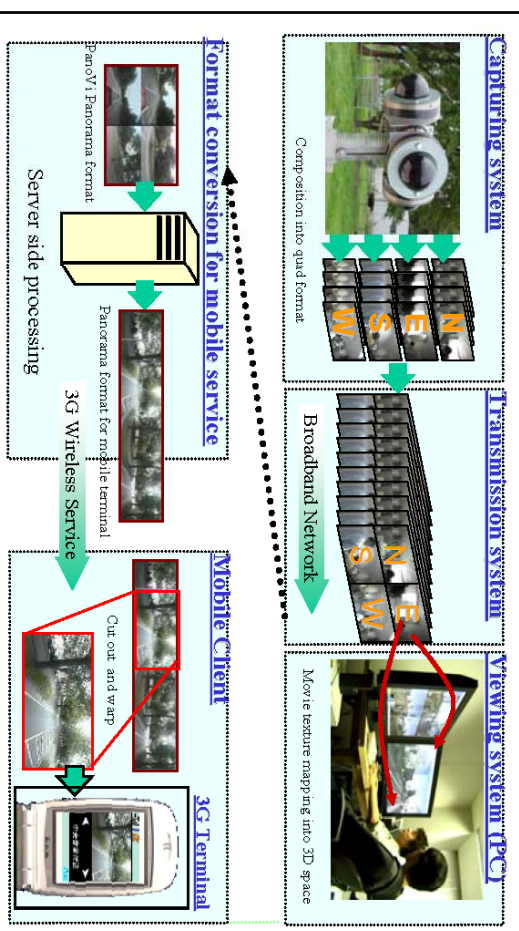
## Google Earthで発見された古代ローマ遺跡



## PanoViで記録された倒壊前の北大ポプラ並木



## PanoVi 環境映像を網羅的に収録・再生



## Google Street Viewのデータ規模

- 前提
  - 札幌市の全道路路上を10m間隔でサンプリングし、その全周画像をデータベース化する
  - 仮定
    - 20km × 20kmの地域に100m間隔で道路がメッシュ状にあると仮定すると、総道路延長は200 × 2 × 20km = 8000km
    - サンプリング点は8,000,000/10 = 80万点
    - 1点あたり1Mバイトの画像と仮定すれば800GBで札幌市全体の全周画像は記録可能
- 結論
  - 今なら10万円のTB級のPC用サーバー1台で札幌市がカバーできる
- Googleは数年後のストリートビュー環境などを正確に予想して、数年前から技術開発とコンテンツ整備の先行投資をしていたと考えるのが自然

## 個人消費動向を反映するPOSデータベース

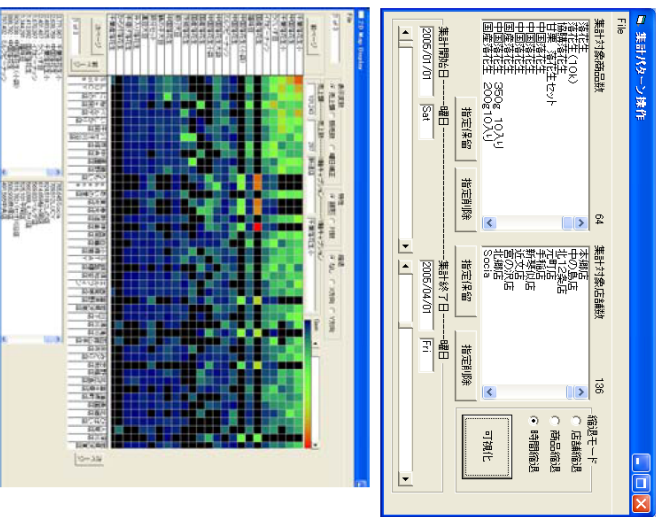
- 1千億円規模の売上げがある大規模スーパーは全てPOSレジスタで売上げ集計されており、一日の売上げ情報は自動的にデータベースに集積されている
- 売上げ300億円で1GBくらいの容量になる
- 単純な数値データの羅列で、それを見ても何がなんだかわからない

(C) T. Yamamoto, Hokkaido Univ.

21

## 二次元マップの表示例 (店舗-商品マップ)

1. 時間縮退により、Y軸-商品、X軸-店舗のマップが作成される
2. 落花生の売上げを2005/1/1-2005/4/1の期間について時間縮退を指定
3. 店舗によって取り扱い商品のパターンが違ってくる



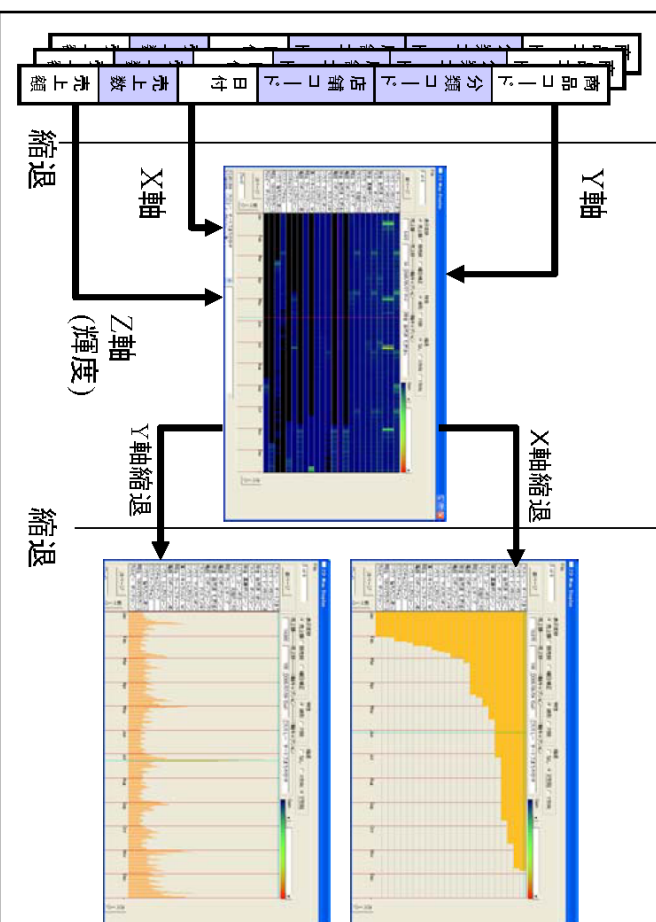
## POS情報ブラウザ

- 縮退(Degeneration)による可視化
  - 売上データの関数空間
  - 売上額関数:  $\text{SaleAmount} = f_1(\text{Shop}, \text{Item}, \text{Time})$
  - 売上数関数:  $\text{SaleCount} = f_2(\text{Shop}, \text{Item}, \text{Time})$
- 変数軸
  - Shop: 店舗軸
  - Item: 商品軸(JANコードは商品コードから派生)
  - Time: 時間軸
- 売上データの変数軸を積分(縮退)することで低次元関数に変換し、可視化する

N次元DB

2Dマップ表示

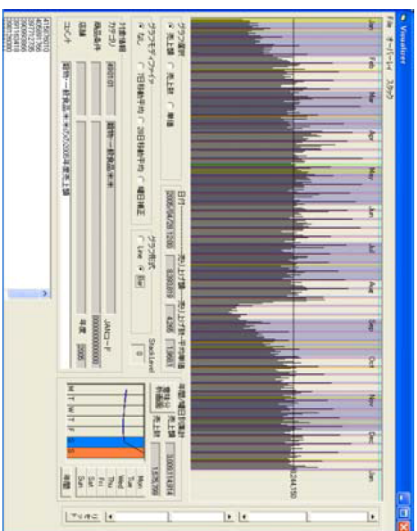
1Dマップ表示



## データが語る地域文化

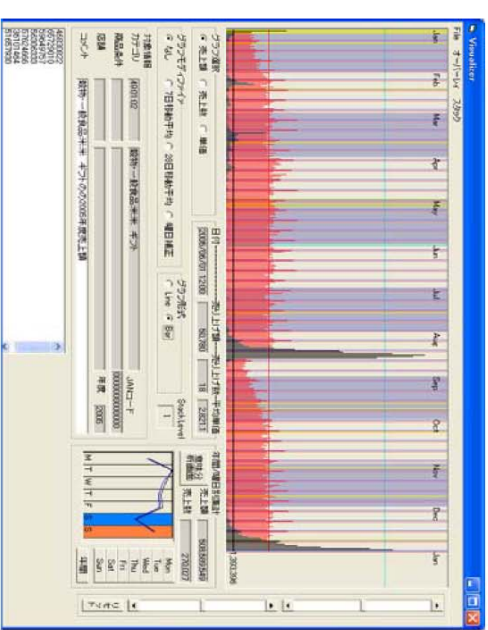
- ・ 沖縄県の「一般米」の売れ方

1月と8月に急激な売り上げ低下がある



## 沖縄の家庭にストックされる米ギフト

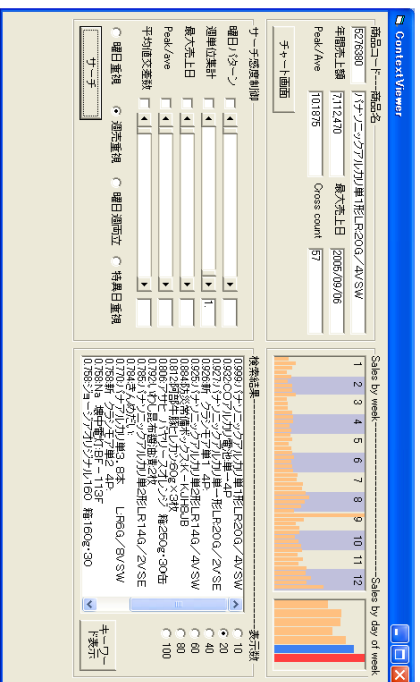
- ・ 沖縄の「米ギフト」習慣が、盆暮れに米の家庭備蓄量を急激に増大させる。それが消費されるまでの売りは一般米の売りに上げがあらない



## 単1電池に見える特異パターン

パナニツクアルカリ電池単一を検索キーとして週単位売り上げ重視で意味検索

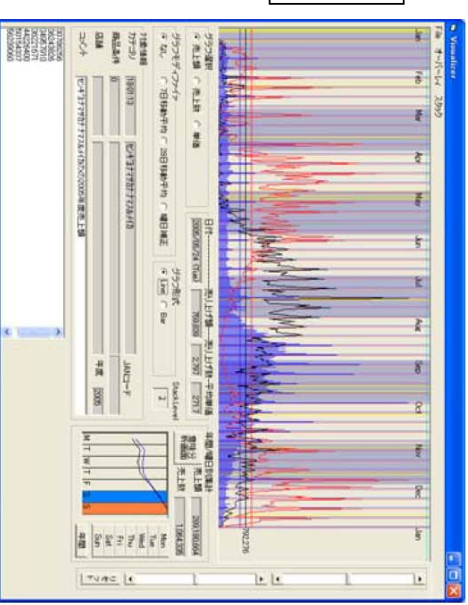
類似パターンとして上位に検索されるのは同様の大型電池と防災常備ボックスであり、これが防災関連グッズとして売れていることがわかる



## 北海道のスーパーでの鮮魚売りにげに見える季節感

青: 秋刀魚分類  
赤: まがれい分類  
黒: 生するめイカ分類

早春、晩春のカレイ  
初夏の生イカ  
初秋の秋刀魚



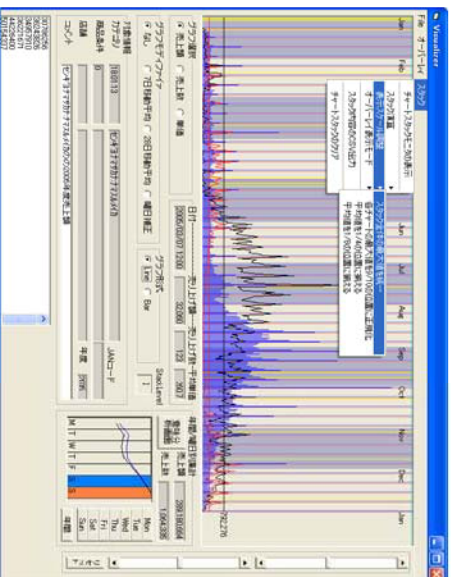
## 金額ベースの比較

青: 秋刀魚分類  
 赤: まがれい分類  
 黒: 生するめイカ分類

グラフ間でスケール統一することにより、金額、個数ベースでの絶対値比較が可能

金額ベースで見ると、秋刀魚、生イカの売上額はカシイよりもだいぶ大きいことがわかる

北海道人は生イカが好き



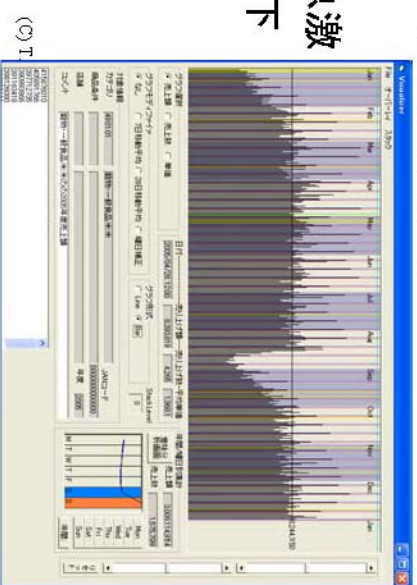
(C) T. Yamamoto, Hokkaido Univ.

29

## データが語る地域文化

・ 沖縄県の「一般米」の売れ方

1月と8月に急激な売上げ低下がある

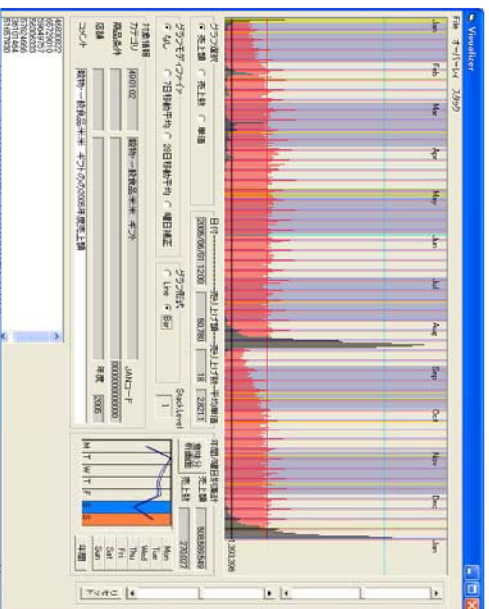


(C) T.

## 沖縄の家庭にストックされる米ギフト

・ 沖縄の「米ギフト」習慣が、盆暮れに米の家庭備蓄量を急激に増大させる。

・ それらが消費されるまでは一般米の売り上げがあがらない



## 新聞全文データベースによる単語出現頻度分析で見えるメディアの嗜好

各紙が1985-2007年の間に天地真理、小柳ルミ子、南沙織という文字を含む記事を出した件数(日経テレコン21による検索)

	天地真理	小柳ルミ子	南沙織	南沙織/天地真理
朝日新聞	28件	92件	49件	1.75
読売新聞	35件	108件	29件	0.83
毎日新聞	27件	67件	39件	1.44
産経新聞	50件	103件	40件	0.8
道新	19件	40件	19件	1.0
日経	7件	23件	13件	1.85

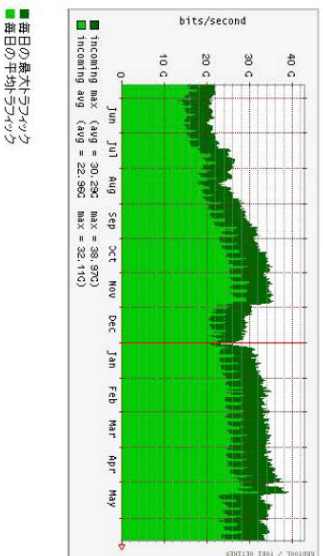
(C) T. Yamamoto, Hokkaido Univ.

32



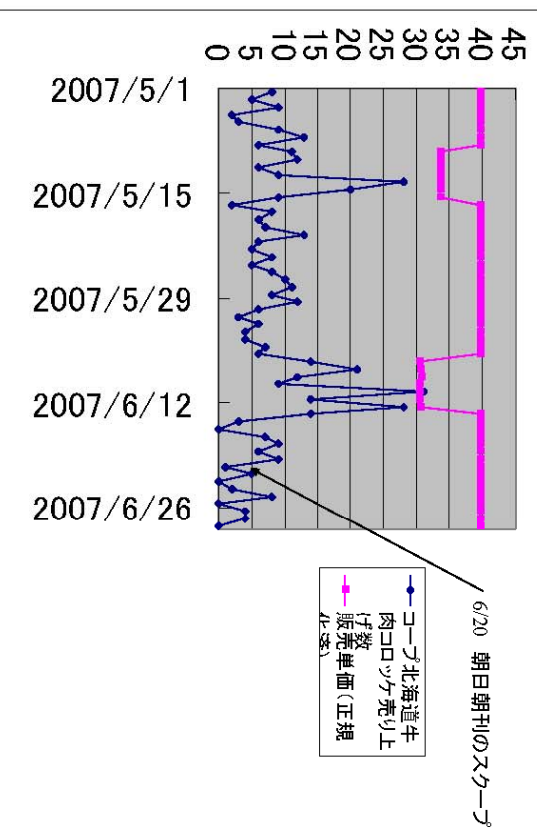
## 事件と情報空間 2004年11月にインターネットトラフィックが急激な減少を示す

トラフィック  
JPNAP total traffic

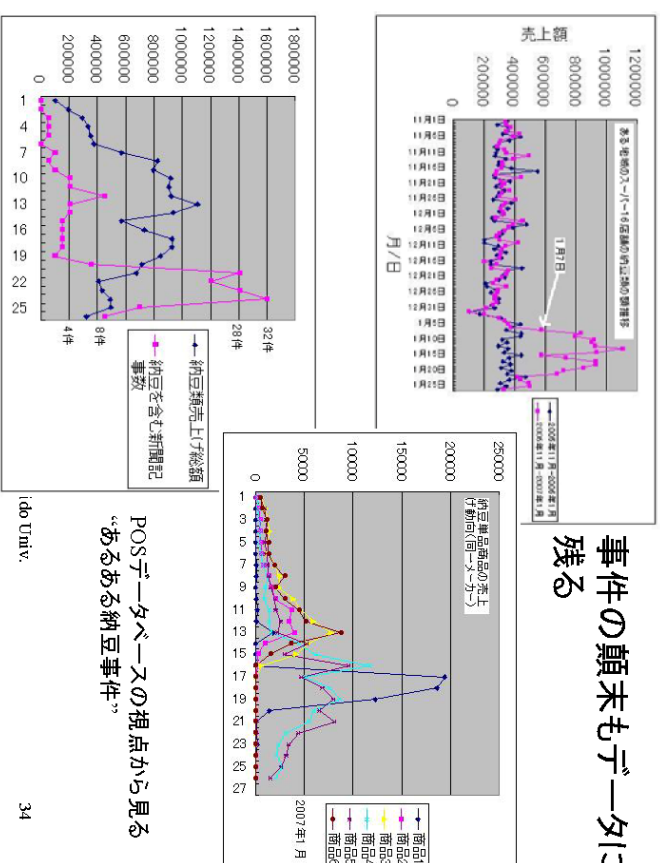


2004年11月27日  
にフアイル交換ソ  
フトウェア Winny  
の不正利用の摘  
発および開発者の  
家宅捜査があった

この現実をどう見るか



## 事件の顛末もデータに 残る



## IntelligenceとInformation

- Intelligence
  - 情報の価値はその意味、解釈にある
  - 意味は人間だけが理解できる
- Information
  - 情報の価値はその量(bit)にある
  - 機械(コンピュータ)が処理(分析)できる

爆発する情報処理能力、情報伝達能力が量的情  
報の価値を高めている

## IntelligenceとInformation

- かつてメディアはIntelligenceを伝えた
  - メディアが運べる情報量(bit)に限界があった
  - メディアは濃縮した情報(Intelligence)を運ぶ専用線
  - コンテンツによって物理メディアが異なるのは当然
- いまメディアはInformationを伝える
  - メディアが運べる情報量には限界がない
  - メディアは量的な情報(Information)を運ぶ透明なパイプ
  - メディアとコンテンツは独立

(C) T. Yamamoto, Hokkaido Univ.

37

## 20世紀末のデジタル革命は頭脳労働の産業革命なのではないか

- 単純知的労働の機械化
  - データベース化、検索システム
  - 輸送システム(ネットワーク)
- 資本家の誕生
  - デジタル資産を蓄えた新しい資産階級の出現
    - Microsoft, Amazon, Google
- 社会構造の変革
  - 情報流量の増大
  - 双方向情報流の発生(生産者と消費者の対等化)

## デジタル化で考古学や社会学の方 法論に新手法が加わる

- 遺物の保存モデル
  - アナログ: 保存状態の良い原本が残る
  - デジタル: 大量のコピーが残る
- 発見の技術
  - アナログ: 人間による洞察、物理的な発掘
  - デジタル: 仮説+コンピュータによる検索
- 情報の未来への伝達
  - アナログ: 図書館、博物館など記録環境整備
  - デジタル: 情報空間(インターネット)へのコピー配布

## おわりに

- 情報通信網のデジタル化が誰でも巨大で網羅的な情報空間へのアクセスを可能にした
- 大量デジタルコンテンツの解析から作成者が意図しないメッセージが見つかることがある
- デジタル情報空間を利用する新しい文系研究の手法が加わるかも知れない

(C) T. Yamamoto, Hokkaido Univ.

40