

Interação Humano Dados

Thiago Adriano Coleti

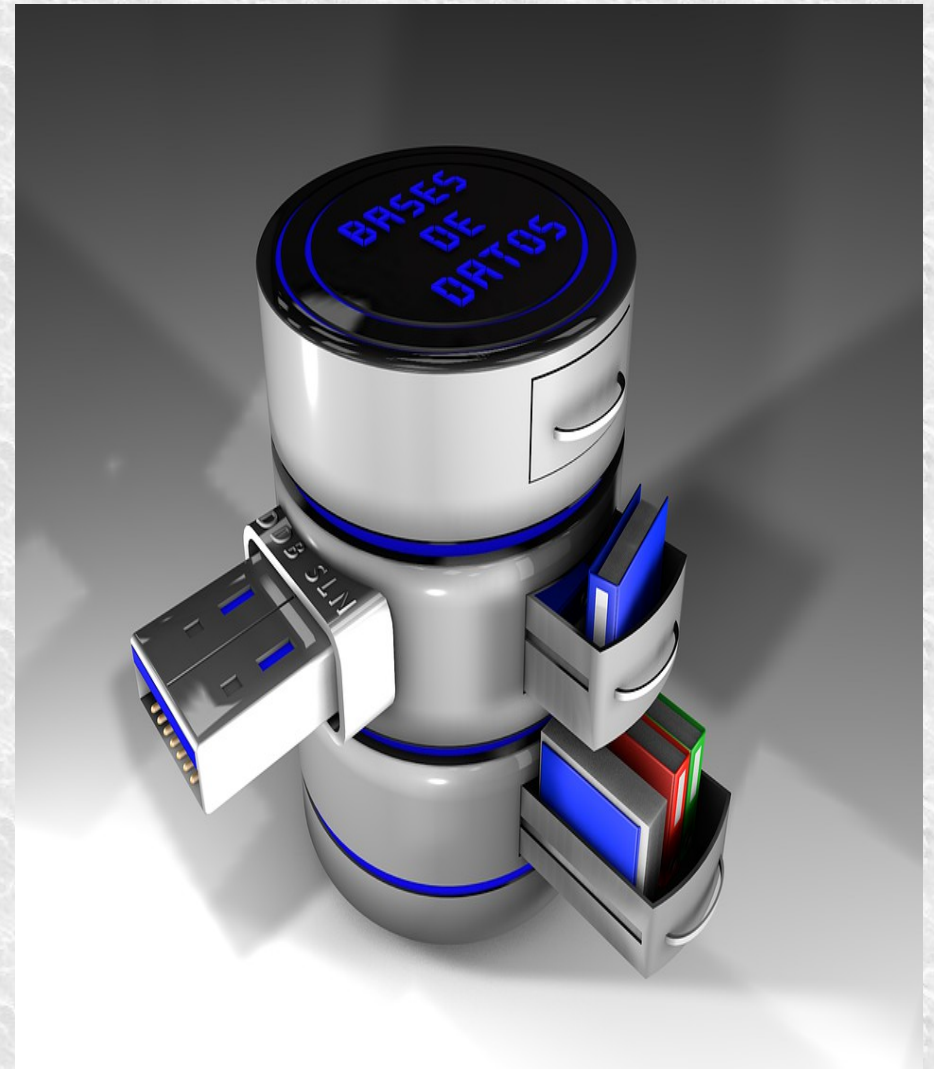
Doutorando em Engenharia Elétrica – area de Engenharia da Computação
Escola Politécnica da USP

Orientador: **Prof. Dr. Marcelo Morandini e Prof. Dr. Pedro Pizzigatti Corrêa**

“...É praticamente impossível você interagir com algum ambiente computacional sem ter algum dado pessoal produzido e/ou coletado...” [Maus, 2015]

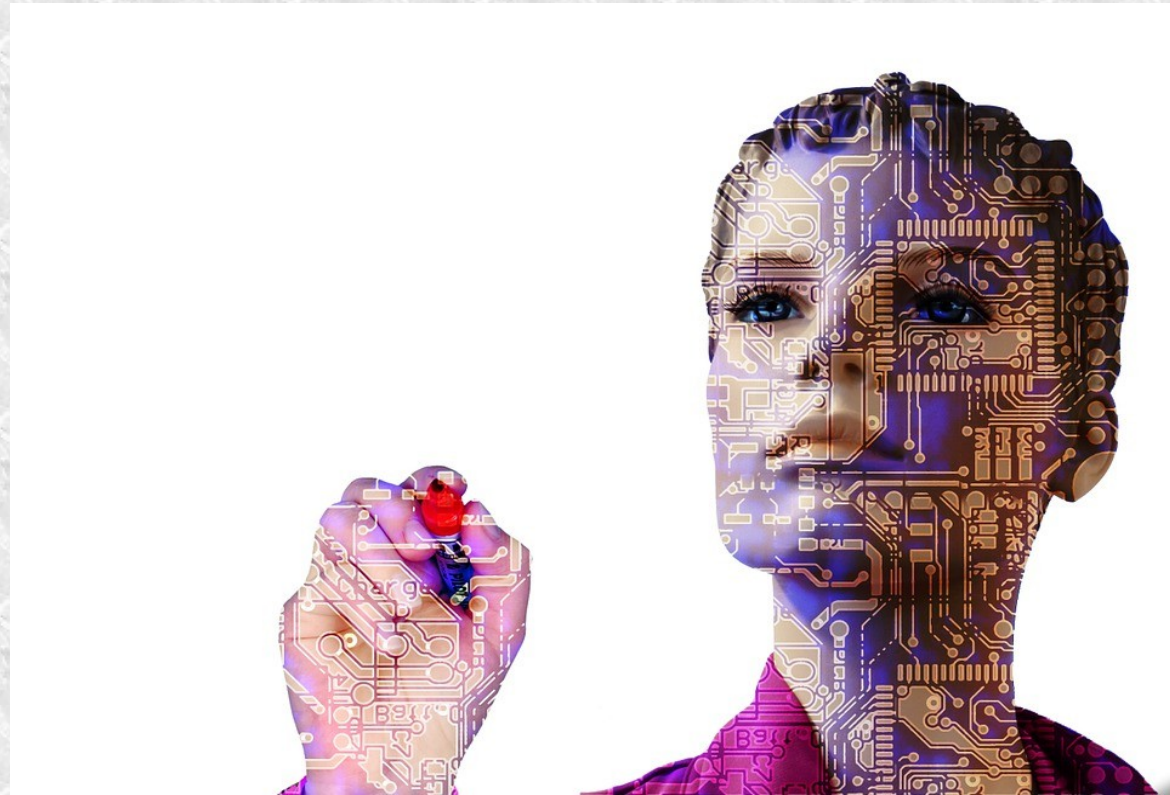
A Era do Big Data

- Vivemos a Era do Big Data no qual a vida das organizações e **das pessoas** se tornou *orientada a dados*.
- A produção de dados e o consumo de informações atinge volumes expressivos



A Era do Big Data

- Dados são produzidos por empresa, pessoas e computadores.
- Podem refletir comportamentos, preferências, ações



Vida orientada a dados



Mas o que são dados??

- São registros de eventos isolados que representam um cenários específicos. Se analisados de forma isolada não significam muita coisa, mas se combinados com outros dados podem prover informações relevantes [Stair, 2013].

O que são dados pessoais

- É um dado produzido por uma pessoa ou por um *software* (por exemplo) que pode representar um comportamento, ação ou interesse de uma pessoa.
- Pode ser classificado como
 - *My Data*
 - *Data about me*

Dados pessoais na prática

- Registrando e interagindo com redes sociais e/ou aplicativos específicos
- Tarefas diárias são realizadas refletindo determinados comportamento



A Interação Humano Dados

- O uso de dados pessoais se intensificou de uma forma que Haddadi [2013] e Mortier [2016] propõem uma área de estudos deste fenômeno.

Interação Humano Dados (IHD)

- Tem correlação com as áreas de:
 - Big Data
 - Personal Big Data
 - Interação Humano Computador

Por que precisamos de IHD ?



Por que precisamos de IHD ?



"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here."

Portas da linha-4 do metrô vão fazer reconhecimento facial de usuários



Por que precisamos de IHD

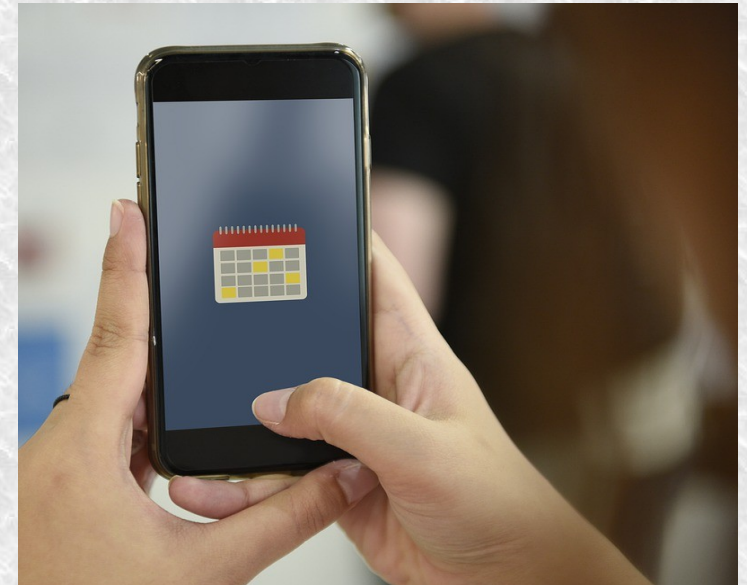
- One thing should be clear, even though we live in a world in which we share personal information more freely than in the past, we must reject the conclusion that privacy is an outmoded value ...we need it now more than ever
 - Barack Obama, President of the USA (US Consumer Privacy Bill of Rights 2012)
- Quanto mais dados revelamos, mais privacidade queremos
- Pouca clareza e dificuldades técnicas de expor o uso dos dados proporcionam as organizações a capacidade de “desviar” de eventuais leis.
- Internet das Coisas como mecanismo de produção e uso de dados pessoais.
 - Sensores
 - Muitos mecanismos pessoais na Internet

Por que precisamos da Interação Humano Dados

- Com a Interação Humano Dados acredita-se que o indivíduo poderá ter maior capacidade de compreensão sobre o uso dos seus dados.
- O indivíduo poderá agenciar e negociar junto aos processadores de dados a utilização dos mesmo.
- Privacidade, direitos e deveres sobre o uso dos dados poderão ser melhor desenvolvidos e compreendidos



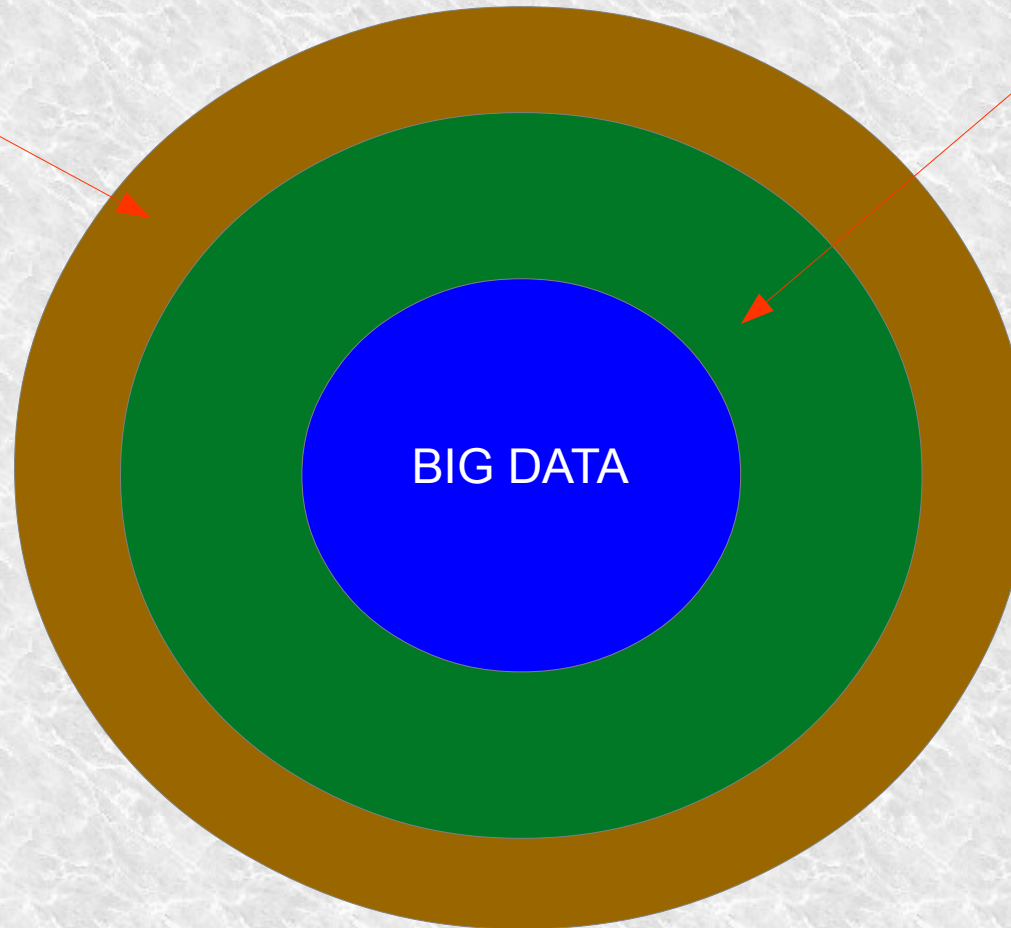
Interação Humano Dados e Personal Big Data



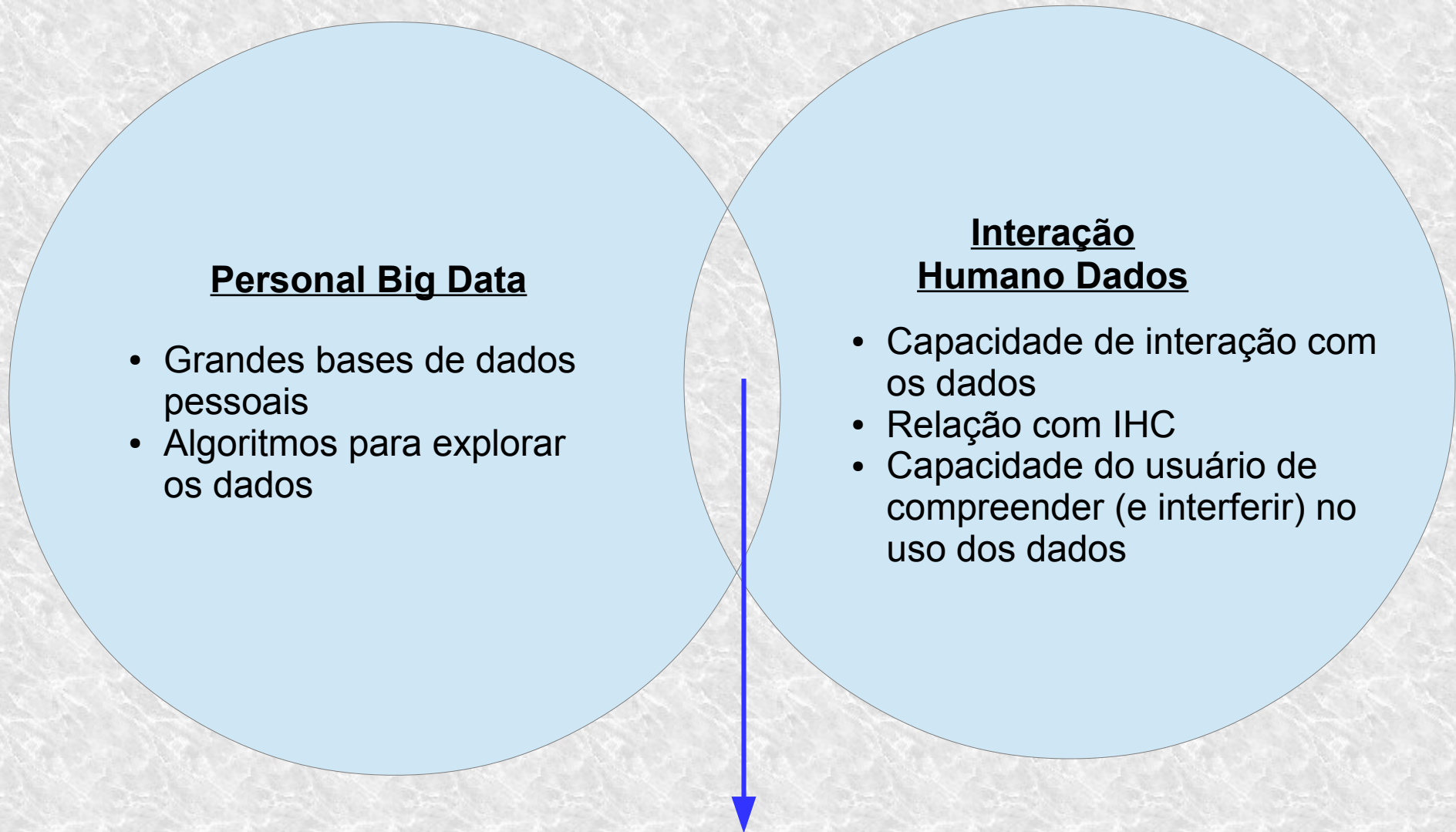
Personal Big Data e Interação Humano Dados

Human Data Interaction

Personal Big Data

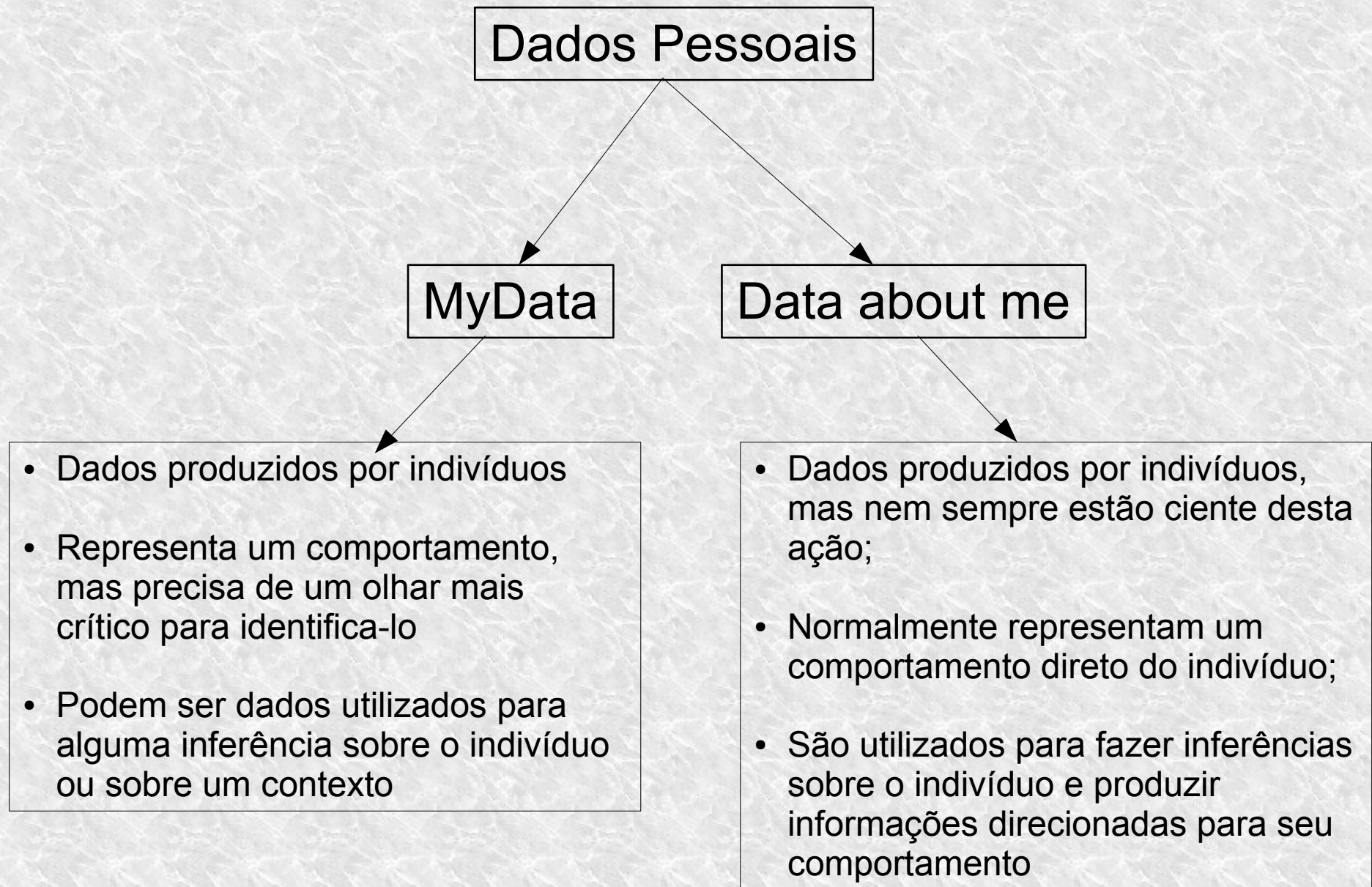


Conceitos Fortemente Relacionados



Coleta, manipulação e uso de dados pessoais em diversos contextos

Como podem ser os dados pessoais



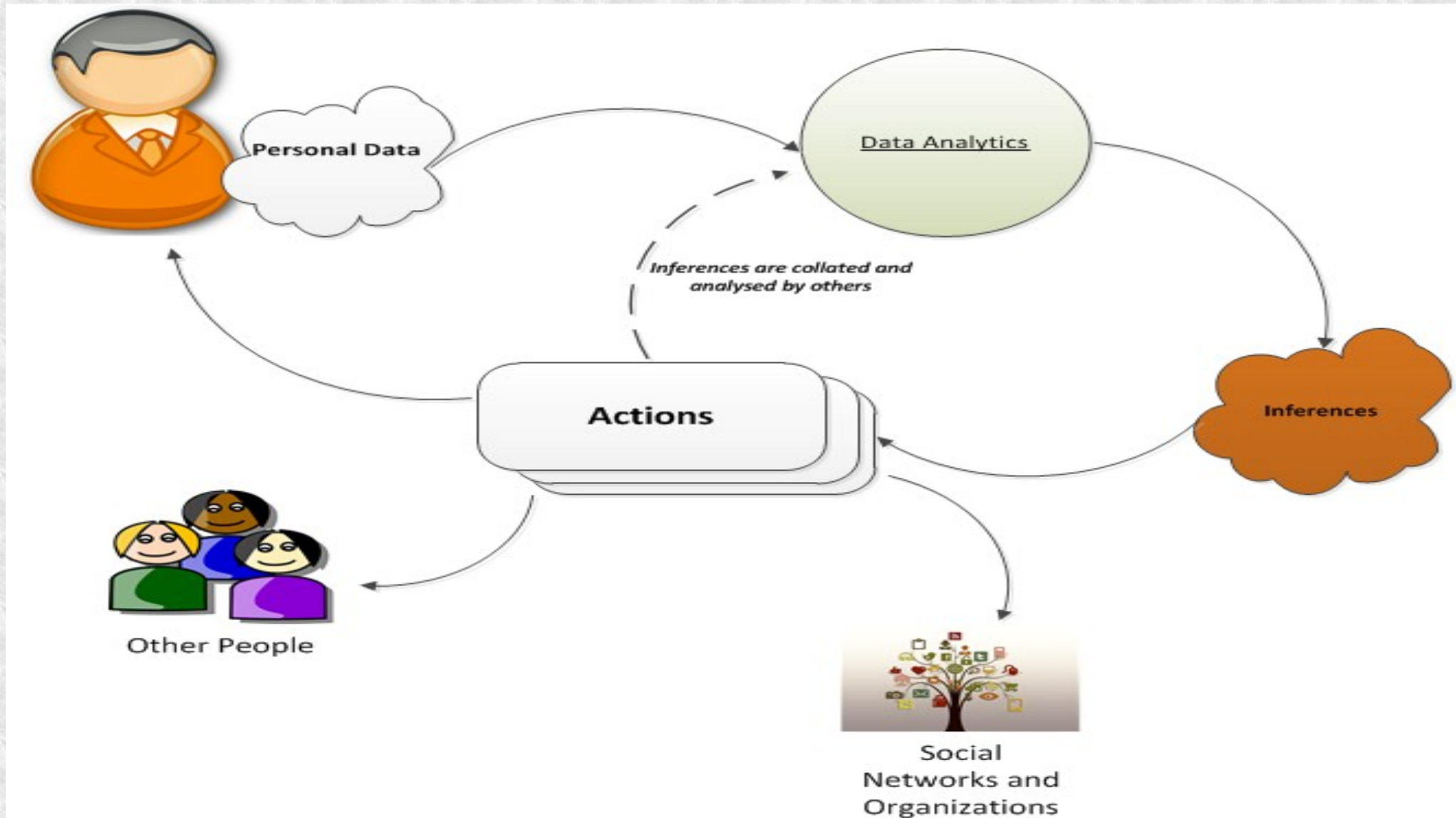
O que é Interação Humano Dados (IHD)

- IHD está relacionada a manipulação humana, análise e tomada de decisão em bases de dados grandes, complexas e desestruturadas [Elmqvist, 2011].
- IHD é a entrega de dados personalizados, dentro de contexto e compreensíveis para os usuários [Cafaro, 2012].
- IHD é promover acesso a compreensão de dados por indivíduos e como os dados afetam seu comportamento.
- IHD é a capacidade de gerenciar diversas e diferentes bases de dados pessoais e habilitar o usuário para controlar eventos de uso de dados [McAuley, Mortier e Goulding, 2011]

O que é a Interação Humano Dados

“IDH é a área que estuda fenômenos de interação entre indivíduos com grandes bases de dados contendo dados pessoais ou corporativos com objetivo de prover mecanismos para dar suporte à coleta de dados pessoais, transparência, compreensão, visualização de informações por indivíduos e tomada de decisão”

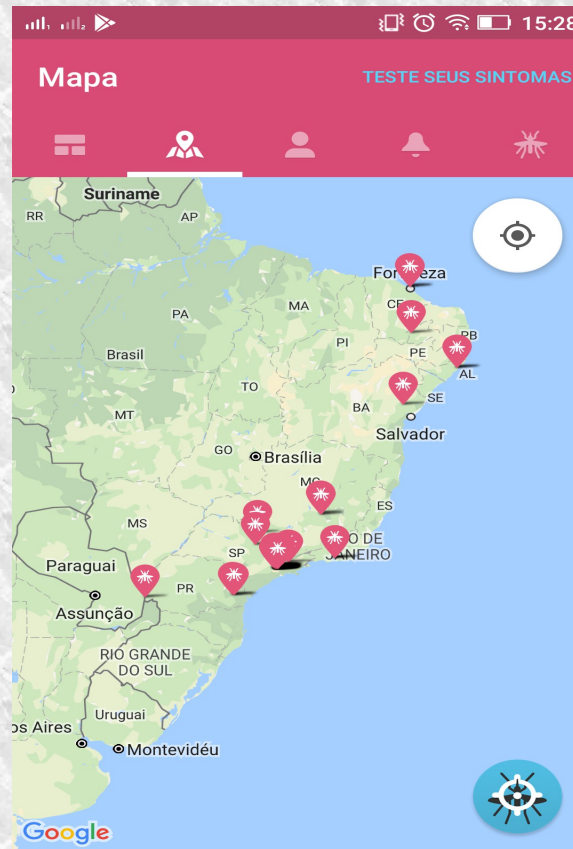
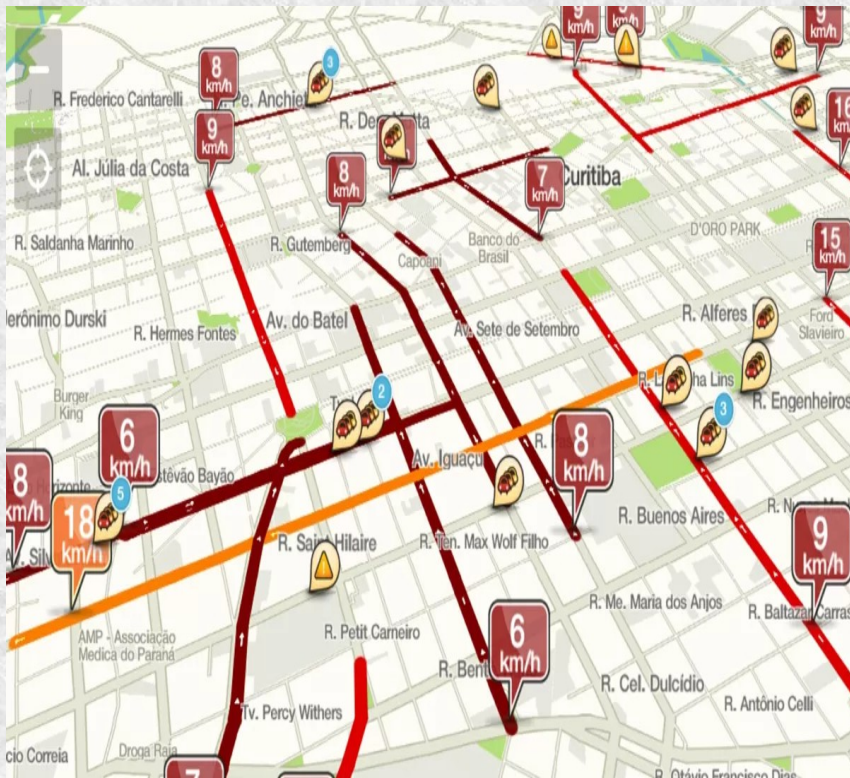
A Interação Humano Dados



Como os dados pessoais podem ser utilizados

- *O Sistema Urubu reúne dados das mais variadas fontes: usuários de rodovia, pesquisadores, concessionárias, órgãos governamentais, entre outros. Todos podem contribuir e cada informação, independente da sua origem, tem a mesma importância.*
- *O Waze usa as informações do seu percurso para calcular a velocidade média, encontrar erros e melhorar o traçado das vias, além de aprender os sentidos de condução e conversões. Você não precisa fazer nenhuma viagem especial com o Waze. Na verdade, o Waze funciona melhor nas suas rotas diárias.*
- *Fonte:*
http://cbee.ufla.br/portal/sistema_urubu/index.php
- *Fonte:*
<https://support.google.com/waze/answer/6078702?hl=pt-BR>

Como os dados pessoais podem ser utilizados



Como os dados pessoais podem ser utilizados



Exemplos anteriores foram rápidos

- Um dado pessoal pode ser utilizado nas seguintes abordagens:
 - Pessoa para própria pessoa (*Quantified Self*)
 - Pessoa para outra pessoa (rede social, Waze..)
 - Pessoa para empresas (CRM, Perfil de clientes, proposta de compras e vendas de produtos).

Pessoa → Própria pessoa

- Abordagem na qual indivíduos produzem dados sobre algum comportamento seu e utilizam para tomar decisões sobre sua vida ou para compartilhar, divulgar ou comparar com outras pessoas que fazem o mesmo.
- *Personal Informatics*
- *Quantified Self*

Personal Informatics & Quantified Self

- ***Quantified Self (QS)*** is a worldwide community of people who are interested in using personal informatics tools to track and make sense of personal data [Fan, 2013].
- ***Personal informatics*** is a class of tools that help people collect personally relevant information for the purpose of self-reflection and self-monitoring. These tools help people gain self-knowledge about one's behaviors, habits, and thoughts. It goes by other names such as living by numbers, personal analytics, quantified self, and self-tracking. [<http://www.personalinformatics.org/>]

Ambientes de *Personal Informatics*

**Feed**
VOCÊ

**Thiago Adriano Coleti**
15 de fevereiro de 2018 às 19:09

Corrida

Distância
6,8 km

Ritmo
9:31 /km

Conquistas
   **3**

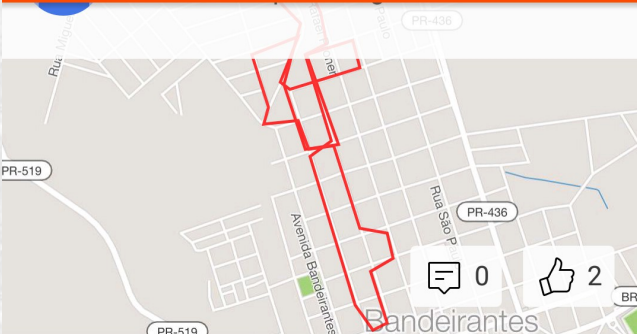


 2 deram kudos

**Thiago Adriano Coleti**

**Corrida** 



Com alguém que não gravou? **Adicionar amigo**

1:04:52
TEMPO EM MOVIMENTO

6,8 KM
DISTÂNCIA

9:31 /KM
RITMO MÉDIO

112 M
GANHO DE ELEVACÃO

Inserir

RESULTADOS

2º melhor tempo
½ milha
Mais 2 conquistas



**Corrida** 

225   

Tudo conta

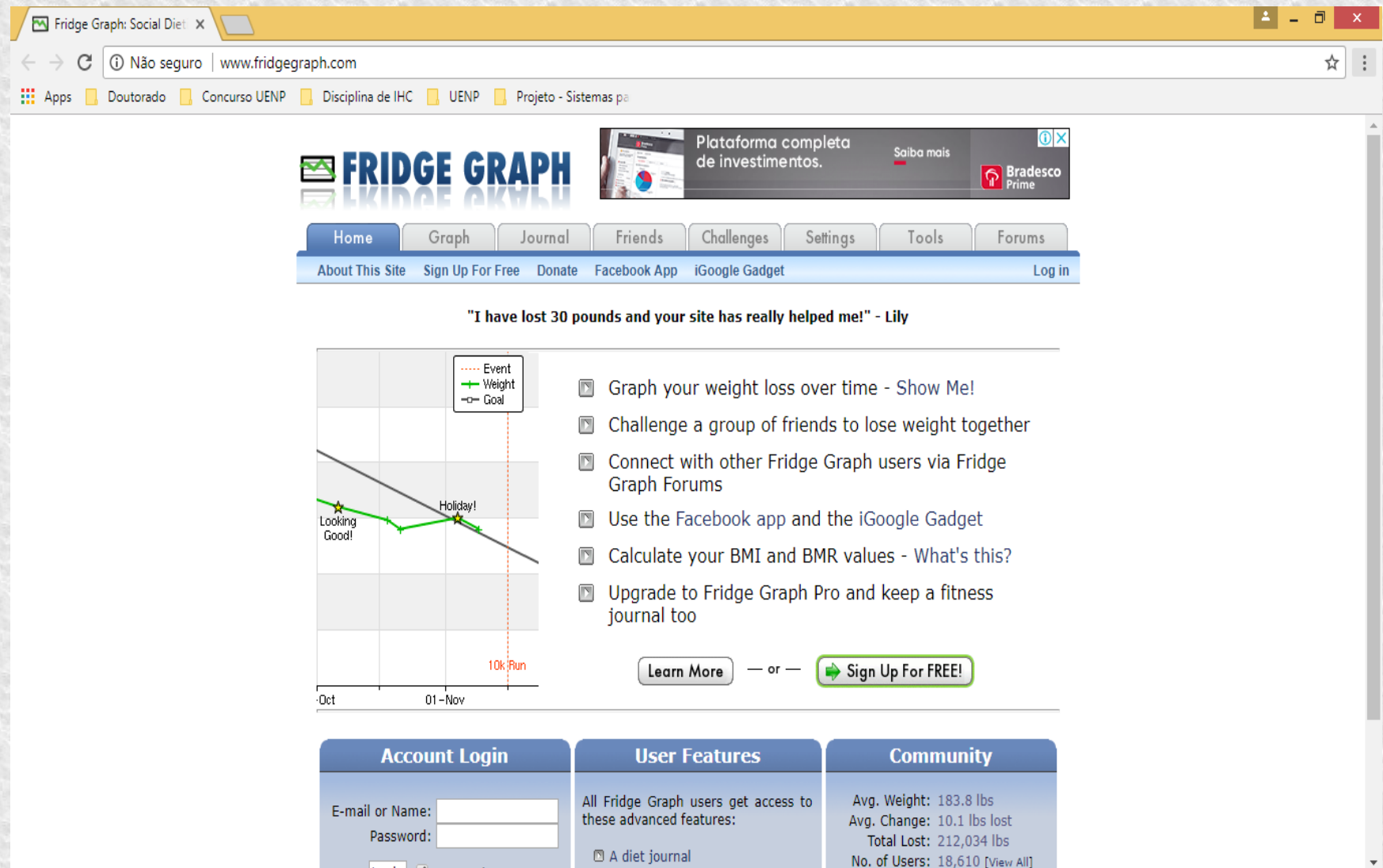
Confira sua tendência de desenvolvimento da forma física com o Esforço relativo e compare suas atividades entre esportes e até mesmo com outros atletas.

Saiba mais >

KM	RITMO		ELEV
1	10:42		-12 m
2	7:41		10 m
3	10:49		1 m
4	8:25		-23 m
5	10:49		11 m
6	7:13		-4 m
0,8	11:13		18 m

Ambientes de *Personal Informatics*



Fonte: <http://www.fridgegraph.com/>

Ambientes de *Personal Informatics*

SugarStats.com - Simple

Seguro | https://sugarstats.com

Apps | Doutorado | Concurso UENP | Disciplina de IHC | UENP | Projeto - Sistemas pa

SugarStats

Home | Tour | My Account

SugarStats is a free blood sugar tracking and online diabetes management tool.

Every diabetic knows that better control means better health, lower medication intake and thus lower medication costs.

With SugarStats you will be able to:

- **Visualize your progress** with easy to read graphs and trends
- **Meet other diabetics** with community groups and discussion forums
- **Share your statistics** with your family, friends or physicians
- **Get rid of that pen & paper log** It's all online!
- **Drill down** into specific timeframes to get a clear picture of your stats over time

SugarStats gives you a simple and easy-to-use interface to input and access your data from home, school, work, when you're on the road or just about anywhere at all. **All you need is a web browser** and the desire to take control of your diabetes.

Register or Take a tour

Signing up takes less than a minute.

SugarStats

Home | Graphs | Stats | Friends | Messages | Journal | Logs | Store

Dashboard

Create New: Stats for Today | Journal Entry | Message | HbA1c Entry | Exam Entry | Doctor Checkup

Recent 5 Day Sugar Graphs

Sugar Entries Graph

Top Tags - Sugar Averages

Last 5 Day Averages

Last 5 Messages

Last 5 HbA1c Entries

Last 5 Journal Entries

Friend Comments

Quick Links

Dashboard

Today

Yesterday

Last 7 Day Stats

Last 7 Day Graphs

Last 30 Day Graphs

View Your Shared Profile

Graphs and Trends

Quick and Simple

Easy to Use

Share your Statistics

Fonte: <https://sugarstats.com/>

Ambientes de *Personal Informatics*

Where Do You Go - crea X

www.wheredoyougo.net

Apps Doutorado Concurso UENP Disciplina de IHC UENP Projeto - Sistemas pa

Where Do You Go

Login with Google!

Where do yo... 4:33

where am i?

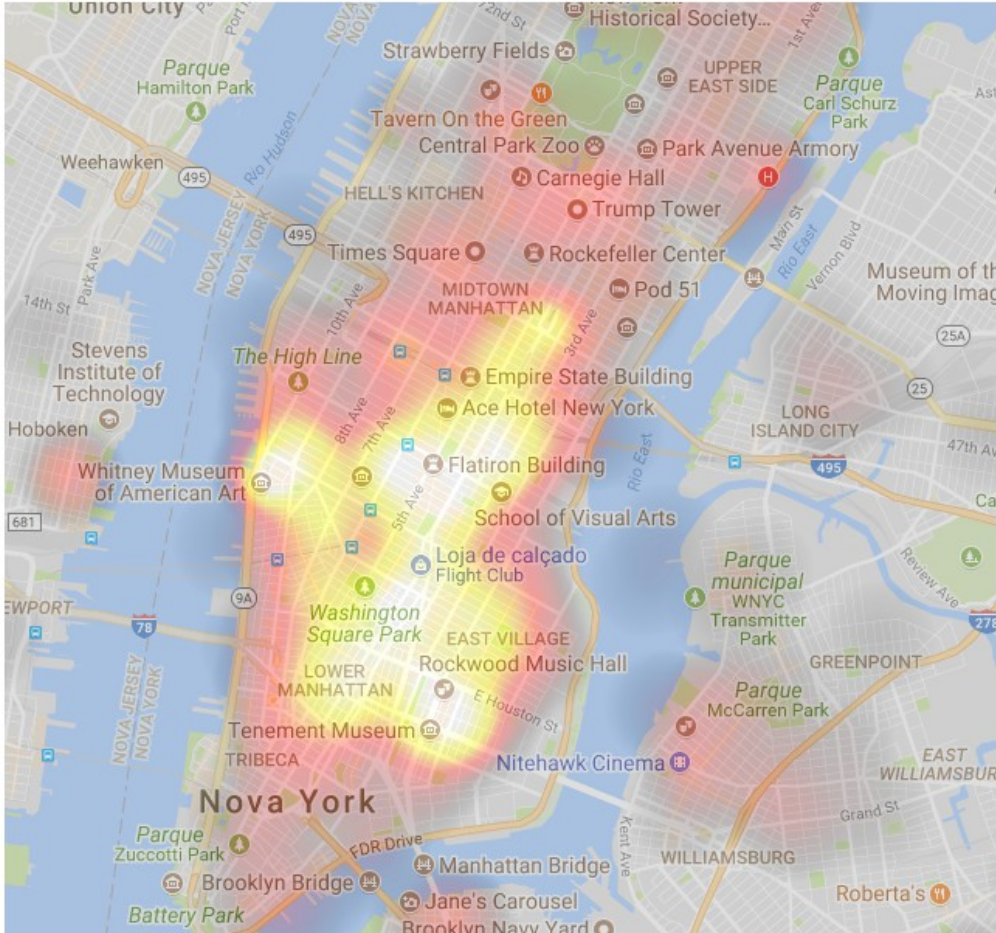
Use this site to create and share a heat map of your Foursquare check-ins. Login with your Google account first, and then authenticate with Foursquare - don't worry, we don't see either password.

foursquare says:

"We're all about helping you find new ways to explore the city. We'll help you meet up with your friends and let you earn points and unlock badges for discovering new places, doing new things and meeting new people."

a heat map? huh?

A heat map is a way to visualize location data using color on a map. Denser concentrations of color indicate more Foursquare check-ins at a venue at that particular location.

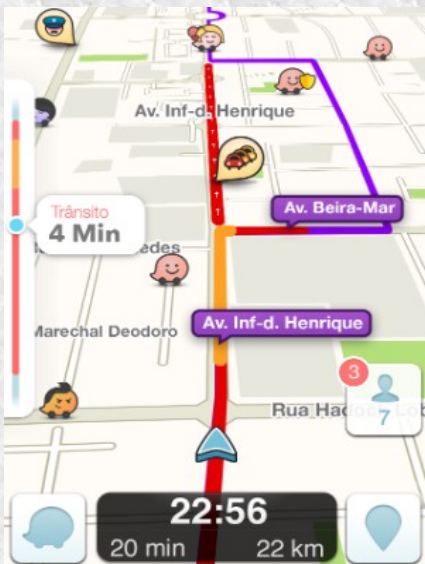
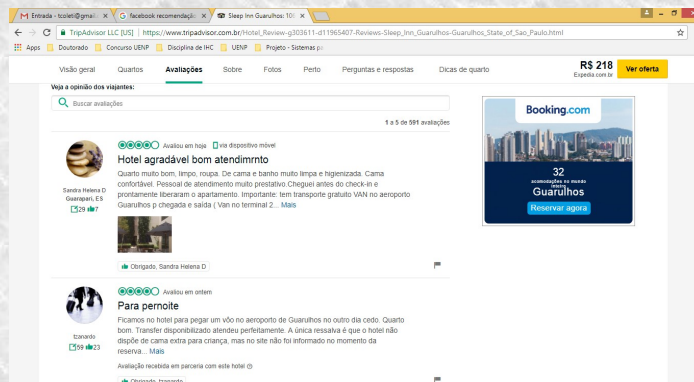
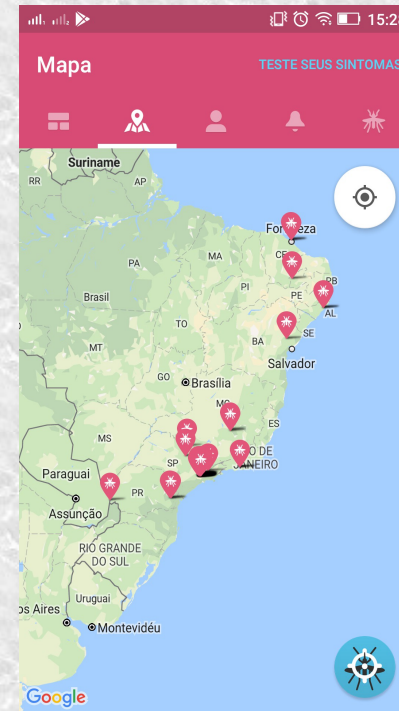


Fonte: <http://www.wheredoyougo.net/>

Pessoa → Outra Pessoa

- São sistemas que atuam sobre dados pessoais a fim de processá-los e proporcionar informações para que outras pessoas possam tomar decisões sobre suas ações.
- Podem ser dados coletados automaticamente
- Podem ser dados gerados pelos usuários

Pessoa → Outra Pessoa



Avaliações (675)

Faça uma avaliação

Pontuação dos viajantes

☐ Excelente		202
☐ Muito bom		240
☐ Razoável		84
☐ Ruim		29
☐ Horrível		36

Tipo de viajante

- ☐ Famílias
- ☐ Romântico
- ☐ Sozinho
- ☐ Negócios
- ☐ Amigos

Época do ano

- ☐ Mar-Mai
- ☐ Jun-Ago
- ☐ Set-Nov
- ☐ Dez-Fev

Idioma

- ☐ Todos os idiomas
- ☒ Português (591)
- ☐ Espanhol (52)
- ☐ Inglês (28)
- [Mais idiomas](#)

Veja a opinião dos viajantes:

Pessoas → Empresas

- Abordagem na qual as empresas coletam dados do usuário para gerar informações que possam ser aplicadas para identificar perfil de compras, sugerir produtos, serviços e outros.



Hotéis com até 25% de desconto
<https://hotels.latam.com/>
Cada R\$2 gastos = 1 ponto MULTIPLUS.



MongoDB Atlas: Try Your Free Cluster
mongodb.com
MongoDB Atlas is the best way to run MongoDB in the cloud and the easiest way to try the n...

Pessoas → Empresas

- Abordagem na qual as empresas coletam dados do usuário para gerar informações que possam ser aplicadas para identificar perfil de compras, sugerir produtos, serviços e outros.

The screenshot displays the Decolar.com website interface. At the top, there's a navigation bar with links for 'Apps', 'Doutorado', 'Concurso UENP', 'Disciplina de IHC', 'UENP', and 'Projeto - Sistemas pa'. Below this is a blue banner with the text 'Torne sua viagem realidade! Ative as notificações e inspire-se com nossas ofertas e novidades.' and an 'Ativar' button. The main header features the 'decolar.com' logo and links for 'Iniciar Sessão Minha Conta', 'Minha Reserva', and 'Ajuda'. A secondary navigation bar lists various travel services: 'Hospedagens', 'Passagens', 'Pacotes', 'Ingressos', 'Carros', 'Disney', 'Seguros', 'Transfers', 'Ônibus', 'Aluguéis', and 'Cruzeiros'. The main content area is divided into three sections. The left section, titled 'Encontre seu Voo', includes filters for 'Ida e Volta', 'Só ida', and 'Multidestino'. It shows search criteria: 'Origem: Londrina, Brasil', 'Destino: São Paulo, Brasil', 'Datas: Seg, 23 de Abr 2018 to Ter, 24 de Abr 2018', and 'Passageiros e classe: 1'. A 'Procurar' button is at the bottom. The middle section features a hotel offer for 'Hotel Sleep Inn Guarulhos' with a '11% DESCONTO' and a price of 'R\$ 196'. The right section is a promotional banner for 'OFERTAS DO DIA' with the text 'VOE MAIS BARATO' and 'PARIS A PARTIR DE R\$2348'.

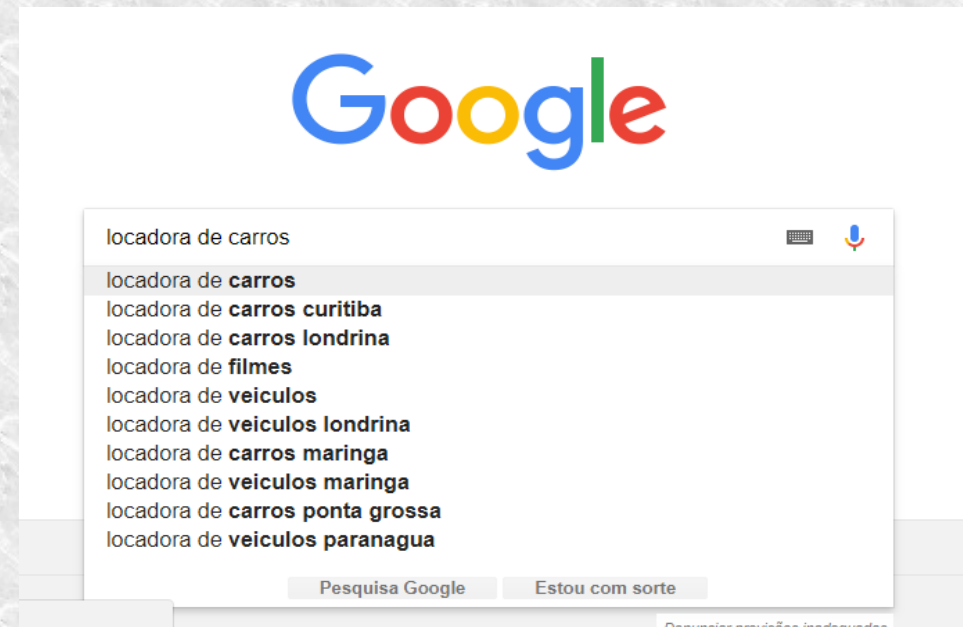
Viajantes como você compraram estas ofertas em Guarulhos

O Volume e Organização dos Dados Pessoais

- Uma massa de dados pessoais pode ser classificado de acordo com 3 níveis:
 - *Data Trail*
 - *Data Smog*
 - *Big Data*

Data Trail

- É um registro particular de uma atividade de um indivíduo em um dispositivo que coleta dados.
- Pode ser uma máquina de cartão de crédito, uma rede social, um telefone, um e-mail, um registro de busca na Internet



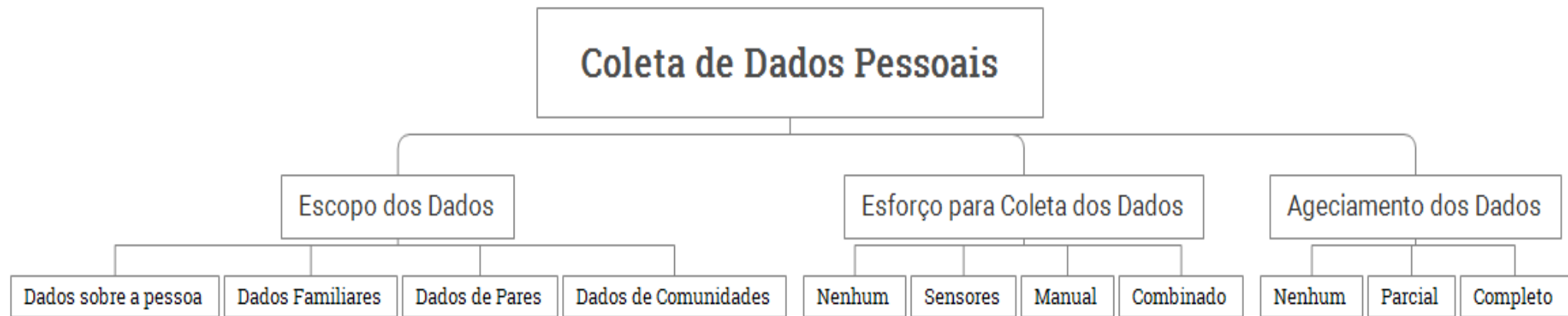
Data Smog

- É uma massa de grande volume de dados pessoais geralmente proveniente da Internet e de sistemas de processamento de dados pessoais.
-
- A massa de dados costuma implicar em dificuldades para separar dados relevantes de dados insignificantes.

Big Data

- Conjunto de dados pessoais (grande volume de dados) que na medida da sua maniputalação começa a apresentar desafios, significados, possibilidades de melhorias e refinamento.
- Envolvem áreas de Computação, Ciência dos Dados, Sociologia e Psicologia.

As dimensões da Coleta/Produção de Dados Pessoais



Fonte: Huang. 2014

Escopo dos dados

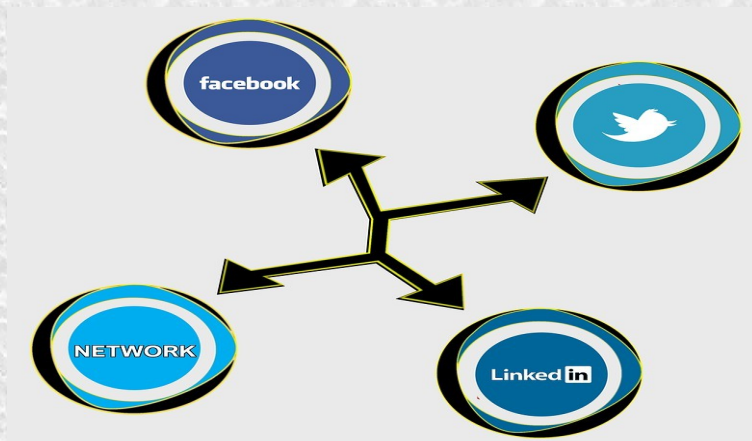
Pessoas: Dados do Monitoramento da Saúde



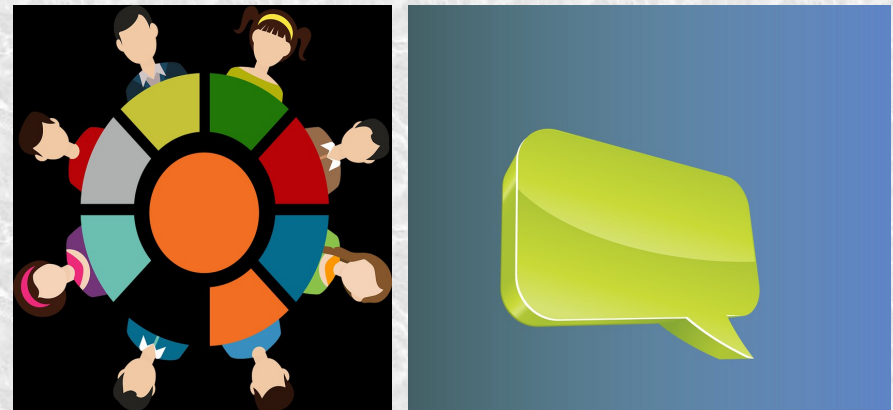
Familia: Dados sobre viagens ou comportamento na casa (sensores)



Dados de Pares: Relacionamentos e interesses comuns em redes sociais



Dados de Comunidade: Conversas em apps de bate papo, comunidades de prática



Esforço de Coleta

Nenhum: Buscas em sites de consulta



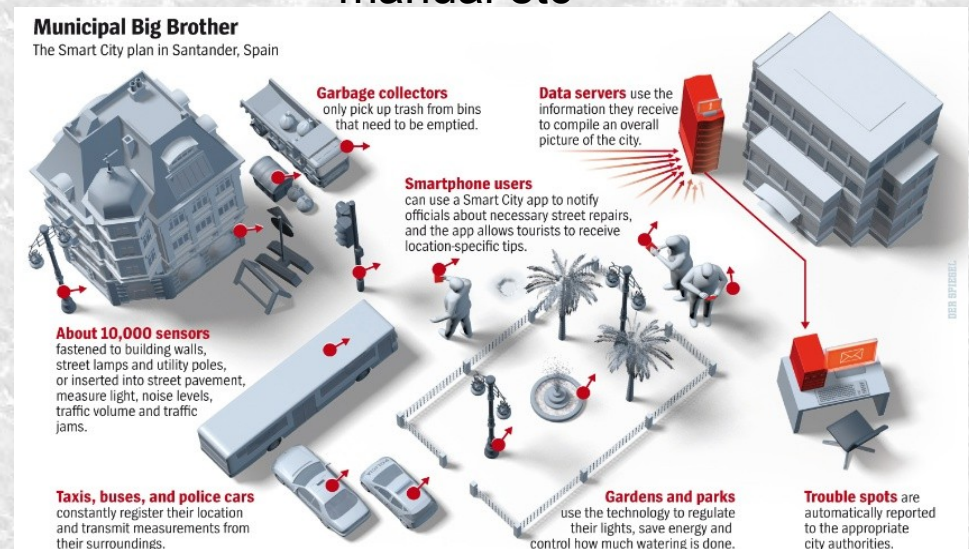
Sensores/Apps: Uso de sensores pouco intrusivos em locais diversos ou em aplicativos móveis.



Manual: Geração de dados por indivíduos de forma manual



Mix: Combinação de geração por sensores, manual etc

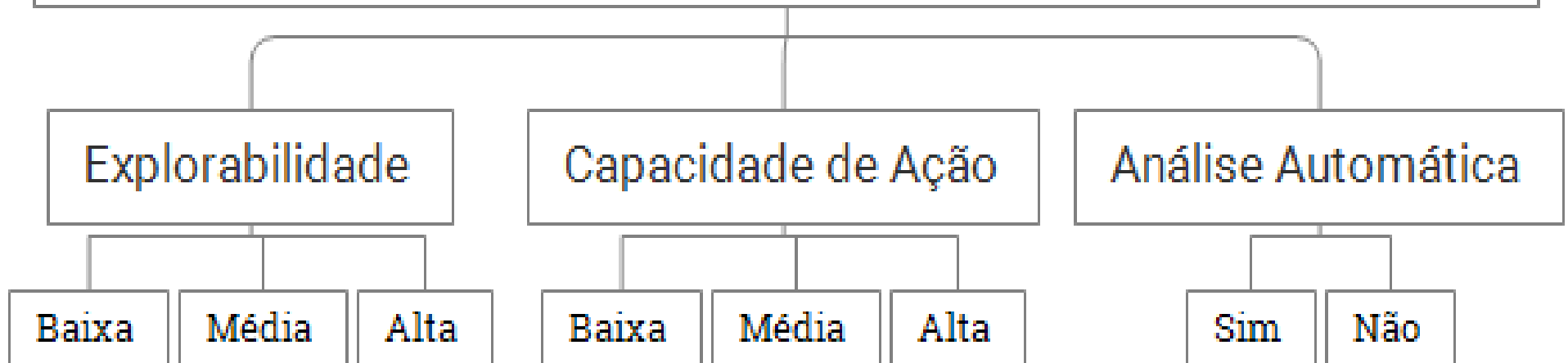


Agenciamento

- **Nenhum:** dados de conversas em apps de bate papo. O usuário não tem (ou tem pouco) poder de gerenciamento dos seus dados
- **Parcial:** Tem controle de parte do agenciamento e pode indicar o que pode ou não ser coletado
- **Total:** tem total controle sobre a coleta de dados → aplicativos infantis e fotos de pesquisas científicas

Dimensões da Visualização de Dados

Visualização de Dados para Indivíduos

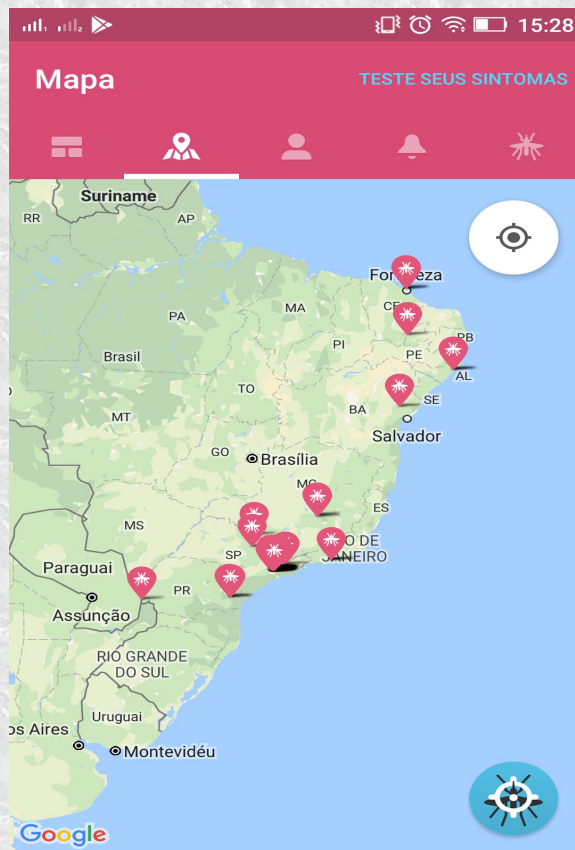


Explorabilidade

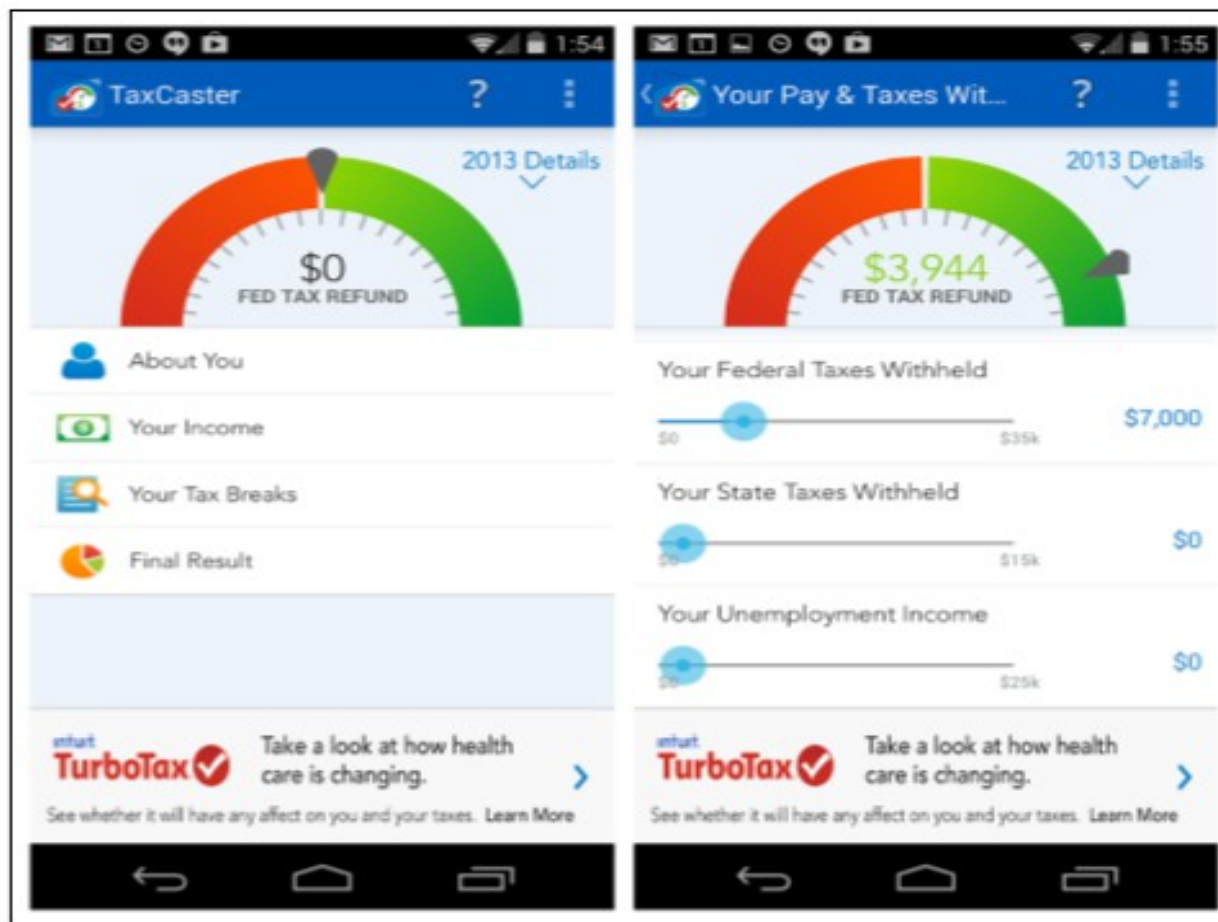
Baixa



Média



Explorabilidade



Fonte: Neil, 2014

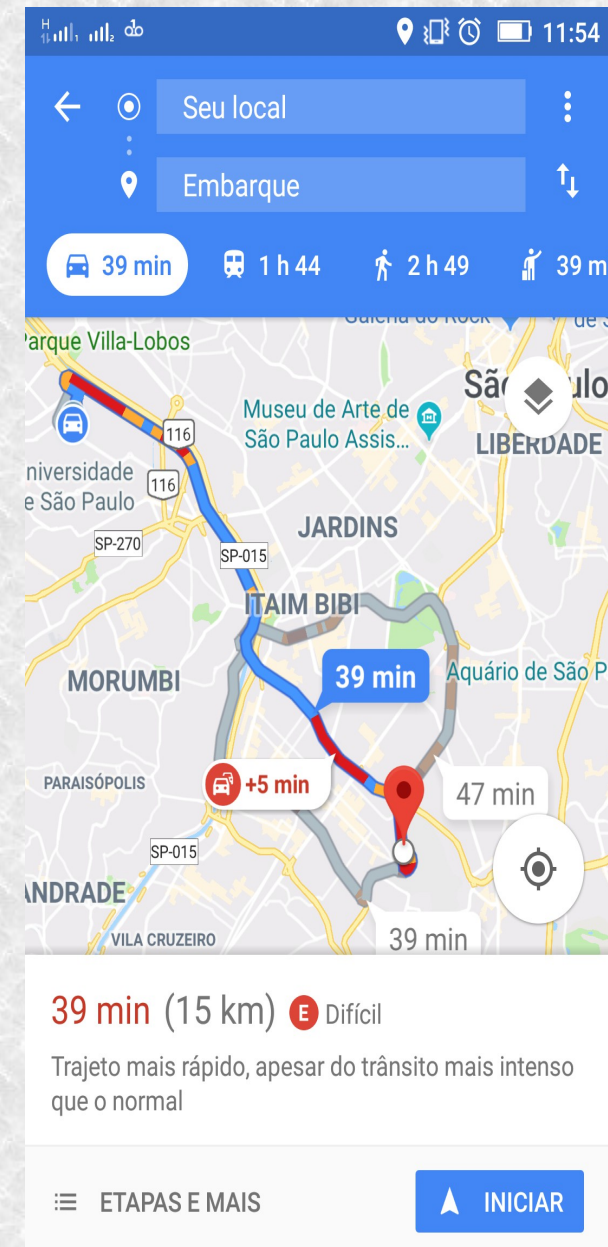
Capacidade de ação

Alta



Baixa

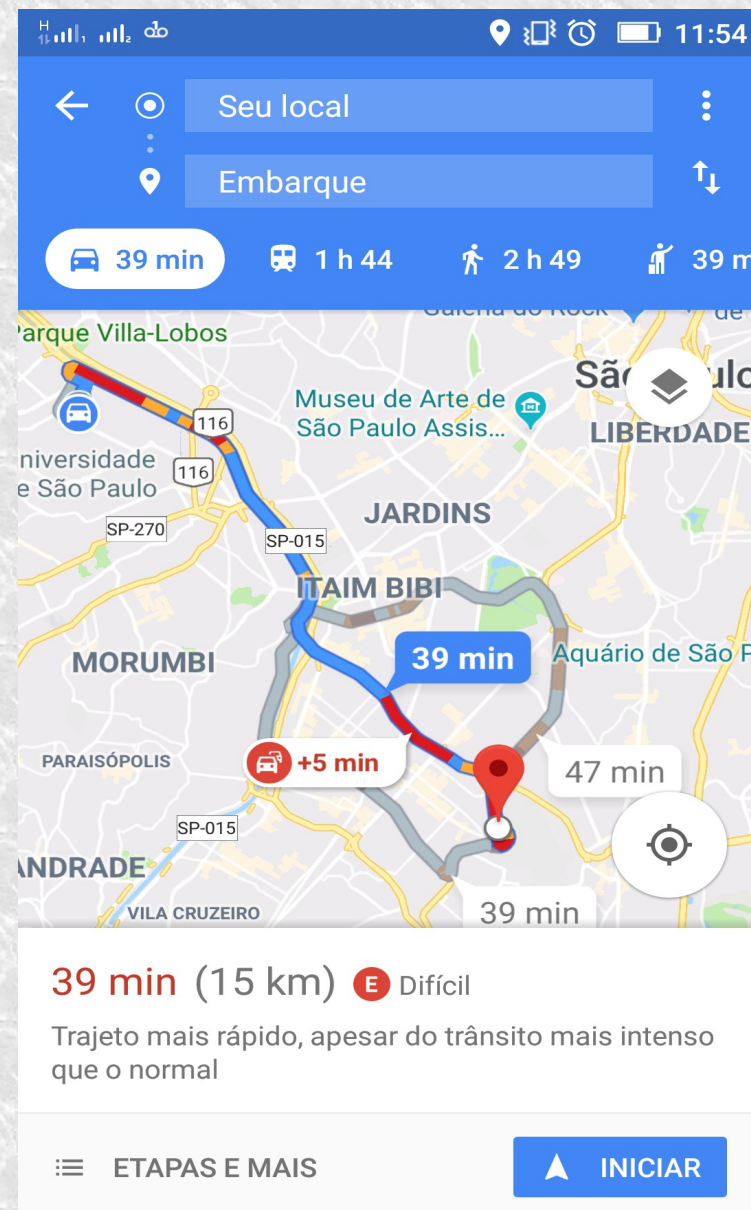
Média



Análise Automática



Sim



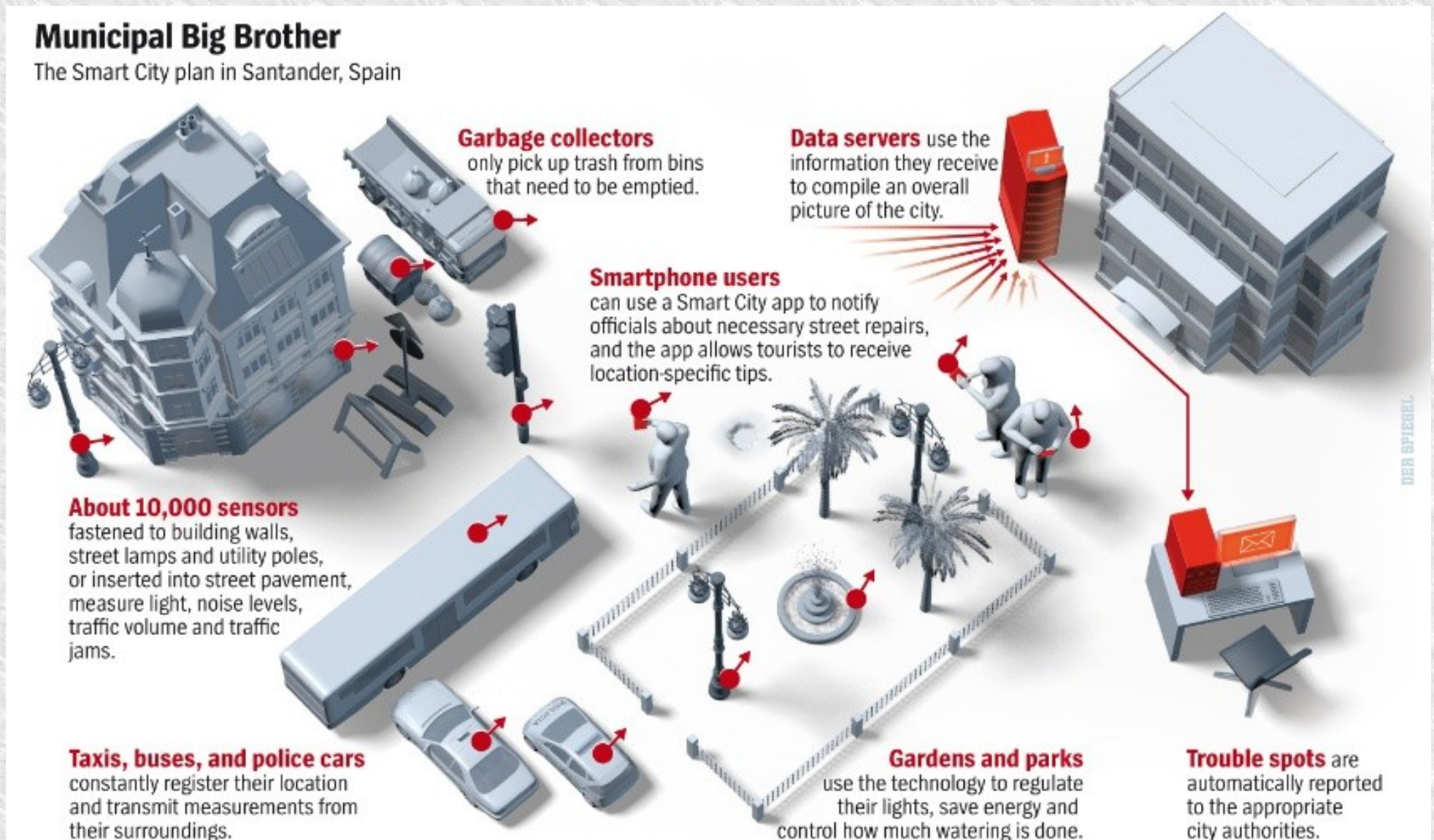
Não

A Internet das Coisas como forte elemento para a Interação Humano Dados

O que pode acontecer

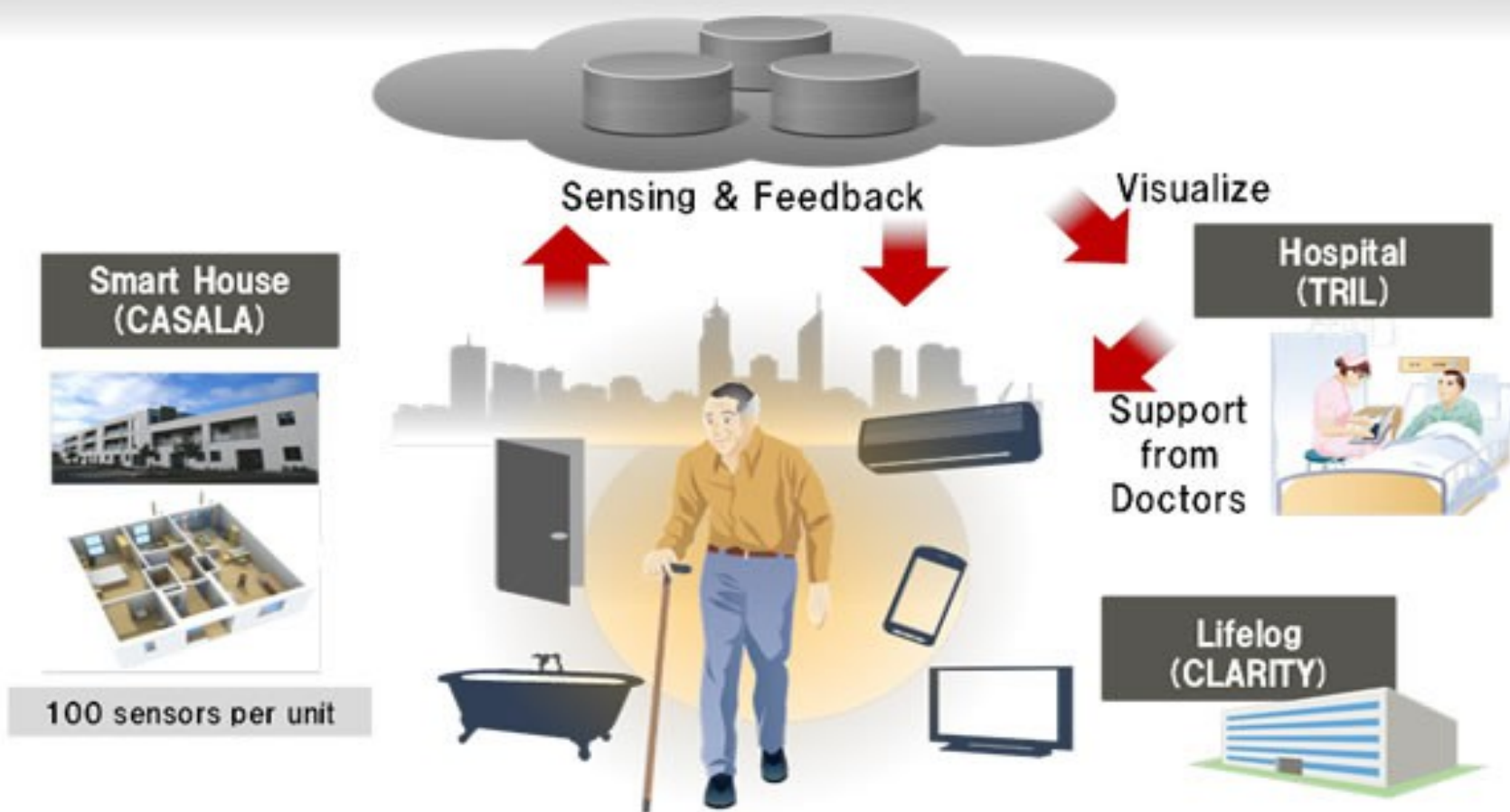
- Uma inimaginável possibilidades de redes de sensores ligados aos mais diversos componentes como carros, casas, pessoas, aparelhos domésticos, cartões etc pode produzir uma quantidade de dados inimaginável.
-
- As pessoas ficariam totalmente conectadas e cientes do estado de outras entidades, mas forneceria seus dados em troca.

A cidade inteligente Santander



Atendimento Hospitalar Inteligente

- Joint project with hospital and smart house
- Monitoring and assisting independent living for senior citizens and patients
- Application of various sensing technologies on Fujitsu Laboratories' processing platform



A nossa privacidade diante os sensores e câmeras

- Como nos prevenirmos e compreendermos como dados a nosso respeito podem ser coletados na rua??

O que acontece com os dados dessas câmeras ??



Uniformes inteligentes





Desafios da Interação Humano Dados

Por ser uma área de exploração relativamente jovem, a IHD provê uma ampla gama de desafios técnicos e científicos

Desafios

- Scheiner alega que o limite para exploração dos dados pessoais é o limite que a Internet que a Computação impor.
-
- Não há um *deadline* em relação ao que acredita-se poder chegar com dados.

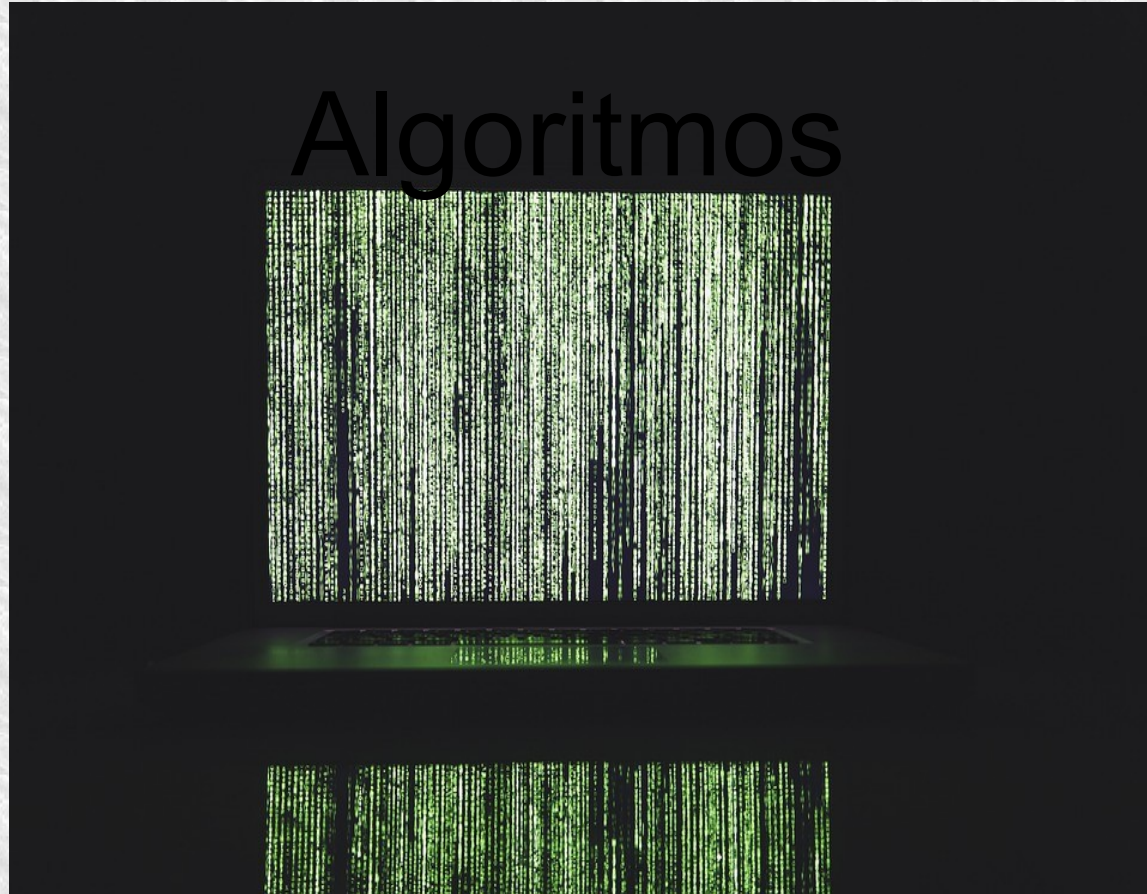
Infraestrutura

- Hardware
- Internet
- Banco de Dados
- Poder de Processamento



Todos podem necessitar de melhorias no decorrer do crescimento do uso dos dados

Algoritmos



- Novos dados = novas possibilidade de informações
- Novas necessidade de algoritmos para minerar, classificar, processar etc tantos dados.

Transparências de ações

- Um dos pontos mais questionados sobre Interação Humano Dados é o fato das ações que ocorrer nos dados são pouco visíveis, compreensíveis e transparentes para o usuário.
- Necessita-se de mecanismos com o objetivo de tornar mais claro o que acontece com os dados.



Interação Humano Computador



Questões Legais / Ética / Responsabilidade

- Quanto mais dados proporcionamos, mais sigilo queremos sobre os dados.
- Testemunhamos casos de grandes empresas que fazem mal uso dos dados pessoais.
- Leis foram (e estão sendo) criadas.
- Desafio de proporcionar a adaptação dos sistemas as leis

Considerações Finais

- IHD é uma área que estuda:
 - A interação de indivíduos com grandes bases de dados
 - A capacidade de produzir e consumir dados pessoais por parte dos indivíduos e os impactos destes dados.
 - Area nova, muita a aprender e desenvolver
 - A interação humano dados lança desafios envolvendo questões técnicas e também humanas em relação ao uso dos dados.

Perguntas???

Proposta de um exercício

- Em grupo
- Escolha um website, aplicativo ou plataforma na qual um indivíduo possa interagir e eventualmente ter um dados produzido e/ou coletado.
- Verifique como esta plataforma informa ao usuário sobre o processo de coleta de dados.
 - Qual dado será registrado
 - Como?
 - Até quando, etc...

Proposta de um exercício

- Verifique também, quais seriam os possíveis dados coletados pela plataforma e como a mesma poderia usufruir deste dado.

Obrigado

Thiago Adriano Coleti

thiagocoleti@usp.br / thiago.coleti@uenp.edu.br

Perguntas....