AITC
先端ＩＴ活用推進コンソーシアム

# ～第１回中間活動報告会～
# クラウド･テクノロジー分野の活動状況について

2011年02月23日

アドソル日進株式会社
荒本道隆

# 背景/活動目的

- 多種多様なクラウド関連技術がオープンソースや製品として出てきているが、「クラウド」という言葉の定義自体が幅広いため、企業内システムでの適用できる領域・効果が分かりづらい。
- そこで、クラウド関連技術の幅広い情報収集をおこなって、調査結果を情報共有するとともに、Hadoop やNoSQLなどの中で利用可能なものを実際に試用し、それらを使ってプロトタイプを試作してみることで、クラウドを企業内システムで活用するヒントや具体的なイメージを得るための活動を行う。

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

AITC
先端ＩＴ活用推進コンソーシアム

# 活動実績の報告

# 2010年10月21日 14:00～18:00

- 第1回 セミナー＆勉強会
- 参加者：14人
- タイトル
  - 情報交換会
  - 『Hadoopの概要と利用動向』
    - Tom White 著「Hadoop」（O‘reilly社刊）翻訳者 玉川 竜司 様
- 内容
  - クラウドに関する様々な情報交換
  - 今後の活動方針：「技術と、テーマ（具体的な実装モデル）の両方からやっていく」

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# 2010年11月22日 14:00～18:00

- 第2回 セミナー＆勉強会
- 参加者：18人
- タイトル
  - 医療保健情報の紹介と課題整理
    - 愛媛大学　大学院医学系研究科　　木村　映善　様
  - 気象庁GPVデータの紹介と課題整理
    - 気象庁　予報部　　浜田　浩　様
- 内容
  - MML（Medical Markup Language）とGPV（Grid Point Value）の概要と、現状の利用方法の紹介

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

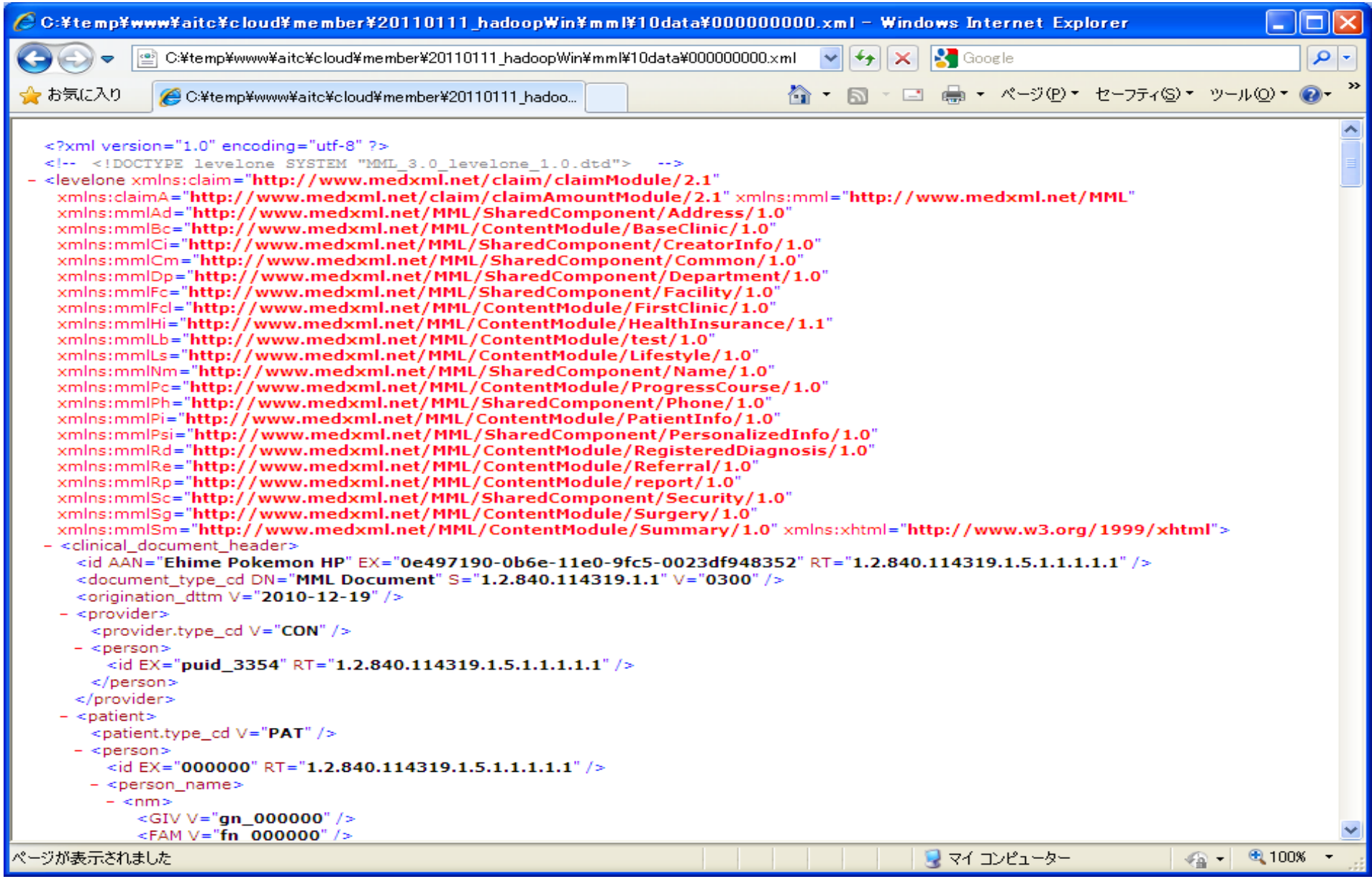
# 2010年12月20日 14:00～18:00

AITC 先端ＩＴ活用推進コンソーシアム

- 第3回 勉強会
- 参加者：１０人
- タイトル
  - MMLとGPVをHadoopでどう使ってみるか？
- 内容
  - 各データの特性と、Hadoopとの相性
    - MML
      - １回診察を受けるごとに1ファイルが生成される、XMLデータ（DTDあり）
      - 蓄積し続ける
      - 病院ごとに使われ方が統一されていない。個人を識別できるIDがない
      - 分析・統計をする場合、基本的に全件スキャン
      - 利用例：Hadoopを使って、興味のある要素だけを抽出する
    - GPV
      - 地球全体、20km間隔の数値データ、バイナリ（MSMは5km間隔で日本付近のみ）
      - 6時間ごとに、全件入れ替え（MSMは3時間ごと）
      - 緯度・経度を元に、特定領域や要素のデータだけ参照したい
      - 利用例：数値データを元に可視化

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# MMLのサンプル



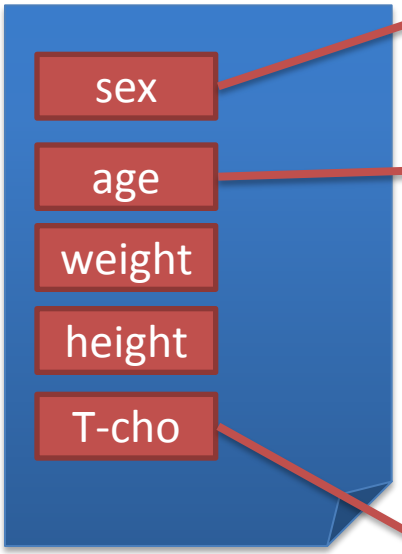

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# MMLの利用例－１

- 例題
  - 全国民のBMIとコレステロールの関係を性別、年齢別に調べて、回帰直線を求めて下さい。
    - BMI(body mass index) 肥満度のindex
    - 体重(kg)/身長(m)^2
  - 性別、年齢別に、BMIとコレステロールを集計しておく。
    - 回帰直線を描くのは統計処理ソフトウェアの仕事

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# MMLの利用例－2




/levelone/body[1]/section[1]/paragraph[2]/content[1]/local_markup[1]/mmlPi:PatientModule[1]/mmlPi:sex[1]

/levelone/body[1]/section[1]/paragraph[2]/content[1]/local_markup[1]/mmlPi:PatientModule[1]/mmlPi:birthday[1]

/levelone/body[1]/section[1]/paragraph[2]/content[1]/local_markup[1]/mmlLb:TestModule[1]/mmlLb:laboTest[1]/mmlLb:item[50]/mmlLb:itemName[1]

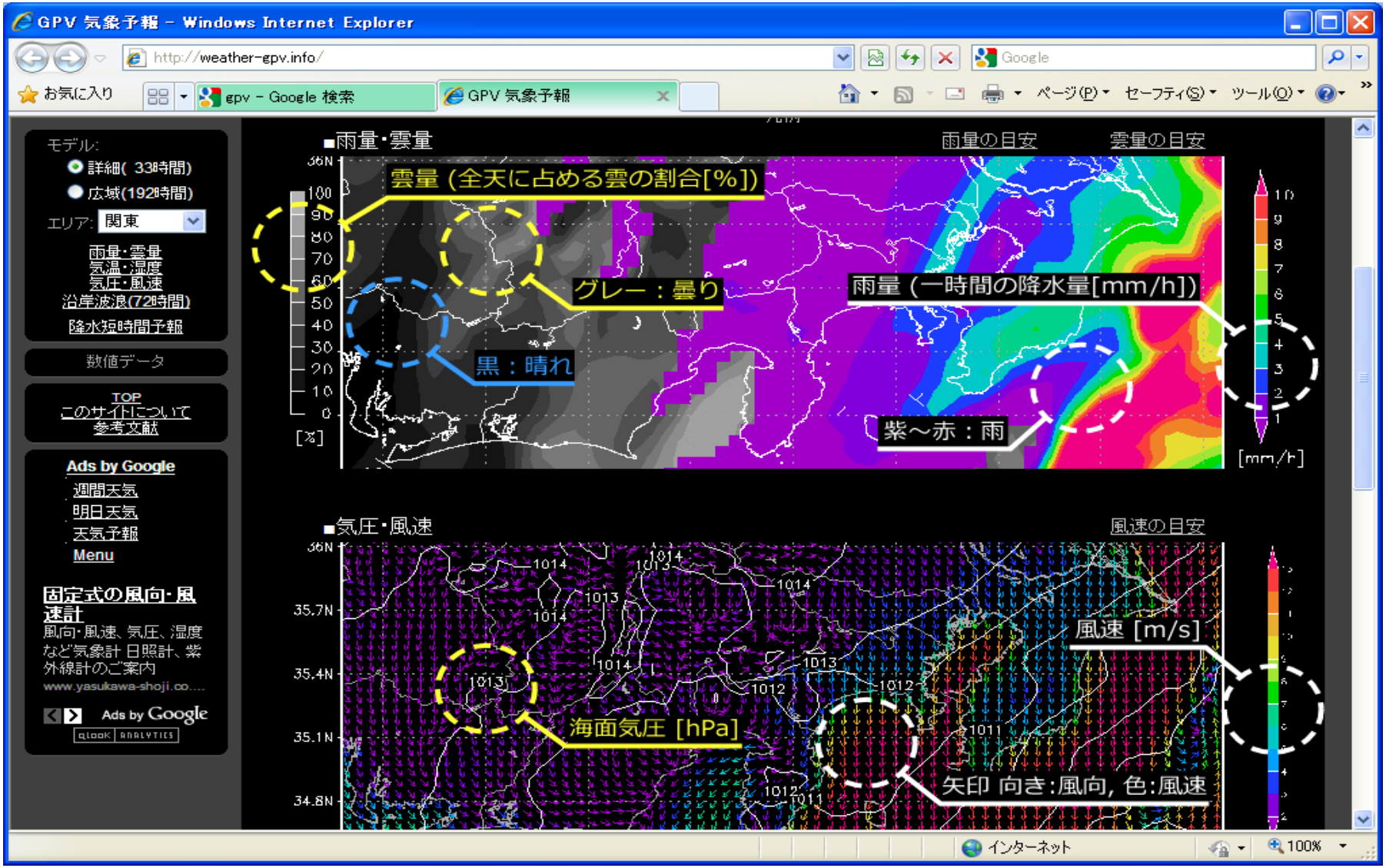
/levelone/body[1]/section[1]/paragraph[2]/content[1]/local_markup[1]/mmlLb:TestModule[1]/mmlLb:laboTest[1]/mmlLb:item[50]/mmlLb:value[1]

```
<mmlLb:item>
    <mmlLb:itemName mmlLb:itCode="610" mmlLb:itCodeId="KRTKC"
        >T-Cho</mmlLb:itemName>
    <mmlLb:value>112</mmlLb:value>
    <mmlLb:numValue mmlLb:out="L">112</mmlLb:numValue>
</mmlLb:item>
```

注意：上記のXpath上のインデックスはサンプル文書上の位置であり、最初から決まっているわけではない。

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# GPVの利用例




http://weather-gpv.info/

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# 2011年01月21日 14:00～18:30

AITC 先端ＩＴ活用推進コンソーシアム

- 第4回 勉強会
- 参加者：17人
- タイトル
  - Hadoopを使って医療データ（MML）の集計を試行
- 内容
  - ハンズオン形式
  - MML（XML）から必要な要素だけを抽出
  - 実際にMapとReduceのプログラムを書いて、動作させる
  - 年齢と性別ごとに、ファイルへ出力

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# 2011年02月18日 16:00～18:00

- 第5回 勉強会
- 参加者：11人
- タイトル
  - 今後の活動内容検討会
  - MMLとGPVをどう活用するか？
- 内容

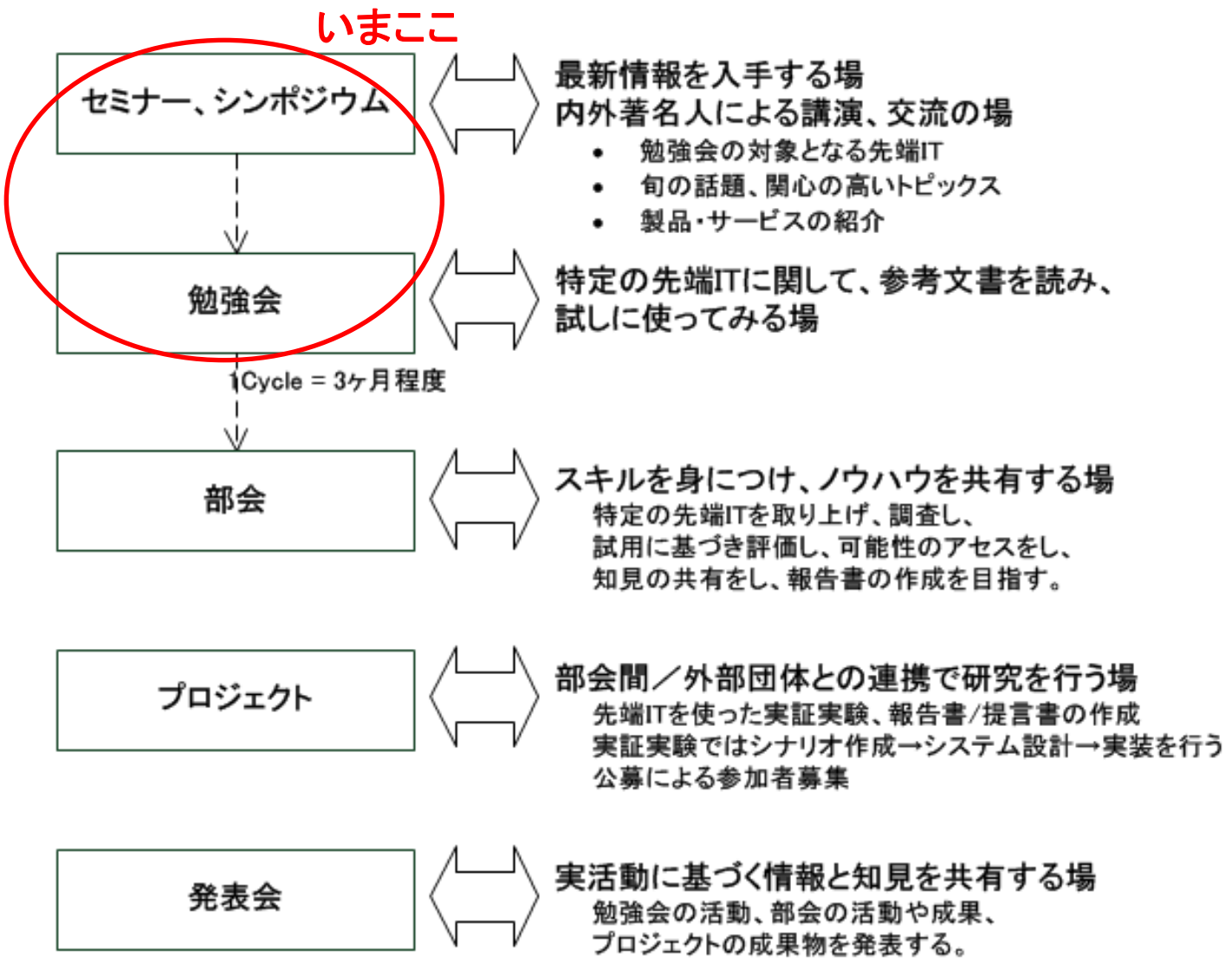
  今後の活動計画について
  - 部会化するのか？
  - 勉強会はどうする？
  - 部会化した場合の、具体的な内容

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

AITC
先端ＩＴ活用推進コンソーシアム

# 今後の活動予定

# 部会化するのか？




『先端IT活用推進コンソーシアム設立総会議案書＜活動の種類＞』より

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# 部会化するメリット

- 目標を明確化
  - 最終成果物（ポートフォリオ？ガイドライン？報告書？等）
  - 作業のゴール＆マイルストーン
- 参加者の半固定化
  - もちろん、途中からの参加も大歓迎
- 成果物の承認と公開
  - 理事会承認をもって成果物を公開
  - 成果物には、参加会員企業名を掲載
  - 公開に伴いプレスリリースを実施

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# 継続して勉強会も開催予定

- 部会での活動成果を共有・伝播する場
  - より詳細な内容を共有
  - 実証実験の一部をハンズオンで実体験

- 第4回 勉強会ライクなハンズオンを継続開催？
  - 「Hadoopを使って医療データ（MML）の集計を試行」
  - 同じ内容で開催（？）
  - プラス、完全分散モードの実体験（？）

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

AITC
先端ＩＴ活用推進コンソーシアム

# 部会化した場合の活動内容

# MMLを使った実証実験

- 大量の、小さなXMLデータに対する集計処理
  - Hadoopと、他実装との比較
    - RDB, XMLDB, その他のクラウド技術
    - 実装のし易さ、性能を上げるための注意点、など
    - 今までの設計方法との違い
  - 性能
    - スケールアウトした時の、性能向上の傾向
    - Amazon Elastic MapReduce の試用
    - 性能をリニアに上げるための注意点、など
  - 名寄せ
    - 集計に必要な前処理もHadoop内で行う

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# GPVを使った実証実験

- 大量データを人が理解し易いよう可視化
  - 大量のバイナリデータを格納
    - 緯度・経度・高度という位置情報を持ったバイナリデータ
    - 3時間ごとに全件入れ替え
    - HDFS, HBase, その他
  - 可視化とUI → 他勉強会との連携
    - スマートフォン
    - リッチクライアント
    - AR での活用

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# クラウド関連技術調査

- Hadoop以外のクラウド関連技術を調査
  - 有効そうであれば、実証実験で使ってみる

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

# ガイドラインの調査

- クラウドが、各ガイドラインに適合するかを調査
  - 「医療情報システム関連ガイドライン」
    - 医療情報システムの安全管理に関するガイドライン4.1版 (厚生労働省)
    - 医療情報を受託管理する情報処理事業者向けガイドライン
    - ASP・SaaS 事業者が医療情報を取り扱う際の安全管理に関するガイドライン
    - 今年から医療情報システム安全管理評価制度(PREMISs)
  - 医療情報と金融情報が、一番厳しいはず
  - 今後、クラウド技術を実業務で使っていく上では重要

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.

AITC
先端ＩＴ活用推進コンソーシアム

# 次回の活動予定

# 2011年03月18日 14:00～18:00

AITC 先端ＩＴ活用推進コンソーシアム

- 第６回勉強会（仮）
- タイトル

  部会化に伴う成果目標・日程を検討・決定

  – 各実証実験／調査の実施内容と役割分担を決定

**皆様の参加、お待ちしております。**

Copyright © 2010 Advanced IT Consortium to Evaluate, Apply and Drive All Rights Reserved.