# ソーシャルゲーム× データ解析

KLab勉強会 2012-06 資料

## 自己紹介

高田敦史

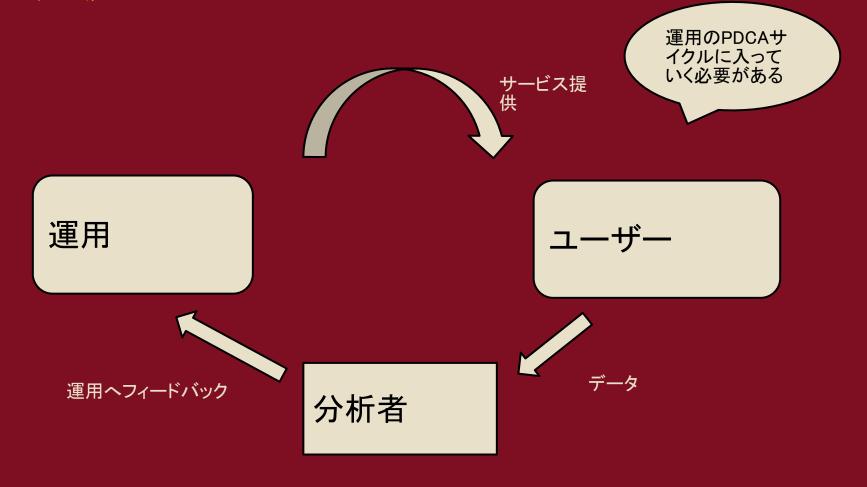
開発部アーキテクトG+Kラボラトリー所属です ソーシャルゲーム開発 PHP、Ruby、Python、JavaScript 最近はデータマイニングに興味があります

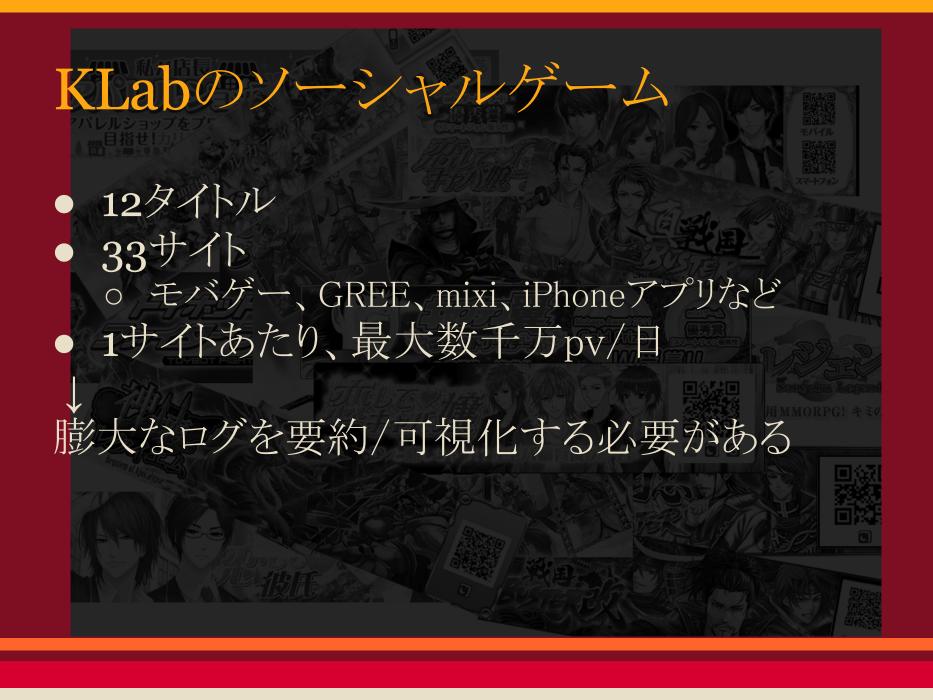
## 最初に

何のためにデータの収集・解析が必要なのか

- ゲームの運用改善、サービス向上に役立てる ため
- 運用・企画のフローの一部に組み込まなければならない
- 技術だけで完結する話でもない

# 最初に

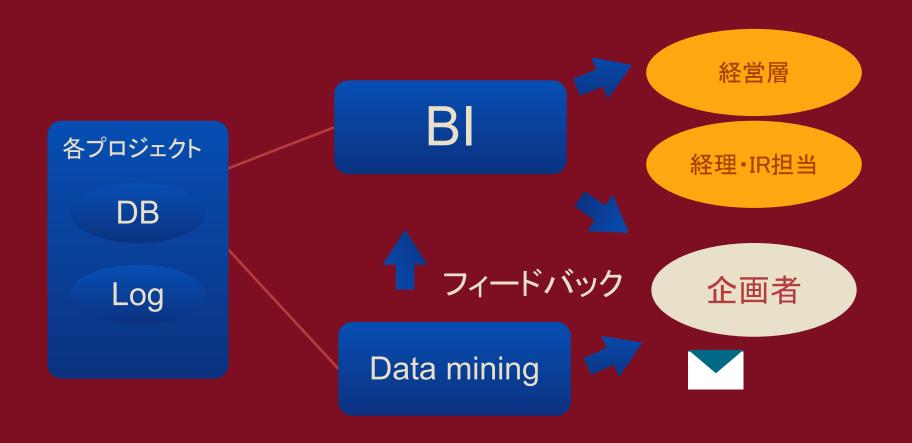




#### 体制

BI用ツール(kg\_report):
WEBサービス
定常的な売上管理
データマイニングチーム(kg\_kpi):
マーケティング担当者 + 開発
各ゲームの詳細なデータを収集 + 蓄積 + 分析
新しい分析軸の模索

## 体制



#### BIチームのミッション

- kg\_reportの開発・運用
  - o 社内限定のWebサービス
  - o 経理方面の人々が売上の確認に利用する
  - ケームの企画・運用にも利用されているが、メインではない。

BI: ビジネスインテリジェンス

#### 売上管理の側面

- 全プロジェクト網羅的であることが重要。
- 正確であること重要。
- 必要な数字は合計だけで、厚い情報は必要ない。
- 担当者が確認できるツールが必要。

# kg\_report

TOP カレンダー・ ユーザレポート・ 充土レポート・ アンケート・																			
mixi	‡ 各アプリ	合算	<b>‡</b>																
									売	上レポート(継	続ユーザ別	)							
売上レポート » 売上	レポート(継続ユーザ別)																		
集計単位	<ul><li>日毎 ○週毎 ○月</li></ul>	毎																	
期間	2012/05/3	1 ~ 20	12/06/14																
アイテムカテゴ	リ 全体	<b>\$</b>																	
			検索																
2012/05/31 (木)~	~ 2012/06/14 (木)																		
								1日継続フー	·tf 4	3日総	締ューザ		7日継続ユーザ	14日継続ユーザ	_	30日継続ユーザ	•	60日継続ユーザ	

日付 ◆	SNSポイント⇔	インストール数⇔	UU ¢	課金UU⇔	課金率(%)¢	ADDULA	ARPU≄	PV ±	1日継続ユーザ		3日継続ユーザ φ		7日継続ユーザ ⇔		14日継続ユーザ ⇔			30日継続ユーザ 💠			60日継続ユーザ 4		<b>ザ</b> ¢	メモ⇔			
14 Y				除土工びじ章	sx.ш-1°(70)∓	ARFFU	ANTO	PV \$	ARPPU≑	ARPU\$	売上 💠	ARPPU\$	ARPU\$	売上 💠	ARPPU\$	ARPU\$	売上 💠	ARPPU\$	ARPU\$	売上 ⇔	ARPPU¢	ARPU\$	売上 ⇔	ARPPU\$	ARPU\$	売上 ⇔	7 2 4
2012/05/31 (木)	30,000,000	40,769	40,00	44,800	984	795.76	33.66	190,000,700	191.53	40.50	3,643,650	1,206.55	16.00	583,286	1,000,000	19/00	1,075,000	N/Amain	40.75	4,750,000	1,018.00	26,000	1,004,000	1982.00	8186	46,312,681	
2012/06/01 (金)	10.386380	30(40)	4901	88,140	10078	2,640.00	200.60	166,679,386	1,094.0	160,60	1,011,000	3,000,79	284.21	1,000,000	2,004.00	21702	1/10/00	4,00001	387.01	0,680,060	4,07379	101.36	15,000,790	2019034	303.00	110,900,700	
2012/06/02 (土)	86870040	204.0	4000	10,000	7987	3,943,79	137.69	197240389	940.00	90.00	1,010,760	1,586,58	100.09	4,394,504	1,600,000	75500	Letters	2,490,00	140.79	1,041,063	3,481,66	307.3%	1306400	109603	99.27	60.761.004	
2012/06/03 (日)	70940.000	101/000	40.40	4000	9.70	1,540.00	190.01	10041100	961.27	98.0	1,500,000	1,000,00	71.00	4,000,000	1,004,000	98.40	1,400,040	18091	200.01	4004,000	8,000.00	190.00	Section	104000	100.01	91,500,440	
2012/06/04 (月)	36(31),050	20,620	477,640	49,500	19180	187.44	34.86	500,007,000	509.74	43.94	881,680	2013.00	83.86	1,101,101	1,000,70	55.00	96,09	1,000,70	61.84	3,683,660	1,781.79	15.45	48479	900.40	76.84	81,761,000	
2012/06/05 (火)	600000	21,046	40040	37,500	8.71	1,010.07	8746	<b>O1,98508</b>	180.00	80.30	995,500	810.00	63.50	3,104,800	96679	86.00	1,000,000	1,867.00	8071	3,981,980	3,400.60	98.49	1,281,421	1,0000	8133	2079000	
2012/06/06 (水)	30,000,007	193299	40,30	34046	138	1,546.61	94.09	19080010	600.00	4740	794,800	166.17	66.99	1,790,000	1,00046	77.00	4.098880	566633	60.01	3.565,560	3.998.60	96.60	4.798.007	5000004	27.09	23,661,066	
2012/06/07 (木)	10/01/10/9	77000	410,000	0.00	5460	1,000,00	20.00	111040000	941.00	41.00	504,834	980.00	46.00	1,011,003	5,000,00	55.00	4.004.000	5,540,67	01.90	0.600,004	0.000.00	1506	4,000,000	5500000	30.00	30.774,000	
2012/06/08 (金)	20,000,000	10,840	80,00	20,600	5100	1,001.00	69.51	10(816)70	479.64	43.24	869,270	500.50	26.00	583,000	1,426.00	10.30	1,700,000	1,070.00	8541	1,860,864	4,000.00	79.00	1,00,00	5,710,60	63.26	10,800,676	
2012/06/09 (土)	379/9/30	19,600	4030	20,007	100	1,011.00	8180	110,070,000	480.74	3436	364,603	800.00	36.56	864,653	1,000,00	80.00	1,000,000	1,000.67	76.01	2,843,963	1,806.80	10/16	LIBERT	1,670.69	60.03	25/15/02	
2012/06/10 (日)	30,007,784	78409	40.06	200,710	500	1,087.06	48.73	100,007,004	475.00	37.99	400,000	779.00	8.87	508,700	1,460,000	16.00	1,004,000	5,460,07	9109	1,040,041	1,806.56	(4.00	1001090	5,460,76	98.23	23,789,096	
2012/06/11 (月)	100,000,004	7488	4000	190,000	400	1,500,75	99.00	TRANSPORT	999.03	4100	470,700	100.00	90.00	100,000	1,000,000	55.04	LIBERRO	100000	91.00	1,000,000	3,004.00	10.00	preme	1071000	99.59	20,000,000	
2012/06/12 (火)	20,000,000	15,600	4070	17,600	560	1,840,01	99.96	100,000,000	580.67	4135	800,000	888.29	88.79	100,200	1,000 (0)	18.00	1,000,000	101002	67.46	2,549,000	1,718.00	10.70	4/24/009	100000	53.84	17,881,685	
2012/06/13 (水)	2024 (44)	97.9	10.00	19,679	261	1,000.00	40.00	MARKS.	489.00	3748	361,601	1,000,70	38.00	653,796	1,086.00	2000	996,070	2502109	67.86	2,000,000	1,001.70	40.00	1,000,000	107904	40.00	938978	
2012/06/14 (木)	000000	0.800	200,000	HOT	960	795.00	27.81	407940	370.09	76.73	20,000	1,790,00	61.00	105,660	760.00	28-97	190-90	107900	20.49	201,000	1,081,77	19.00	796079	20000	33.44	1,043,000	

#### ※売上(SNSポイント)に関して

- ・1円=1SNSポイントです。
- ・消費税込みの数値となります。
- ・あくまで概算のため、PFへの各種手数料等は考慮していません。
- ・モンスターパスターは、モンパスゴールドは含まれていません。

#### ※各数値に関し

課金UU = 課金したユニークユーザ数,課金率 = 課金UU÷UU,ARPPU = SNSポイント÷課金UU,ARPU = SNSポイント÷UU

#### ※インストール数に関して

・Appleのアプリにつきましては、厳密には「インストール数」ではなく、「初回起動数」となります。

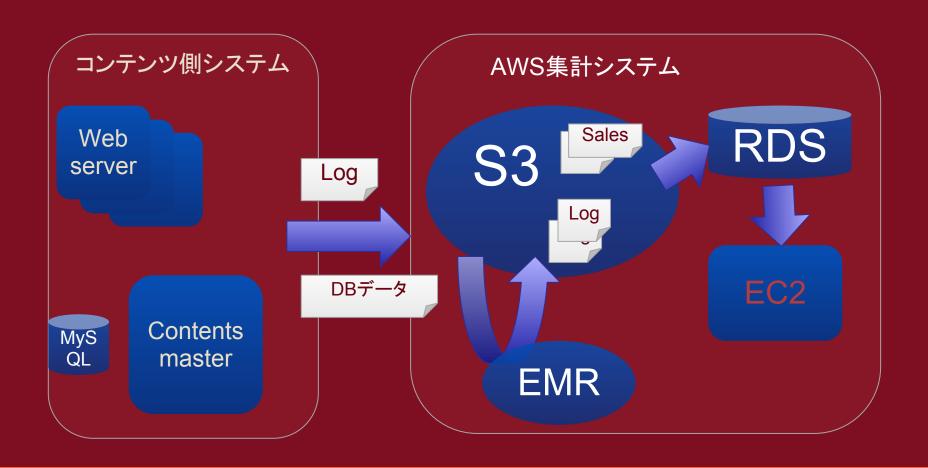
# データマイニングチームのミッション

- 開発
  - 多くのゲームのデータを統一フォーマットで収集・蓄積 → 分析に利用できるように
  - o 定常的なレポートの整備
- 企画
  - で データを分析し、企画・運用チームへフィード バック → 改善へ
  - o 定常的なレポート項目の追加
- KPI: key performance indicator

## ゲームの運用に役立つ情報

- 分析に必要になったときに、使える状態に最低限蓄積されていればよい
- 深い・細カルト情報
- 標準化されたフォーマットで蓄積

#### KPIシステム構成



#### システム構成

- 集計スクリプト(python)
  - 各プロジェクトのサーバーで稼動し、アクセスログ +DBデータをシリアライズ→送信
  - o データフォーマット
    - メタデータ + タブ区切り
      - 売上ログ
      - アクセスしたユーザー
      - 招待
      - 新規インストール
      - レベルのスナップショット
      - 登録日、スマートフォン/フィーチャーフォン、UAなど
    - アクセスログ

#### システム構成

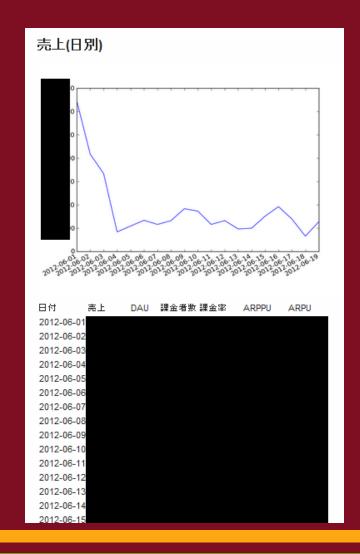
- AWSデータストレージ
  - EC2: スクリプト動作用
    - 集計・データのチェック
    - ■メールレポート送信
    - 分析用マシンとしても稼動
  - o S3: アクセスログおよびKPIデータをファイルで格納
  - RDS: 分析用にデータを格納
  - EMR(Elastic MapReduce):
    - ログをバイナリフォーマットに変換

#### HTMLレポート

- HTML静的ファイル
  - グラフと基本データの表
- 基本KPI
  - 売上、DAU、新規登録など
- アクセスデータ
  - 滞在時間
  - パス別滯在時間
- 登録日別データ
  - 登録日別売上、UU

#### HTMLのメリット

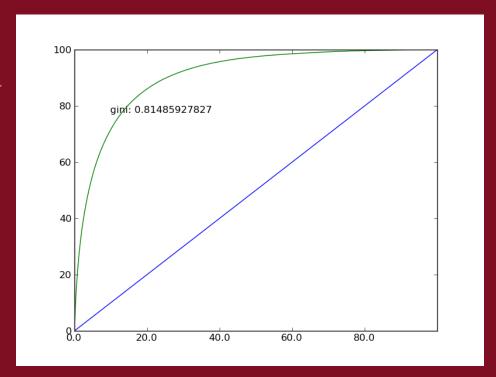
- 迅速に項目追加
- HTMLテーブルはエクセルにコピペできる
  - 分析者が利用しやすい
- ・図を載せられる



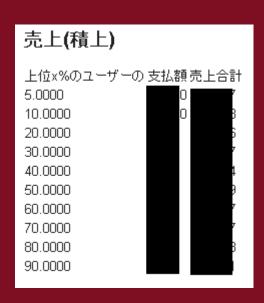
#### ジニ係数

ジニ係数: 主に所得格差の指標として使用される

ここでは購入額の差を数 量化



## デシル分析



上位x%のユーザーで売上 の何%をしめるか

#### アクセスログのバイナリ化

#### msgpack:

JSONのような構造化バイナリフォーマット

#### http://msgpack.org/

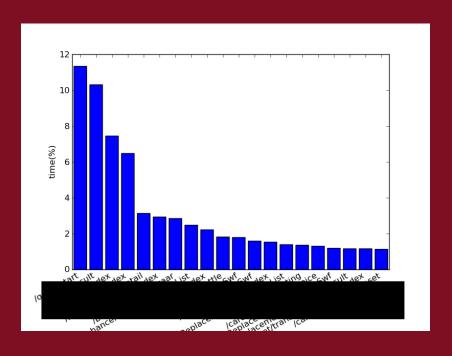
- opensocial\_viewer\_idごとにファイルを100 分割
- 4カラムのバイナリデータに変換

の高速化

ログ検索

時刻	opensocial_vi	URL	GET
(unixtime)	ewer_id	(path)	param

#### バイナリログのメリット



- パス別の滞在時間
- ユーザーの行動履歴 などが高速に計算できる

## Elastic MapReduce

Amazonのインフラ上で動くHadoop 13台の仮想マシンによる分散処理で、ログを変 換

 $0.06 \text{ per hour} \times 13$ 

## 統計パッケージの利用

#### R

- o オープンソースの統計パッケージ
- ο 独自の言語を使用
- o プログラマーには扱いやすい

#### SPSS

- IBM製、有償
- o GUIで統計解析を実行
- o 企画・マーケティングチームの利用を想定

#### 統計パッケージの役割

- 統計パッケージ上で、データを組み合わせ、 解析にかけることで、様々な仮説をスピー ディに検証できる
  - o 前提として、データが取得できる状態になければ ならない
- 有益なレポートは定常レポートへ
  - Python: numpy + matplotlib
- 運用チームへのフィードバックも重要

# 質問など