

SEAGAIA 2018: EHRの現在

機械学習を用いたEHR/PHR分析のためのデータ収集と予備検討



medical informatics

京都大学医学部附属病院
先制医療・生活習慣病研究センター
杉山 治



自己紹介

- 私：京都大学医学部附属病院
- 先制医療・生活習慣病研究センター 特定講師
- 先制医療・生活習慣病研究センターとは
 - 年会費：¥500,000（おおよそ）
 - 病気の予防に貢献する最先端の検診環境
 - 最先端の医療機器による健康診断を受けることができる
 - 有益な医療情報とともに幅広い研究を展開
 - ほとんどの検診データを研究に提供
- 私のミッション：
 - 先制医療センターの情報部門
 - 先制医療センターをはじめとした、京都大学医学部附属病院の電子カルテデータを用いた研究の支援
 - データウェアハウス（DWH）システムの整備・運用



<https://www.kuhp.kyoto-u.ac.jp/info/20160629.html>



<http://www.himedic.jp/facility/kyoudai.html>

機械学習を用いたEHR/PHR分析のためのデータ収集と予備検討

EMMR

- 京都大学医学部附属病院におけるデータウェアハウスシステム
- DWHの今後



PHR

- 長浜健康ウォーク
- 長浜健康ウォークのシステム
- 健康ウォークで収集したデータの解析



medical informatics

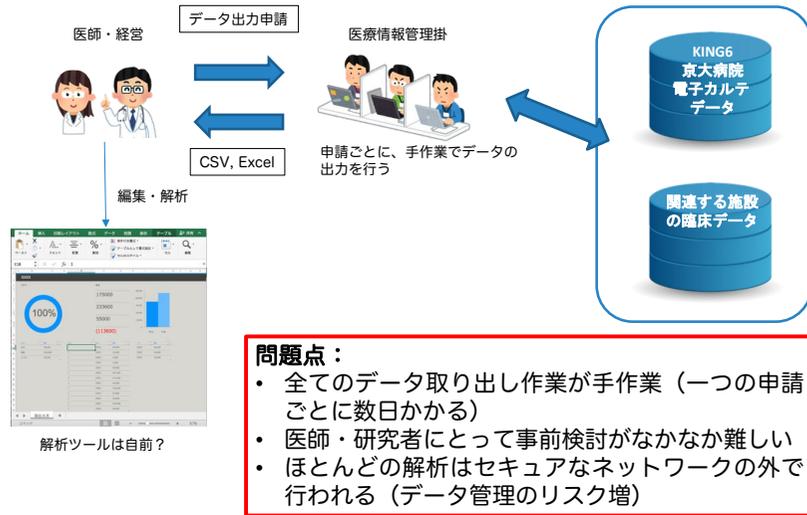
EMR: 京都大学医学部附属病院におけるデータウェアハウスシステム

京都大学医学部附属病院
先制医療・生活習慣病研究センター
杉山 治



データウェアハウス導入の要求 これまでの指標公開・臨床研究の流れ

京大病院電子カルテシステム内の
セキュアなネットワーク



京都大学医学部附属病院におけるデータウェアハウスシステム

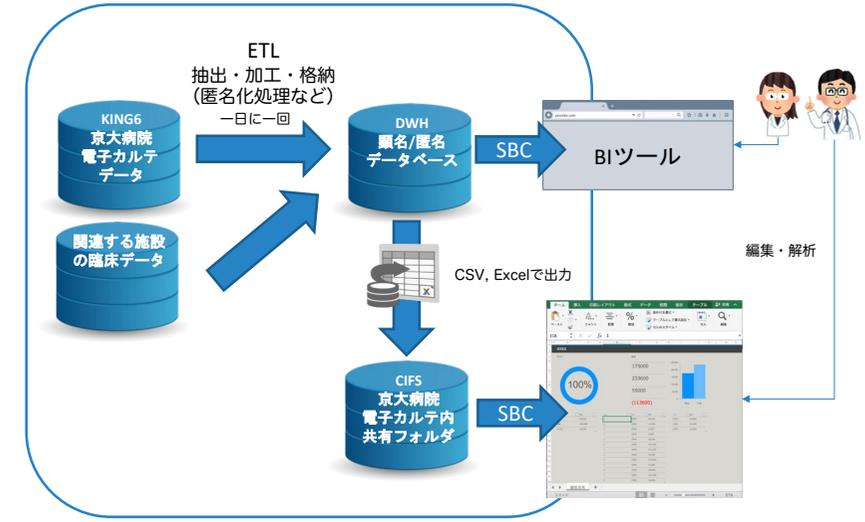


DWHを構成するHW



電子カルテからのデータの流れ

京大病院電子カルテシステム内のセキュアなネットワーク



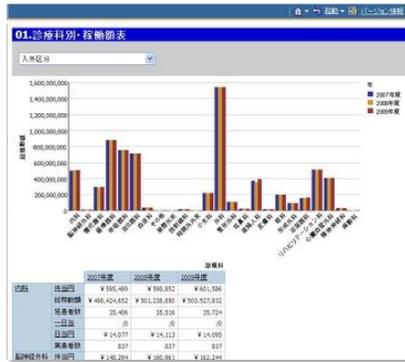
臨床研究・経営支援・指標公開の作業を全てセキュアなサーバ上で実施可能

2つのアプリケーション

集める 貯める 見せる

MD-View

様々なデータを集計し、可視化するBIソフトウェア



CPDP

時系列イベントの分析



MD-View

様々なデータを集計し、可視化するBIソフトウェア

10

帳票開発画面：Query Studio

集める 貯める 見せる

帳票開発画面：Report Authoring

集める 貯める 見せる 12

診療科別 月間アレグラ処方状況

科	1	2	3	4	5	6	7	8	9	10	11	12	累計
内科	3.3	24.1	5	4.5	13.8	24	35.9	33	23	21.5	21.2	13.4	212.9
外科	14.7	19.1	9	24.8	8.9	24.6	32	22.2	3.2	4.7	10.5	179.6	64
産科	5.9	5.2	1.7	7.2									25.8
小児科	3.8	5.4	19.2	14.5	19.8	22.6	48.9	51.5	19.7	34.3	5.1	5	222.8
計	63.4	56.3	81.9	85.9	68.8	95.7	95	92.5	36.4	48.4	52.7	31.7	648.9

ここで上のようなクロス集計をもとに、グラフを作成

DWH内のデータ項目は、ツリー図で表現される。帳票に必要な項目は、クリックして選択し、挿入ボタンをクリックすることで、画面右側の帳票に配置できる。

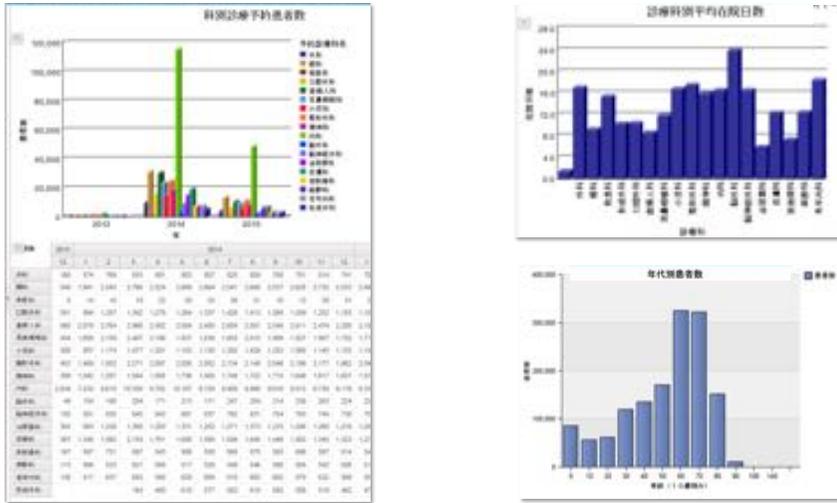
データ取得用のクエリー作成

DWH内のデータ項目がツリー図で表示され、ドラッグアンドドロップで取得対象の項目として配置できる。

CognosおよびDB2で用意されている関数を使って、複雑なデータ加工・編集を行うこともできる。

データ表示用の画面作成

出来上りの見た目そのままに (WYSIWYG)、画面を作成していくことができる。



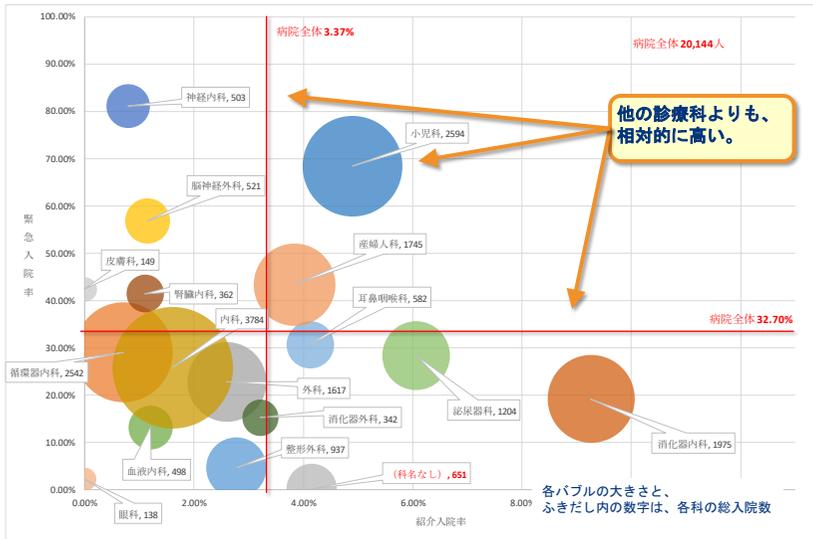
Confidential - ××病院 / IBM Japan

IBM MD-Viewをもとにした データ利活用の応用例

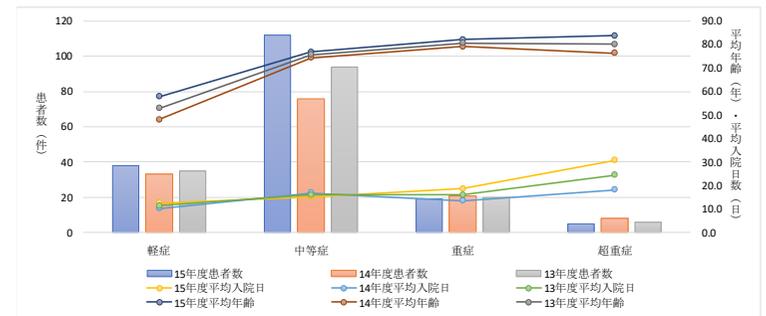
主に経営指標の作成について

本資料のグラフ、および元になった数字は、この資料用に作成されたものであるため、一般的な医療現場のデータの傾向とかい離している場合があります。また、画面はIBM Cognos Business Intelligenceで出力したデータを、Microsoft Excelで加工しています。

17



重症度	2015			2014			2013		
	患者数	平均入院日数	平均年齢	患者数	平均入院日数	平均年齢	患者数	平均入院日数	平均年齢
軽症	38	12.6	57.8	33	10.2	48.1	35	11.4	52.9
中等症	112	15.2	76.8	76	16.8	74.4	94	16	75.6
重症	19	18.6	82.1	21	13.6	79.2	20	16.1	80.6
超重症	5	30.8	83.8	8	18.2	76.4	6	24.5	80.1
合計	174	15.4	73.6	138	14.9	69.3	155	15.1	71.2



CPDP

CPDPの概要

CPDPは、Clinical Pathway Discovery and Planningの略称で、顧客や患者など、人ごとに起こったイベントを、時間と紐付けて分析する環境を提供する。例として電子カルテを入力とした場合、患者ごとに症状、病名、処方等の情報をイベント列として管理し、イベント列のパターンによる検索やイベントの順序、時間差を用いた分析が可能。

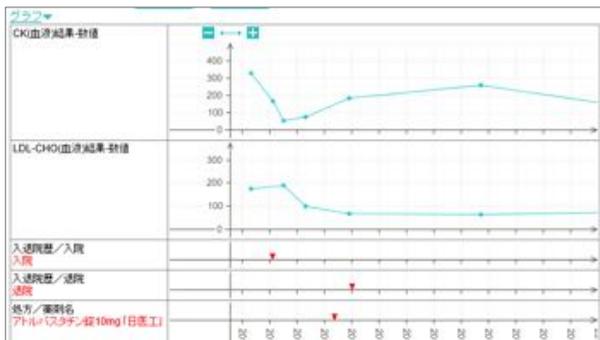


分析例：

- パターン 気管支鏡検査 → 気胸発生(2日以内) に該当する患者を検索
- パターン 脂質代謝異常 → 処方 → 副作用(CK高値) に該当する処方の内訳を提示

イベント検索の例

- LDL-CHO(血液)が140以上となった後、アトルバスタチンを処方した患者の時系列を表示します。
- その後、時系列の表示項目にCK(血液)を加え、アトルバスタチン処方後のCKの上昇を確認します。

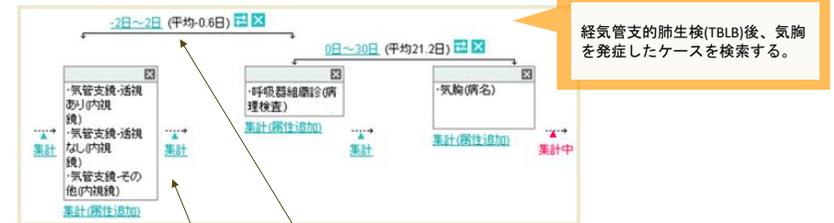


イベント列の分析例

病名マスタには無い合併症を検索

副作用や合併症として起こった症状を通常の病名検索と区別して検索したい

⇒ 指定した順序、時間差でイベントが起こったケースを検索できます

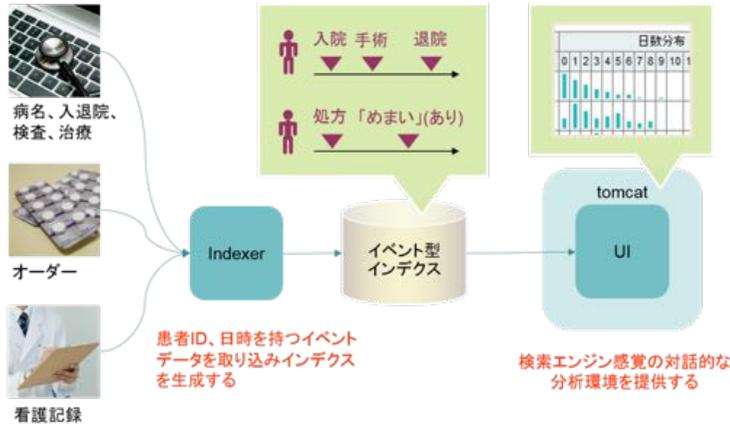


経気管支的肺生検

時間差を2日以内にセット

気管支鏡イベントのOR条件

CPDPはイベントデータの入力からインデクスを生成するIndexerと、インデクスを読み込みユーザーが操作可能な画面を提供するUI(ユーザーインターフェース)から構成されます。



イベント列の検索/分析

病名より詳細なレベルでの患者の絞込み

- 合併症の患者は何名か?

気管支鏡 → 気胸

- 副作用の原因となった薬剤の内訳は?

脂質代謝異常 → ? → 副作用

対話方式

高速なGUIで途中結果を見ながら、検索エンジン感覚での分析が可能

スケーラビリティ

1サーバーで18万名の患者、3000万イベントの動作実績あり

テキスト分析

テキスト中のキーワード、検査値で分析 (診療記録テキスト欄)

体重58kg BP: 171/102

抗がん剤 → 血圧≧150

副作用の直前に使われていた薬剤を検索

イベント分析のビューでは、指定したイベント列の前、後、間に起こるイベントのリストを表示することができます

病名登録とクリアチンキナーゼの検査の間のイベントを分析

0日~30日 (平均4.3日)

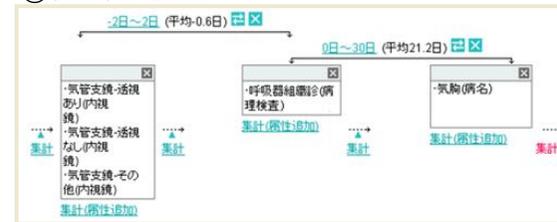
分析ビュー (患者情報 | イベント | 数値 | 時間差 | イベント列 | 集合) CSV

カテゴリ: 処方 / 薬名
集計範囲: 直前のイベントから 0日~30日 直後のイベントまで 0日~30日

選択	イベント	患者数	▼相関	前日数差平均(±SD)	後日数差平均(±SD)
<input type="checkbox"/>	[ALL]	188	0.9	0.8(±3.8)日	4.4(±7.9)日
<input type="checkbox"/>	タナトリン錠	34	19.3	0.2(±1.0)日	1.2(±4.1)日
<input type="checkbox"/>	プラビックス錠	54	17.3	0.2(±0.7)日	2.8(±6.7)日
<input type="checkbox"/>	バイアスピリン錠	65	7.7	0.5(±1.7)日	2.6(±6.4)日
<input type="checkbox"/>	ニコランジス錠	24	7.7	2.2(±6.5)日	5.5(±8.2)日
<input type="checkbox"/>	ブラバスタチンNa錠	9	5.8	2.7(±7.6)日	6.1(±7.4)日
<input type="checkbox"/>	リピトール錠	21	5.0	1.4(±3.8)日	5.0(±8.0)日
<input type="checkbox"/>	バリエット錠	40	3.8	2.1(±6.8)日	1.3(±3.4)日

- 課題: 最終の出力がBIツールだけだと、コーディングできる専門家や、システムインテグレーションができずに使いづらい
- 解決: CPDP機能のAPI化

① 検索条件を編集



② ブックマークを保存 (bookmark id)

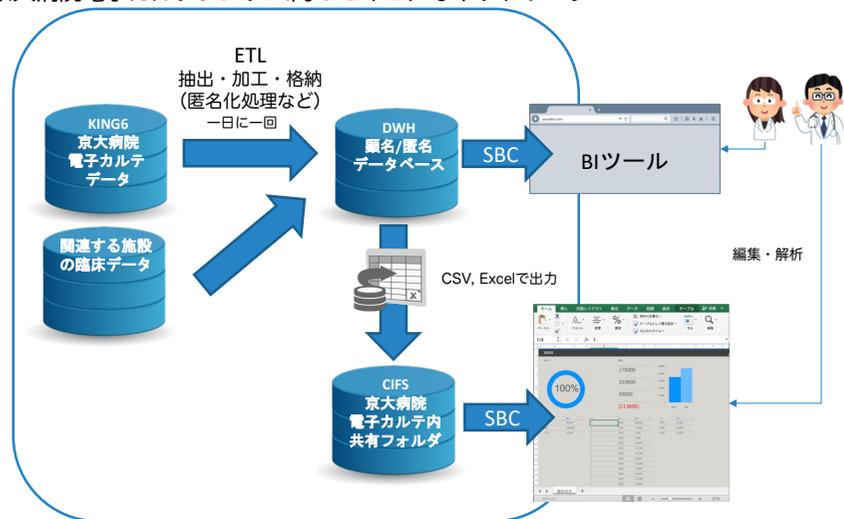
③ ウェブAPIから、機能をいつでも呼び出すことができる

<http://XXX.XXX.XXX.XXX/call?bookmark=YYY>

CPDP機能をもつさらに高度なアプリケーションの開発へ

即応性の高い解析環境の提供が先端医療につながる (と期待しています)

京大病院電子カルテシステム内のセキュアなネットワーク



臨床研究・経営支援・指標公開の作業を全てセキュアなサーバ上で実施可能

DWHを運用する上でわかってきたこと

- 集計・統計・検索・可視化が可能になったとはいえ、まだまだ医師が望む高度な解析を行うためには、いろいろなノウハウの蓄積が必要
 - 使用者が求める情報の単位と、DWHが出力する情報の単位が異なる
 - 使用者が求めるデータ例：患者(N人)ごとの検査値一覧(N行)
 - 実際に出力されるデータ例：患者(N人)・検査値(M個)のレコード一覧(N x M行)

より医師の視点に近い高度な検索機能の実現が求められる

- 様々な観点からのデータクレンジングの必要性
 - 電子カルテの情報は完全なものでなく、様々な抜け落ち用語のずれが存在する
 - 例：TG → トリグリセリド → **中性脂肪**
 - 検査結果も数値の項目に、「0.08ミマン」のような文字列が混入するケースも存在する (検査部に確認すると色々な例外ケースが定義されている)

あいまい検索・単語候補の提示機能の実現が求められる

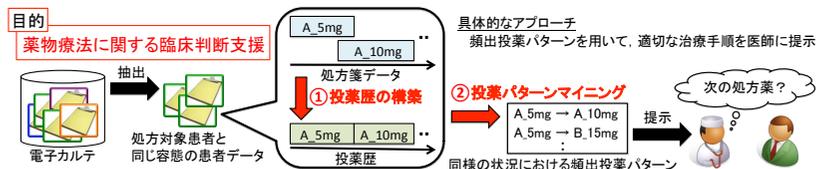
- 検索におけるノウハウ
 - 電子カルテのテキスト検索は、簡単ではない
 - 例えば、「嘔下」というテキストを検索しても、「嘔下なし」といった書き方をしているテキストも検索結果として出てきてしまう。意味単位で検索をかけることは単純なテキスト検索では困難

単語レベルではなく、文脈理解などの自然言語解析が求められる

DWHのこれから

より高度なETL機能の導入 (以下はまだ未導入)

例：電子カルテの処方箋データからの投薬歴の構築と構築された投薬歴を用いた投薬パターンマイニング[森田ら'17]

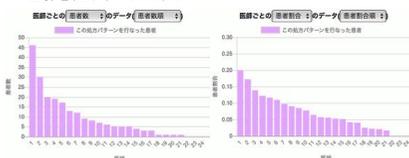


上記のデータの可視化による診療支援 (以下はまだ未導入)

例：他の医師の処方方を提示することによる臨床判断支援システム[加田ら'17]

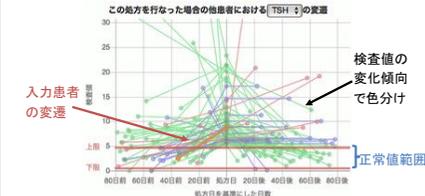
医師ごとの処方患者グラフ

その処方を選んだ患者数や患者割合を医師ごとに表示したグラフ。評価アルゴリズムで用いられていた三つの評価値の理解を促す目的がある。



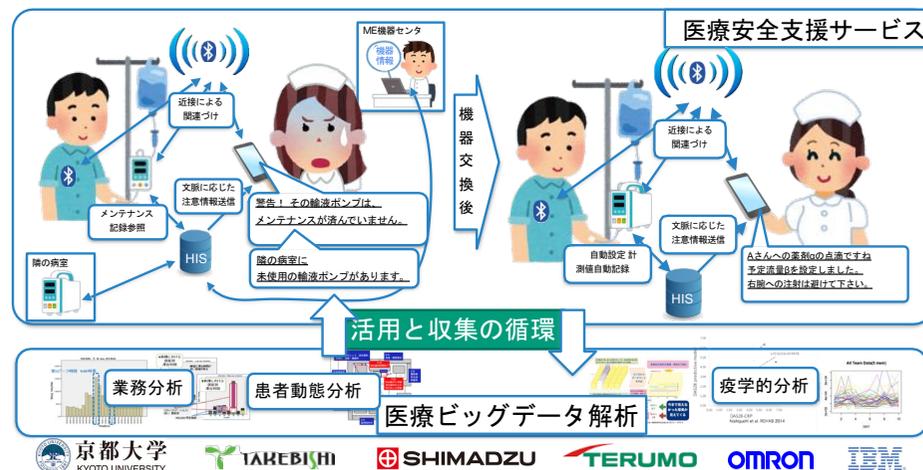
検査値変遷グラフ

その処方を選んだ患者の処方日前後の検査値のグラフ。この処方を行なった場合の他患者における「TSH」の変遷



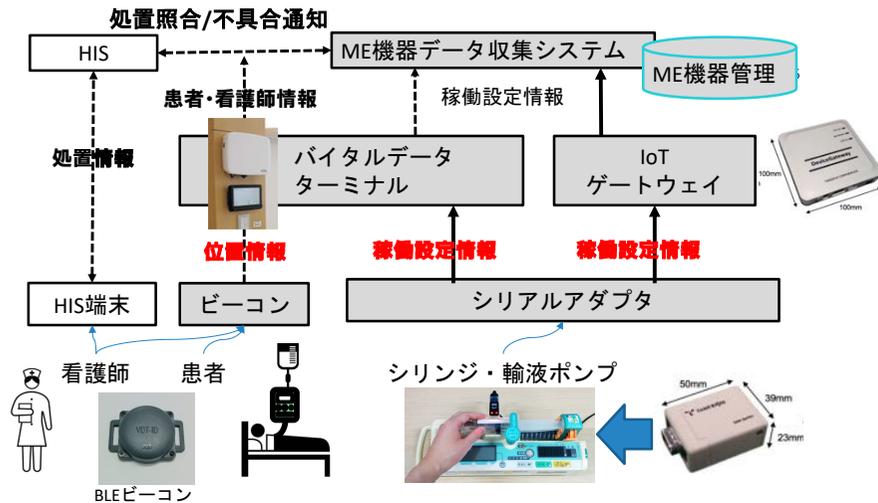
DWHのこれから：医療機器の電子カルテの解析

医療機器をIoTデバイスにして、その情報を収集・解析



DWHのこれから：医療機器の電子カルテの解析

シリンジ・輸液ポンプの位置・動作情報の収集と利用



EMR:まとめ

- 京大病院に導入されたデータウェアハウスシステムの紹介
 - MD-View: 様々なデータを集計し、可視化するBIソフトウェア
 - CPDP: 時系列イベントの分析
- DWHシステム導入にあたってわかってきたこと
- DWHの今後の方針

今後は

- ・これまで実現が困難だった診療科にまたがるデータの解析の実施
- ・ディープラーニング等の最新の機械学習の手法の応用
 - ・件数とスパース性の問題に対応が必要・・・

健康づくりの障害「きっかけ」

- 健康作りは大事なことは知ってます
- でもね、忙しいんですよ、時間が無いんですよ



PHR:ながはま健康ウォーク

medical informatics

京都大学医学部附属病院
 ○杉山 治, 黒田 知宏
 立命館大学情報理工学研究科
 鈴木 真生, 大槻 涼, 若尾 あすか, 松村 耕平, 野間 春生
 近畿大学理工学部
 多田昌裕



ちょっとしたきっかけを
 つくろう

3 きっかけを与える二つの方法

自動計測機器利用の効用

機器配布群が有意に継続率が高い

健康情報活用基盤構築に向けた標準化及び実証実験 平成21年度成果報告書

グループ化の効用

チーム実施で完走率向上

新潟県 NIGATA PREFECTURE

新潟県チーム > 防災・交通安全 > 507...の旅行券や温泉宿泊券が当たる！安全運転

50万円分の旅行券や温泉宿泊券が当たる！安全

今年もドライバー5人1組のチームを組んで、100日間無事故・無違反を目指し、無違反を達成したチームには、達成記念品をもらえなくプレゼント。さらに会社ぐるみはもちろん、友達・家族を誘って挑戦してみませんか？
参加費は1チーム1,000円（1人200円）です。

8,464チーム、42,320名が参加、88%達成（2014年度実績）

「見える化」と「仲間作り」がポイント

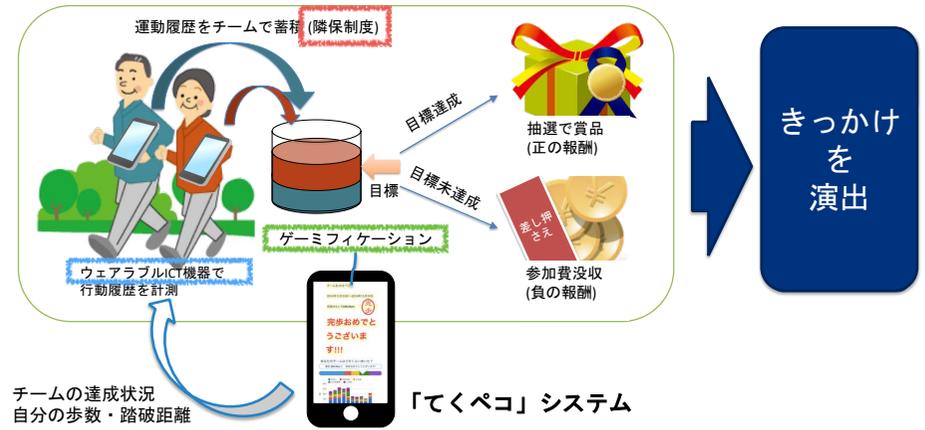
テレビ



2016年11月21日NHK
ニュースほっと関西

4 みんなで一緒にながはま健康ウォーク

10日間で一人40 [km] 歩こう！（一日約8000歩）



- 「イベント」として開催し「きっかけ」を演出
- 正(景品)と負(参加費)の二つのインセンティブ提供
- 「仲間作り」によるインセンティブ強化

4 イベントの結果(2017)

チーム数	300	10 - 11月
完歩	273	91%
未完	27	9%

完歩率 90% 突破

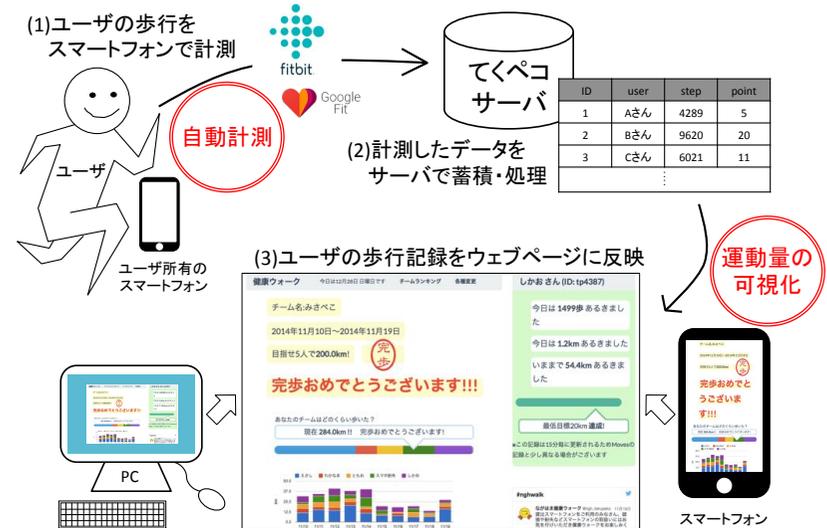
4
4  これまでの結果

	総数	完歩	未完	完歩率
14年11月	170	155	15	91%
15年09月	75	68	7	91%
15年10月	152	136	16	89%
15年11月	161	149	12	93%
16年10月	184	177	7	96%
16年11月	193	172	21	89%
17年10月	142	130	12	91%
17年11月	158	143	15	92%
延べ参加	1235	1130	105	91%

4
5  これまでの結果

	総数	完歩	未完	完歩率
1人参加	529	463	66	88%
3人チーム	381	354	27	93%
5人チーム	325	313	12	96%
全日程	1235	1130	105	91%

 てくペコ システム概要



 長浜健康ウォークのシステム
 medical informatics

京都大学医学部附属病院
 先制医療・生活習慣病研究センター
 杉山 治





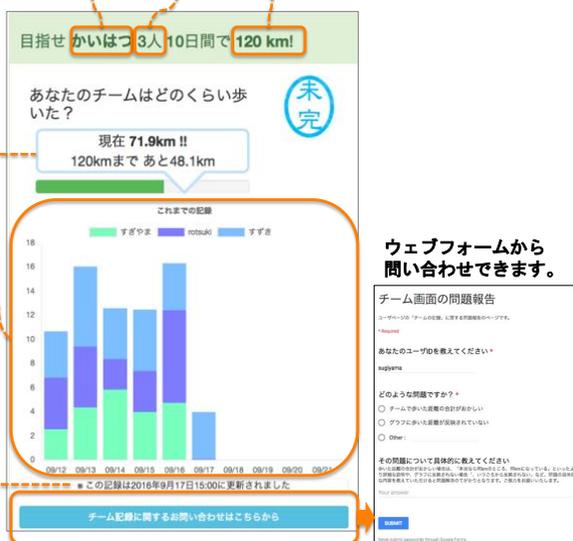
健康ウォークのユーザ画面

個人の期間中の歩行数
グループメンバーの歩行数
グループ全体の歩行数



チームの記録の見かた

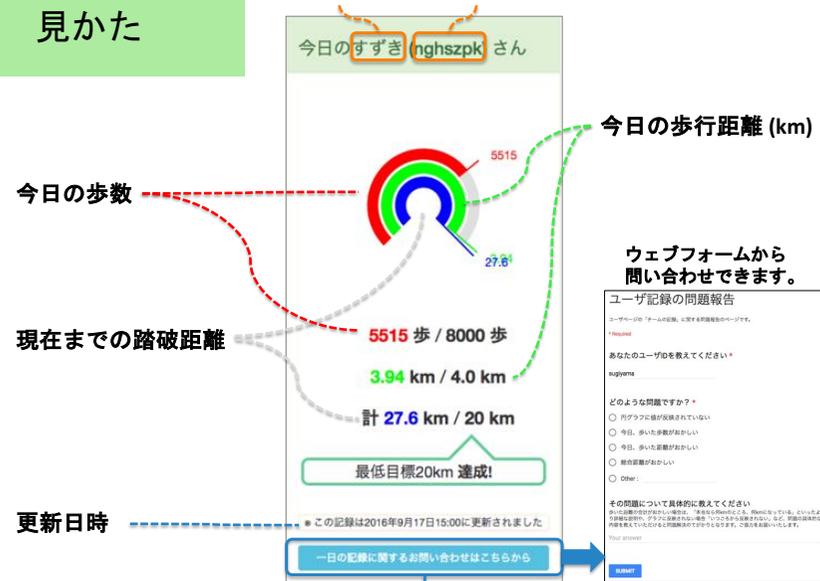
チーム名 人数 目標距離



50

個人の記録の見かた

ニックネーム ID



51

イベント主催者は収集した情報の一覧にオンタイムでアクセス+調整可能。

てくペコカスタム機能

ホームに戻る アカウントのバックアップ アカウントの反映

検索 "Fitbit記録" 実行

☆ 直近のアクション

100個の内ひとつも表示されています

アプリ	検索	User id	作成日	更新日	Api	Raw description
Backend	Fitbitプロフィール	<input type="checkbox"/>	2018年 5月18日 2:12	2018年 5月18日 2:12	https://api.fitbit.com/1/user/5MF6RK/activities/steps/date/2018-05-18/2018-05-18.json	Traceback (most recent call last
	Fitbit記録	<input type="checkbox"/>	2018年 5月18日 2:04	2018年 5月18日 2:04	https://api.fitbit.com/1/user/5ZZBHT/activities/distance/date/2018-05-18/2018-05-18.json	["activities-distance": [{"dateTi
	GoogleFitの記録	<input type="checkbox"/>	2018年 5月18日 2:04	2018年 5月18日 2:04	https://api.fitbit.com/1/user/5ZZBHT/activities/steps/date/2018-05-18/2018-05-18.json	["activities-steps": [{"dateTime"
	GoogleFitプロフィール	<input type="checkbox"/>	2018年 5月17日 7:02	2018年 5月17日 7:02	https://api.fitbit.com/1/user/5ZZBHT/activities/distance/date/2018-05-17/2018-05-17.json	["activities-distance": [{"dateTi
	Movesの記録	<input type="checkbox"/>	2018年 5月17日 7:02	2018年 5月17日 7:02	https://api.fitbit.com/1/user/5ZZBHT/activities/steps/date/2018-05-17/2018-05-17.json	["activities-steps": [{"dateTime"
	Movesプロフィール	<input type="checkbox"/>	2018年 5月17日 7:02	2018年 5月17日 7:02	https://api.fitbit.com/1/user/5ZZBHT/activities/steps/date/2018-05-17/2018-05-17.json	["activities-steps": [{"dateTime"
	Nfc connections	<input type="checkbox"/>				
	アクティビティ	<input type="checkbox"/>				
	トラックポイント	<input type="checkbox"/>				
	手動入力したアクティビティ	<input type="checkbox"/>				



サーバサイドに求められる案件

安価に

- 市役所にウェブアプリ開発費や、サーバ代を捻出してもらうことはほぼ不可能

早く

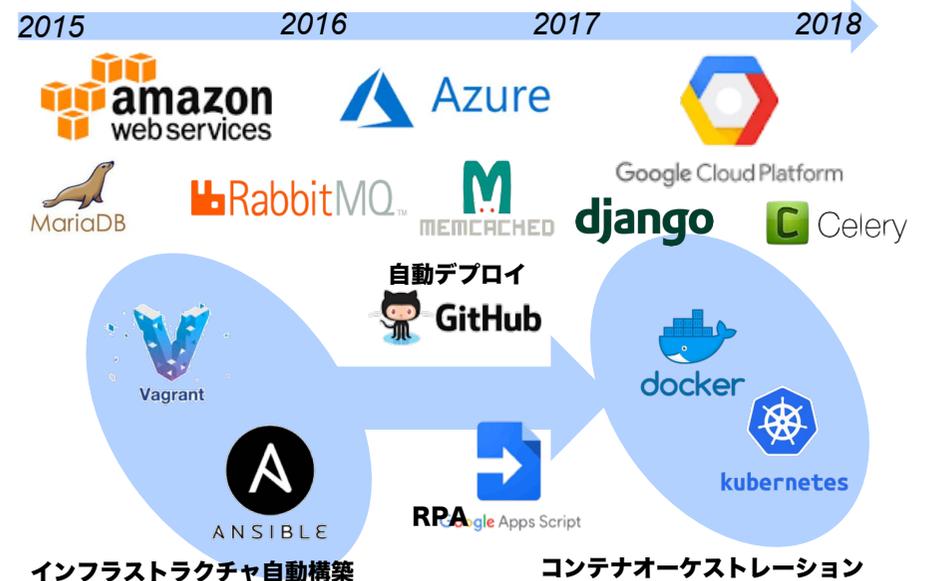
- 問題が起これば、即対応。数時間内にバグフィックスからリリース対応までを実現する必要がある

柔軟に

- この問題はレアケースだから、対応しないなどという判断はありえない（役所の職員は全てに対応しようとします）

だいたい、ソーシャルゲームとかに求められる要求と同じ・・・

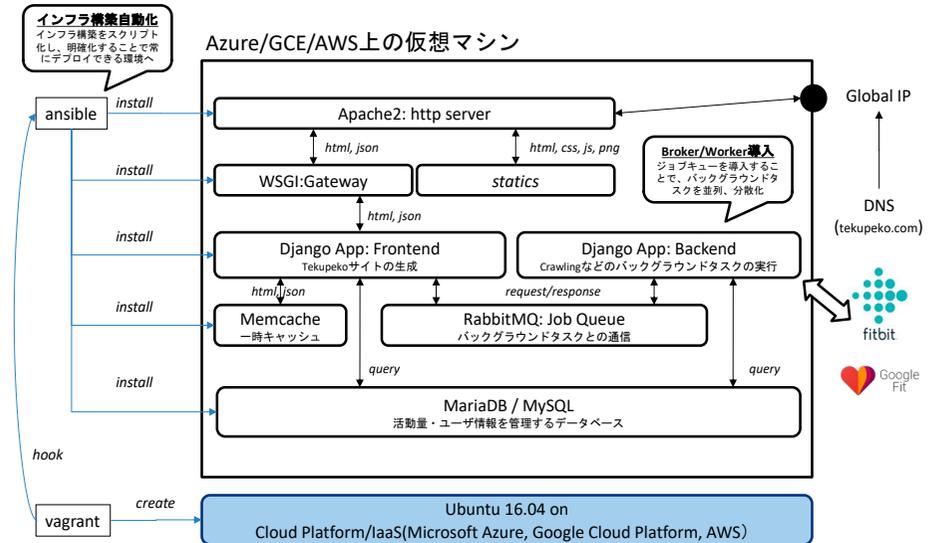
用いているサービス:
クラウドサービス・OSS使いまくり



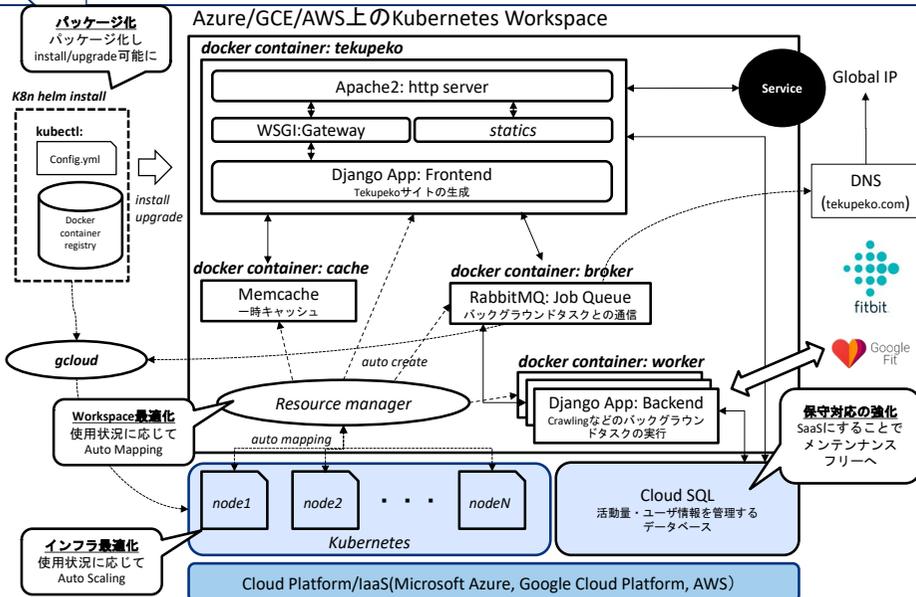
機能対応表

機能	2015	2016	2017	2018 (予定)
個人の活動量の可視化	○	○	○	○
グループの活動量の可視化	○	○	○	○
ゲーム進捗の管理	△	○	○	○
Moves対応	○	○	△	×
Fitbit対応	×	×	○	○
Google Fit 対応	×	×	○	○
Apple Health Kit対応	×	×	×	○
A&D 歩数計対応(NFC)	○	○	○	○
固有スマートフォンアプリ	×	×	×	(iPhone)/Android
OAuth read (データ入力一般化)	×	×	×	○
OAuth write (データ出力一般化)	×	×	×	△
インフラ自動構成	Ansible	Vagrant / Ansible	Vagrant / Ansible	Docker/Kubernetes

2017年度までのシステム構成



2018年度のシステム構成



コストからみる効果

見積もり	見積もり	見積もり
Microsoft Azure Estimate		
Service type	Description	Estimated Cost
Virtual Machines	1 A4 v2 (4 vCPU, 8 GB RAM) x 730 Hours; Linux - Ubuntu; 従量課金制; 0 管理 OS ディスク - S4	¥52,980.48
Backup		¥0.00
Application Gateway	0 instance(s) x 744 hours, 0 GB data processed, outbound transfers: 0	¥0.00
IP Addresses	0 Static IP Addresses	¥0.00
VPN Gateway	staticdynamic tier, 0 gateway hour(s), 0 GB outbound internet	¥0.00
ExpressRoute	Metered data plan, 1gbps port speed, Outbound data transfer: zone2 /Month in japan-west	¥0.00
Azure SQL Database	1 elastic database(s) x 744 hours, size:	¥0.00
Support	Support	¥0.00
	Licensing Program	Microsoft Online Services Program (MOSP)
Monthly Total		¥52,980.48

¥53,000 /month
 ↓
 > 10,000 /month

2017年度第一期の測定性能解析結果

■ 性能測定に用いた指標

- 更新時間：時間ごとにCrawlingしているデータの更新間隔
 - ユーザの体験として、どの程度のタイムラグでデータが更新されているかを確認する
- 歩幅：参加者の測定距離の平均 / 歩数の平均
 - 平均歩幅は大体 50 cm だと言われているが、そのような挙動を示すかを確認する

■ 統計を取ったもの

- iPhone, Android別, 参加者の内訳
- メーカー別の参加者の内訳
- iPhone, Android別, データの更新時間
- メーカー別, データの更新時間
- iPhone, Android別, 歩幅の差
- メーカー別, 歩幅の差

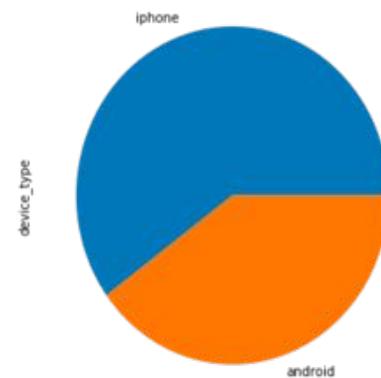
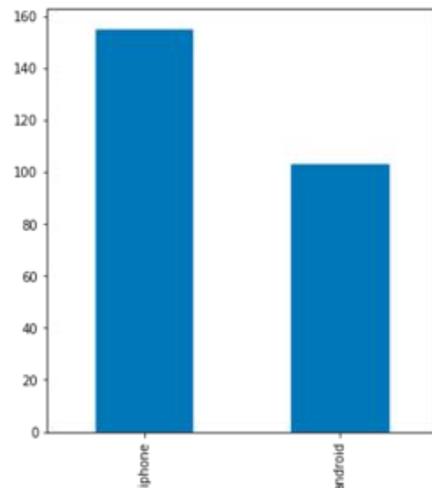
データ収集の裏舞台

medical informatics

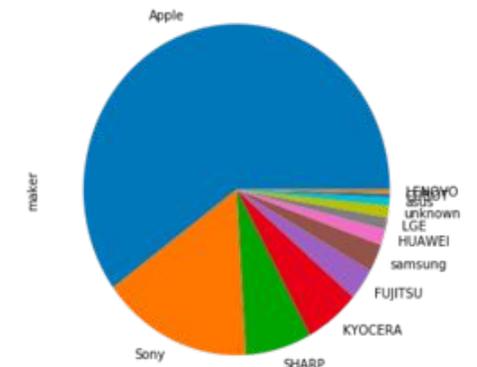
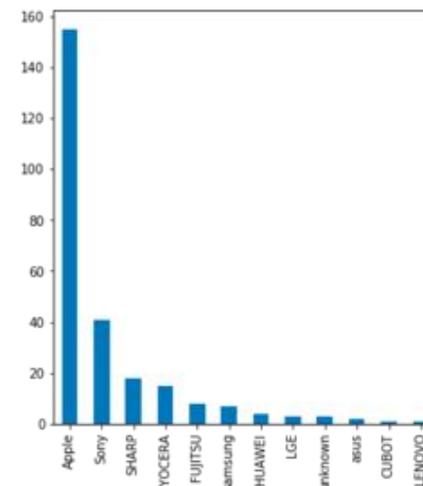


iPhone, Androidの内訳

参加者中、250名がスマートフォンを用いて、活動量を測定。その内訳は次の通り

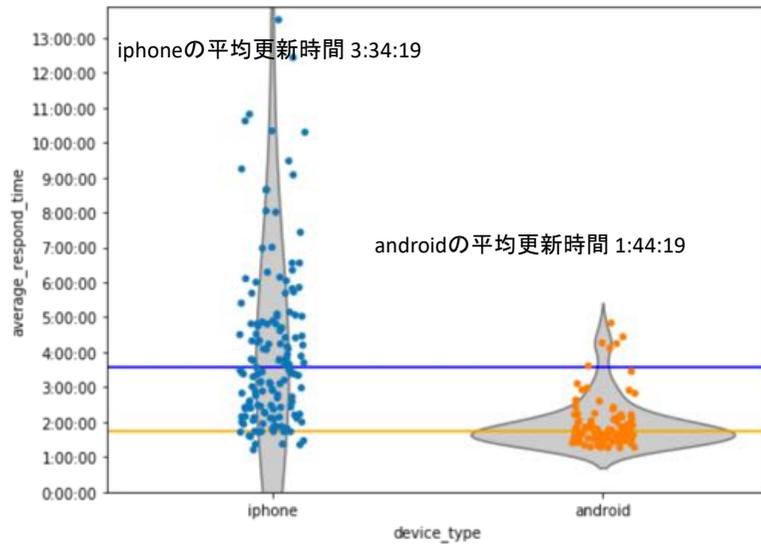


メーカー別の参加者の内訳



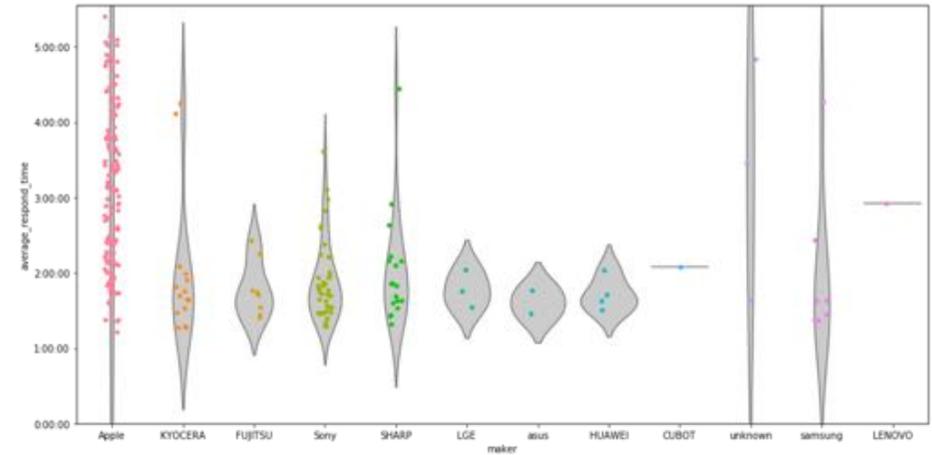
iPhone, Android別, データの更新時間

64



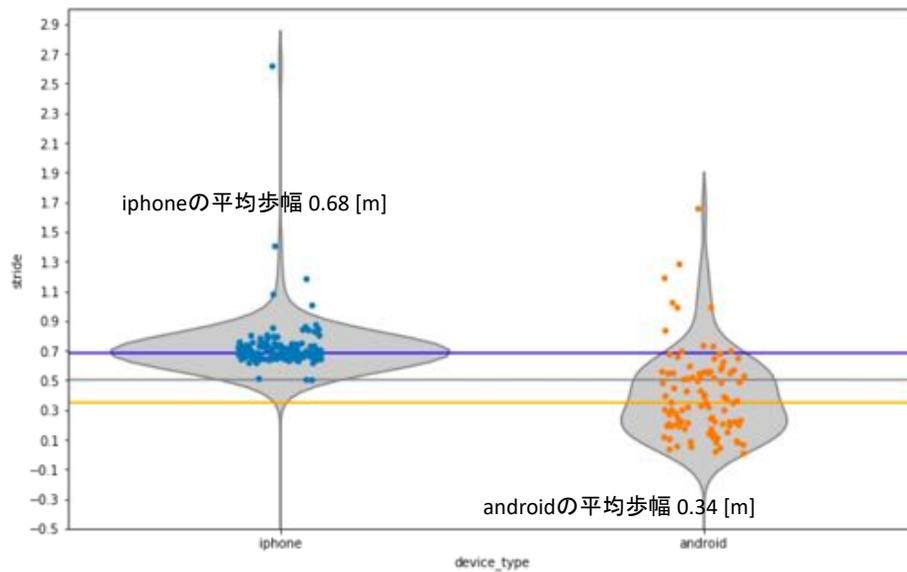
メーカー別, 更新時間の内訳

Fitbitは画面を開かないと同期できないのに対し、Androidはバックグラウンドで同期するので、Androidの方が同期速度が大きい



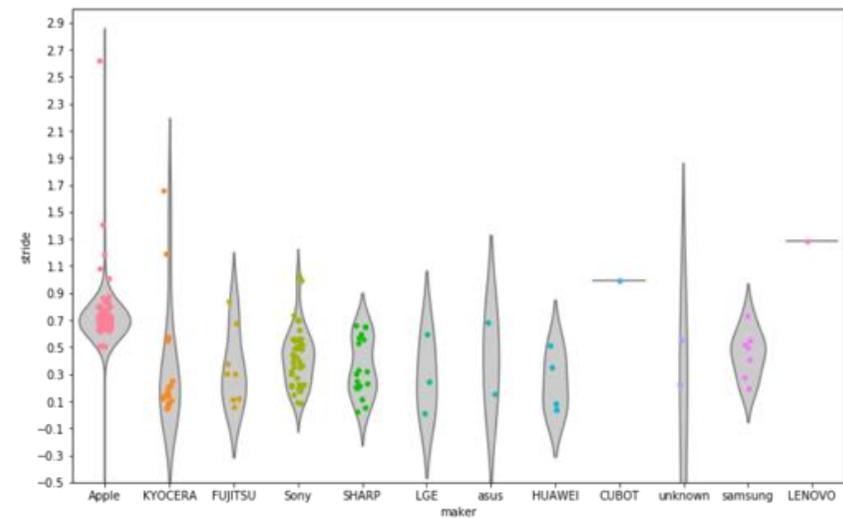
iPhone, Android別, 歩幅の差

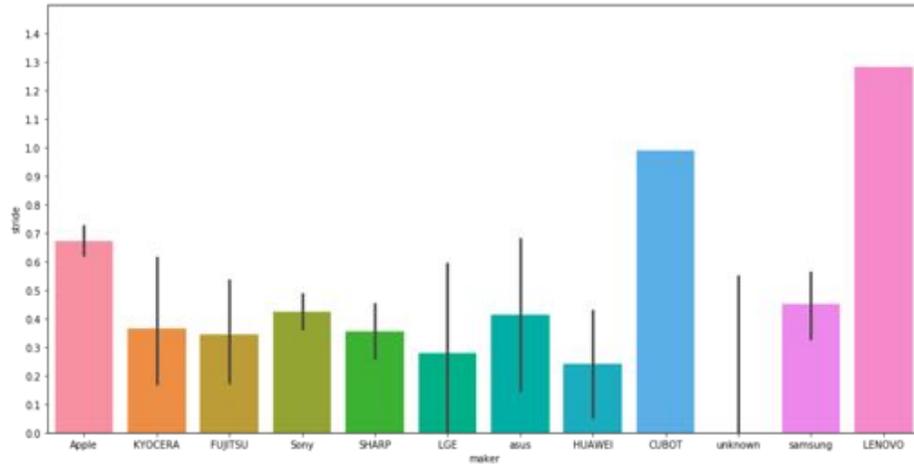
66



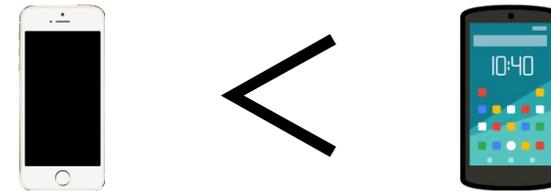
メーカー別, 歩幅の差 (1/2)

67

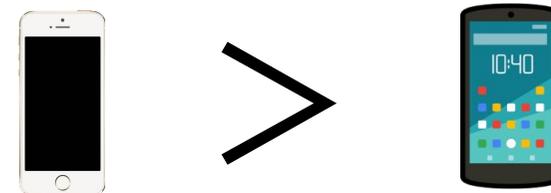




更新時間: Androidの方が更新性能は良い



計測性能: iPhoneの方が計測性能は良い



アンケート概要

- 「長浜健康ウォーク」の開始前、終了1ヶ月後、終了3ヶ月後の3回に分けてアンケート調査を実施。
- 例: 2014年の集計人数
 - 1回目(イベント開始前): 320名
 - 2回目(イベント終了1ヶ月後): 378名
 - 3回目(イベント終了3か月後): 321名

の方にご回答いただきました。

健康行動に対する意識調査 (定性)

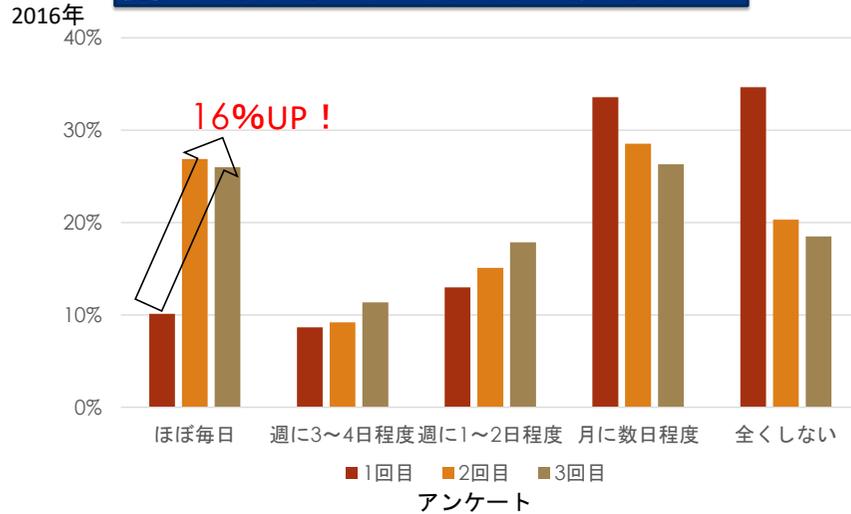
medical informatics

近畿大学工学部
多田昌裕



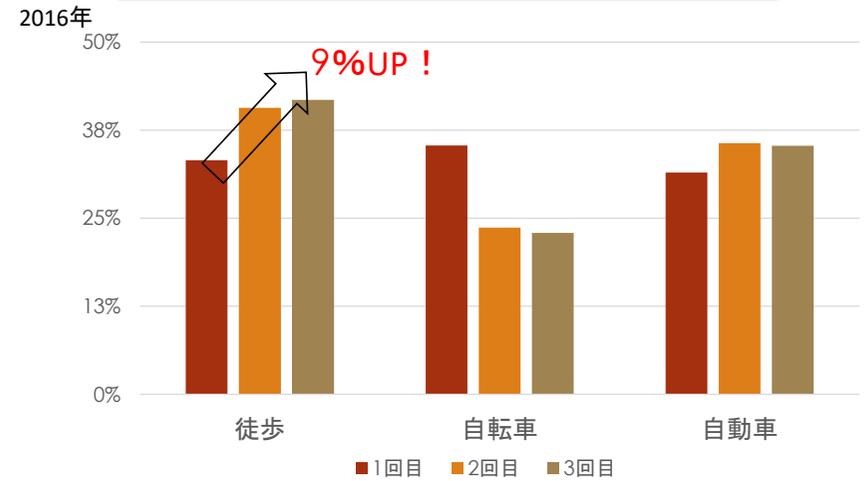
評価：行動変容は見られたか？

徒歩で外出することはどれくらいありますか？



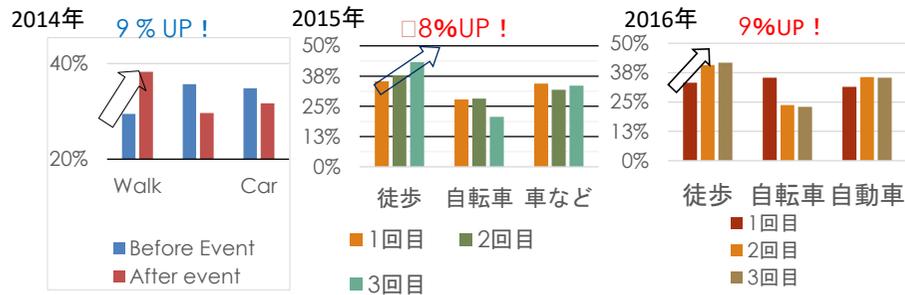
評価：行動変容は見られたか？

徒歩10分で行けるところへ出かけるとしたら、主にどのような方法で行くことが多いですか。



評価：行動変容は見られたか？

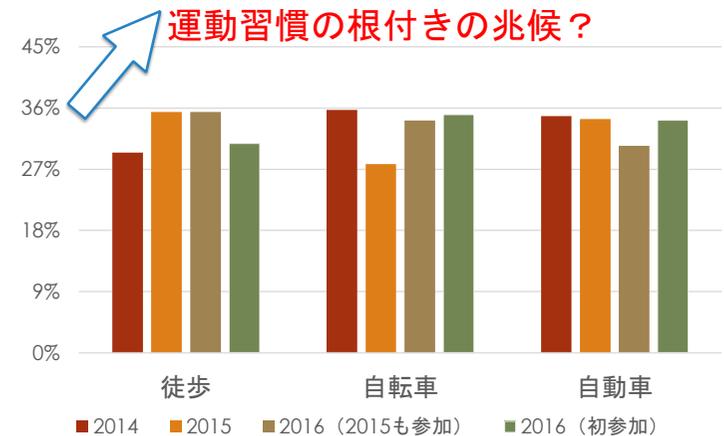
徒歩10分で行けるところへ出かけるとしたら、主にどのような方法で行くことが多いですか。



毎年着実に効果が出ている

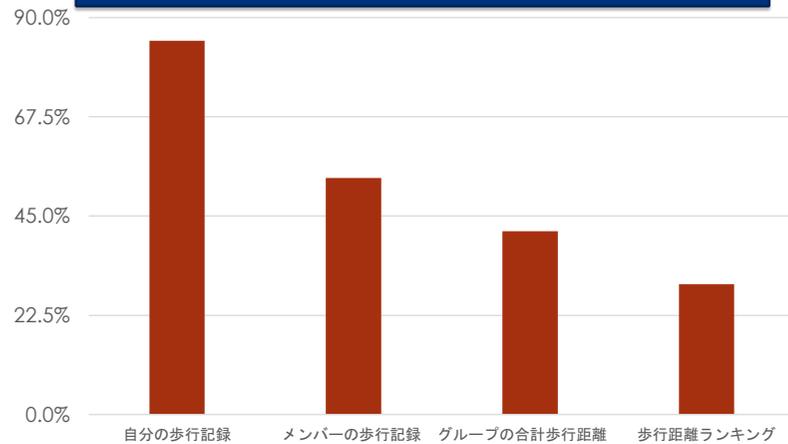
評価：行動変容は見られたか？

徒歩10分で行けるところへ出かけるとしたら、主にどのような方法で行くことが多いですか。



評価：チームでやるのは効果的？

このイベントで用いたシステムで、あなたの「やる気」を上げるものはありましたか？



チームでやるのは効果的？

グループでということ、足を引っ張らないよう頑張ろうと思えた。

5人グループで参加し、グラフを見ながらもう少しがんばろうとか他の人に迷惑をかけないようにという思いもあった。

解析一覧

- 日記実験による行動変容要因の解析
- 外的要因と歩行距離の関係解析
- 参加者の歩行状態の解析
- チーム参加の影響の解析

歩行に作用する要因の解析（定量）

medical informatics

京都大学医学部附属病院

○ 杉山 治, 黒田 知宏

立命館大学情報理工学研究科

鈴木 真生, 大槻 涼, 若尾 あすか, 松村 耕平, 野間 春生



歩行に作用する要素の調査

過去2回のイベントで約90%の人が目標を達成した。
10日間のイベント中の歩行を促すことには成功したが、
どんな要素が歩行に作用しているかは不明である。

そこで

日記調査を実施した

日記の内容をユーザの歩行記録
と照らし合わせ、歩行に作用する
要素を抽出した

回答者：ながはま健康ウォークに
チームで参加している人
の中で同意を得た8名

記入サンプル

朝

朝に書く 今日気分、予定、目標など

記入例) 今日は天気が良いし散歩しよう
ながはま健康ウォークの仲間と会う予定だ
今日はあんまり歩かないで休もう など

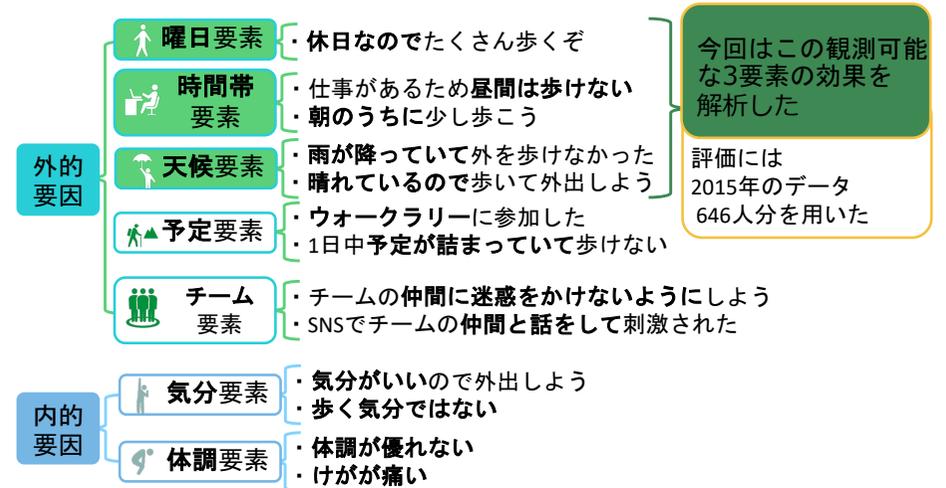
夜

夜に書く 今日一日はどうでしたか?
ながはま健康ウォークについて記録してください

いつ	どこで	どんなことがあった どんなことをした	どうだった 備考
お昼休みに	職場で	チームの仲間と 健康ウォークに ついて話した	他のチームの 人も一緒に 話した

歩行に作用する要素の調査

日記の内容から下記の要素が歩行に作用しているとわかった。



歩行に作用する要素の解析：天候要素

仮説：降雨のある日は無い日より歩行量が少なくなるのではないかと

降雨量が0mmの日の歩行量の平均値とそれ以外の日の歩行量の平均値についてt検定を行なった。

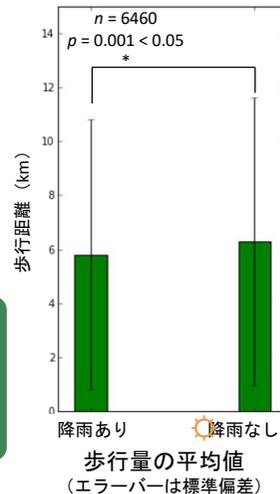
イベント期間における滋賀県長浜市の降雨の有無

	1日目	2日目	3日目	4日目	5日目	6日目	7日目	8日目	9日目	10日目
9月	☀	☀	☀	☀	☀			☀	☀	☀
10月	☀	☀	☀		☀		☀	☀	☀	☀
11月	☀			☀	☀	☀	☀			

降雨がある日の歩行量と降雨がない日の歩行量には5%の有意水準で有意に差があることがわかった。

降雨がある日は無い日よりユーザの歩行量が少なくなっていることを確認できた

全ユーザに対して、晴れ予報の日に積極的に歩くように働きかけることが有効なのではないかと



歩行に作用する要素の解析：曜日要素

仮説：休日は平日よりも歩行量が多くなるのではないかと

土・日・祝の休日の歩行量の平均値と平日の歩行量の平均値についてt検定を行なった。

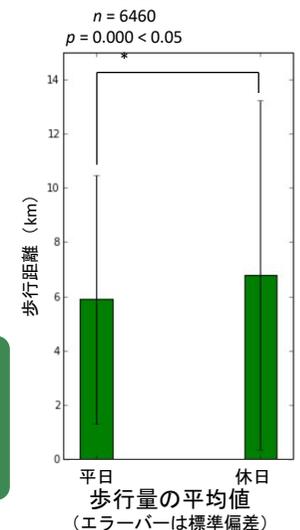
イベント期間における休日の分布

	1日目	2日目	3日目	4日目	5日目	6日目	7日目	8日目	9日目	10日目
9月	金	土	日	月	火	水	木	金	土	日
10月	火	水	木	金	土	日	祝	火	水	木
11月	日	月	祝	水	木	金	土	日	月	火

平日の歩行量と休日の歩行量には5%の有意水準で有意に差があることがわかった。

休日は平日よりもユーザの歩行量が多くなっていることを確認できた

全ユーザに対して、休日に積極的に歩くように促すことが有効なのではないかと



9 歩行に作用する要素の解析：時間帯要素

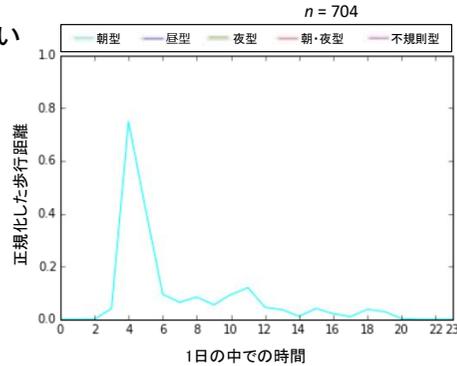
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



9 歩行に作用する要素の解析：時間帯要素

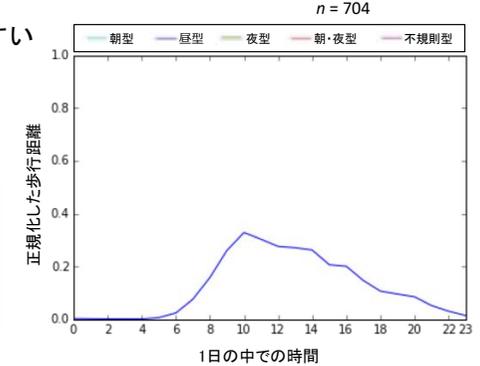
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



9 歩行に作用する要素の解析：時間帯要素

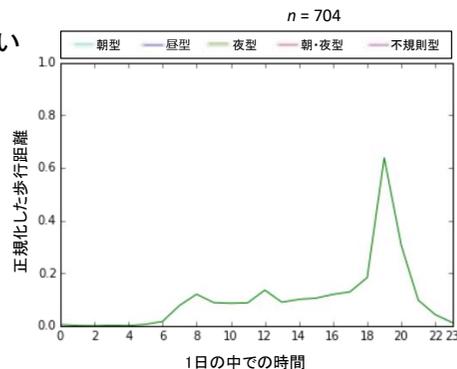
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



9 歩行に作用する要素の解析：時間帯要素

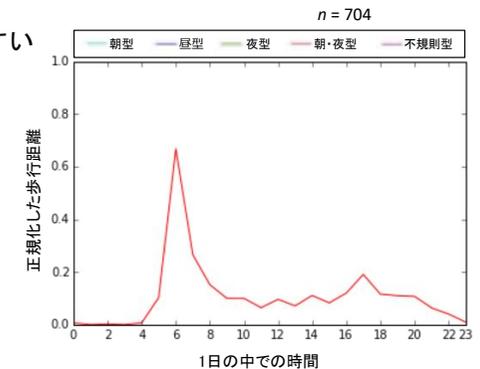
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



1 歩行に作用する要素の解析：時間帯要素

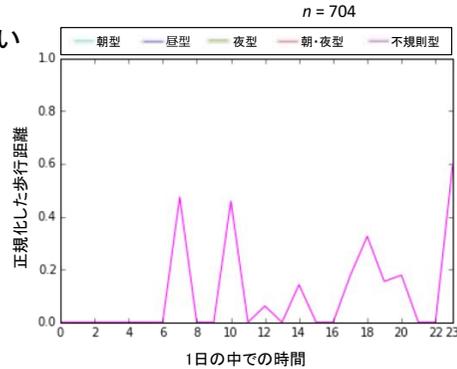
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



1 歩行に作用する要素の解析：時間帯要素

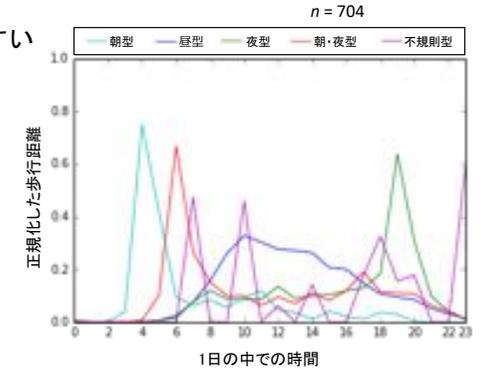
仮説：ユーザによって歩行できる時間帯に特徴が現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると朝型、昼型、夜型、朝・夜型、不規則型の5つのパターンが観測された。

ユーザによって1日の中で歩行しやすいタイミングが異なるといえる。

ユーザによって歩行できる時間帯に特徴が見られ、パターンを観測することができた

個々の歩きやすい時間に合わせて歩行を促すことが有効なのではないか



1 歩行に作用する要素の解析：時間帯要素

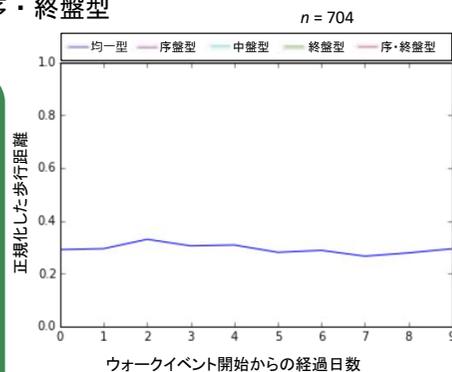
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



1 歩行に作用する要素の解析：時間帯要素

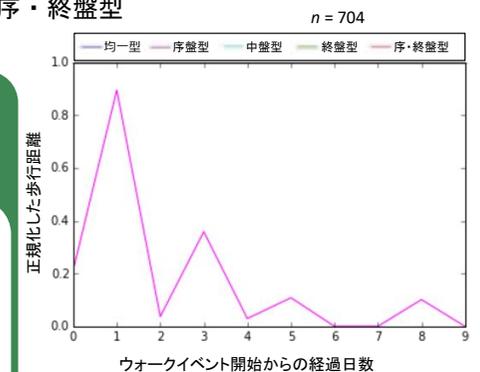
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



1 0 3 歩行に作用する要素の解析：時間帯要素

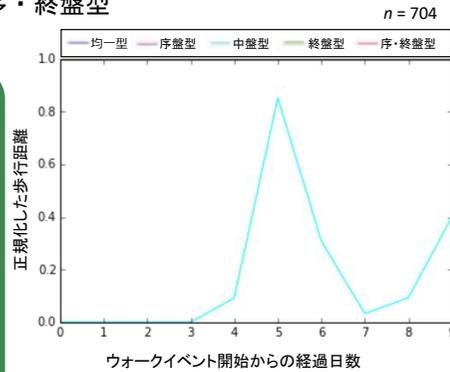
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



1 0 4 歩行に作用する要素の解析：時間帯要素

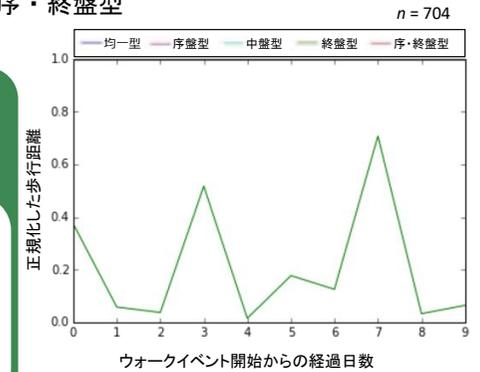
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



1 0 5 歩行に作用する要素の解析：時間帯要素

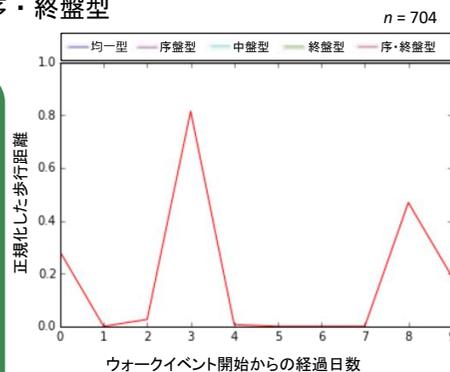
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



1 0 6 歩行に作用する要素の解析：時間帯要素

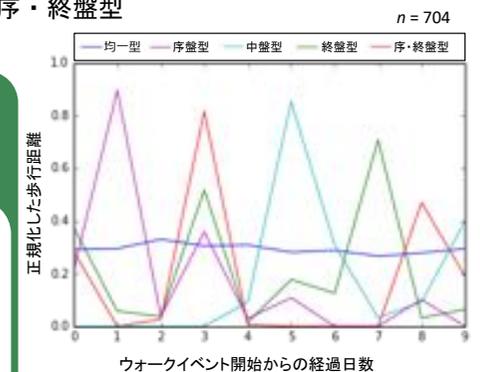
仮説：1日の中だけでなく、10日間を通してユーザによって歩行パターンが現れるのではないか

MeanShift法でユーザの歩行記録をクラスタリングすると均一型、序盤型、中盤型、終盤型、序・終盤型の5つのパターンが観測された。

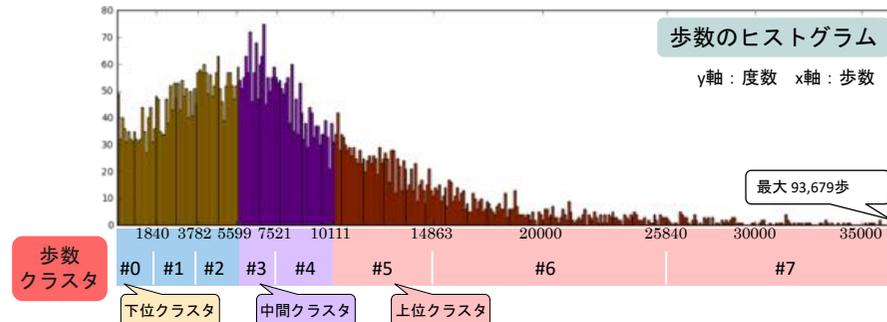
イベント期間の経過日数によってもユーザの歩行パターンが異なることが確認できた

序盤型、中盤型、終盤型は歩行量が少なくなる時期にも歩くように促すなど、

パターンに合わせて働きかけを変えることが有効なのではないか



歩数の分布に非常にばらつきがあるため、順序指標化した。



連続する2日間毎に歩数クラスターの遷移確率を算出する

歩数の遷移に各要素が与える影響を考察する



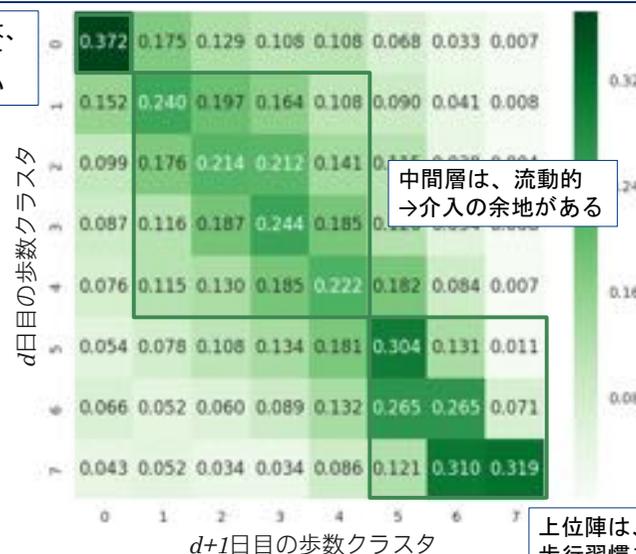
2017/02/12

ヒートマップから見る歩行者の傾向

下位陣は、安定して歩かない

中間層は、流動的 → 介入の余地がある

上位陣は、安定して歩行習慣を継続する

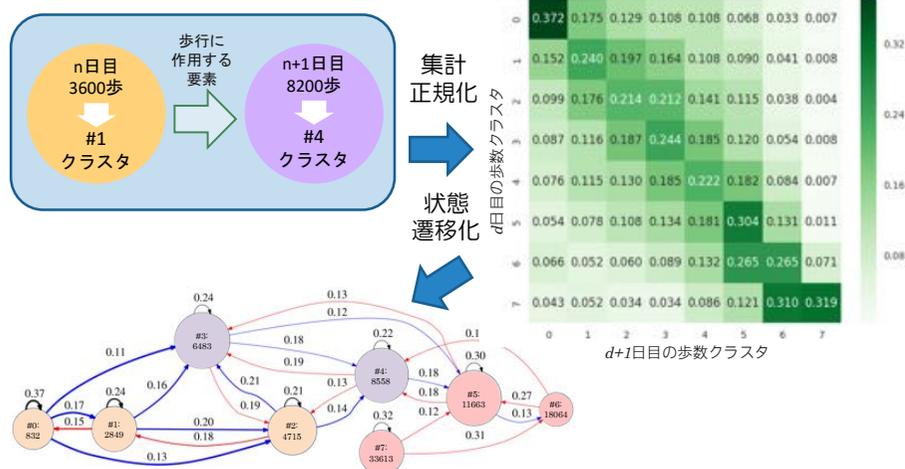


長浜健康ウォークにおける参加者の歩行状態の解析

medical informatics

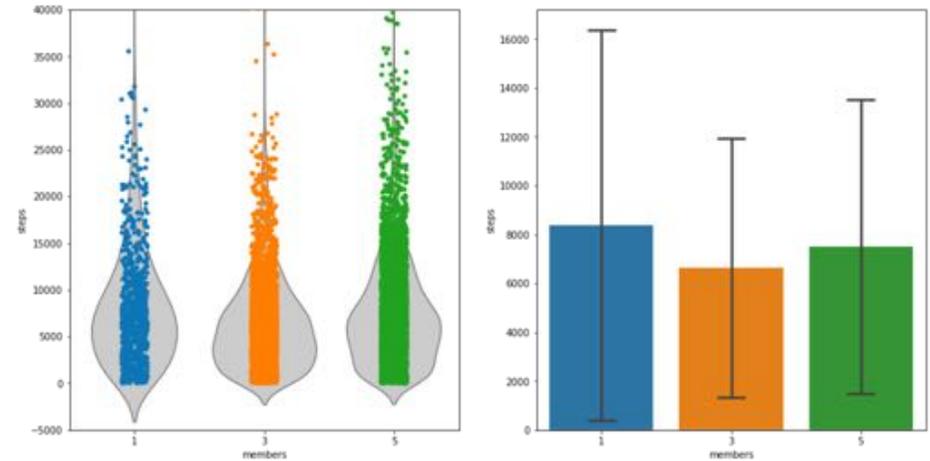


活動量の全体の遷移パターン



- 参加者はある状態に遷移するとその状態に留まる可能性が高い
- 6000-8000歩を中央値とするクラスター#3, #4を中心に、それ以下のクラスターは上昇する傾向が、それ以上のクラスターには下降する傾向が見られる

チーム人数別の散布図、平均と分散（標準偏差）



- 一人の方が平均歩数は多い
→ 踏破距離の共有のルールの利用 → **active、inactiveの群に分かれる？**
- チーム（複数名）での参加の方が分散が少ない → **何らかの協調作用がある？**

チームによる協調作用の検討

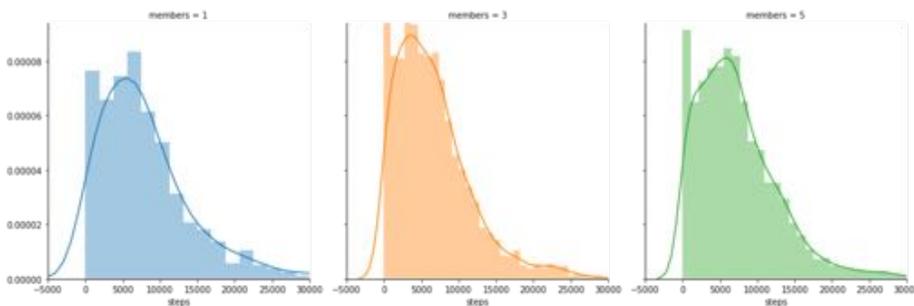
- ユーザの歩行距離を順序指標化した上で、10日間の歩行状態（シンボル）列を作り、チーム内メンバの距離を比較する。
- 比較には、文章間の距離を比較する以下の2つの指標を用いる
 - **コサイン距離**：
 - ベクトル空間モデルにおいて、文書同士を比較する。
 - **編集（レーベンシュタイン）距離**：
 - 二つの文字列がどの程度異なっているかを、1文字の挿入・削除・置換によって、一方の文字列をもう一方の文字列に変形するのに必要な手順の最小回数によって比較する。
 - チームのメンバ同士の作用は、その日中に現れるとは限らない。そのため、一日遅れてチームのメンバの影響が現れるといった作用がある程度反映できる（と考えた）

長浜健康ウォークにおける チーム参加の影響の解析

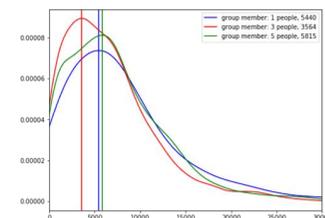
medical informatics



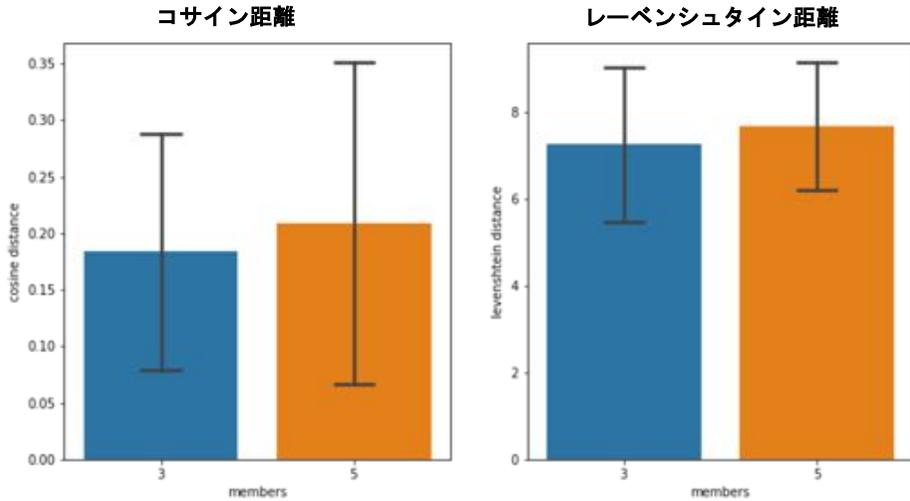
アクティブ群と非アクティブ群に分かれるのか？



カーネル密度推定で得られた分布を見ると確かにチーム人数増えるほど、分布の形が崩れる



♡ コサイン距離と編集距離

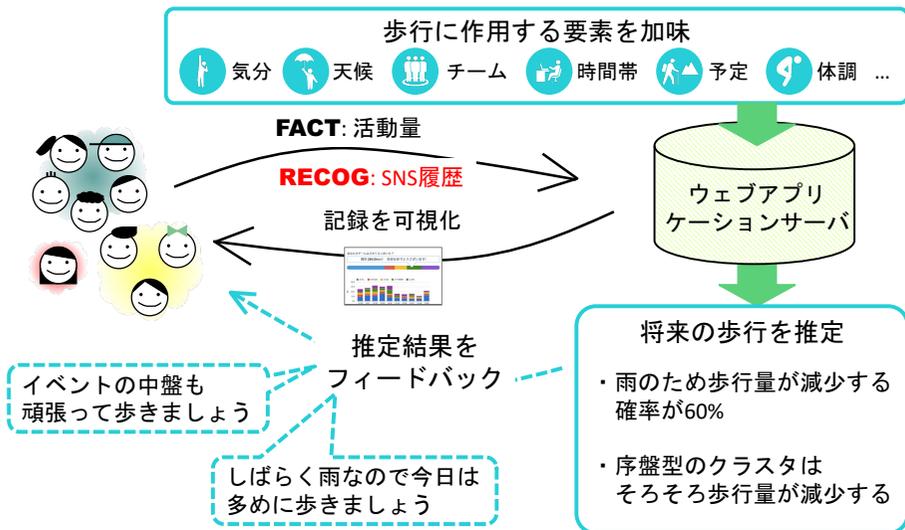


チーム人数が少ない方が、協調作用は強く働く

♡ まとめ

- 自動計測とグループ効果を利用した健康促進イベント・長浜健康ウォークを毎年やっています
- 前後合わせて3回のアンケートを実施
 - 徒歩で歩けるところには徒歩で行くようになったと答える参加者が増加
 - ほぼ毎日徒歩で移動すると答える参加者が増加
 - 仲間との励まし合いが力になった
 - しかし、習慣付けになったのは20%程度
- 期間中の参加者の行動パターンを解析
 - 天候・曜日によって行動が異なることが確認された
 - 歩く時間帯・10日間のどのタイミングで歩くかについて、いくつかのパターンが有ることが確認された
 - ユーザの歩行状態とその遷移パターンを可視化した
 - グループの効果の解析を行い、グループ人数が増えることは単純に歩数の増加に繋がらないこと、人数が増えることによって、チーム内に頑張りムラができる可能性を示唆した
- 今後は、FACTだけではなく、RECOGNITIONの収集と、それらのデータを用いた行動推薦システムに拡張していきたい

♡ てくペコの今後



♡ 機械学習を用いたEHR/PHR分析のためのデータ収集と予備検討について簡単に紹介しました

EXMR

- 京都大学医学部附属病院におけるデータウェアハウスシステム
- DWHの今後

PHR

- 長浜健康ウォーク
- 長浜健康ウォークのシステム
- 健康ウォークで収集したデータの解析

みんなで一緒に！

てくペコ 健康ウォーク