

# Computação por Humanos na Perspectiva do Engajamento e Credibilidade de Seres Humanos e da Replicação de Tarefas

Lesandro Ponciano, Francisco Brasileiro

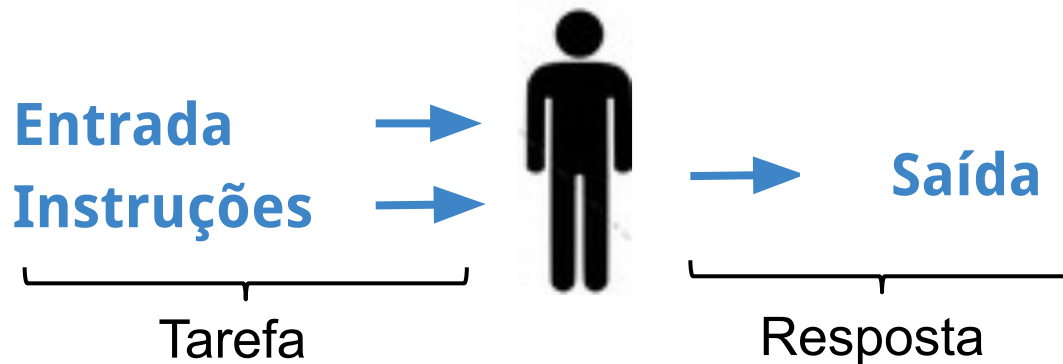


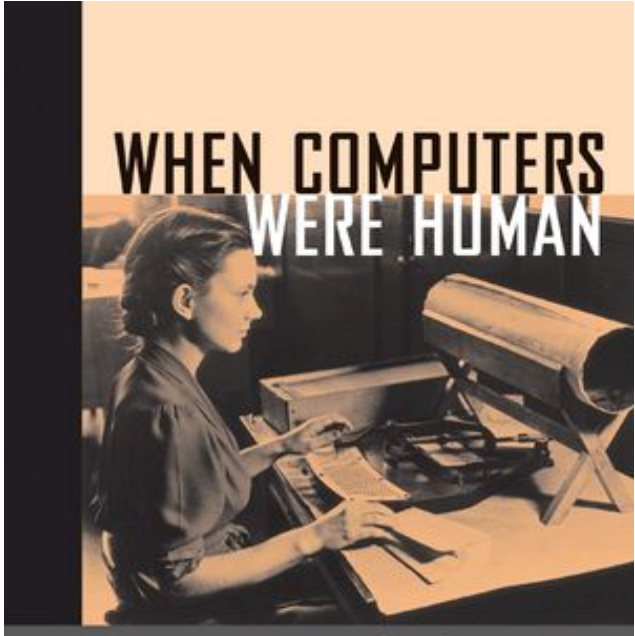
# Computação

- ‘Processo de mapear **entrada** em **saída** usando um conjunto finito de **instruções**’

# ... por Humanos

- O mapeamento é realizado por um ser humano
- O processador é um ser humano





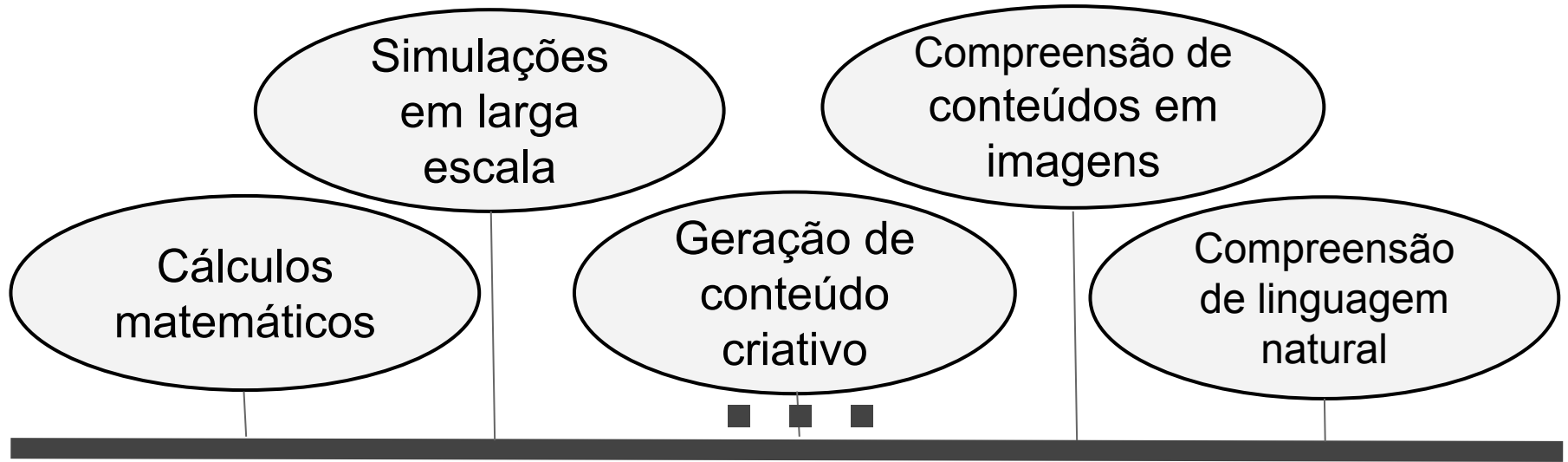
Surgem máquinas capazes de realizar tais computações:  
**computadores digitais**



No início,  
**'computadores'** eram seres humanos que realizavam cálculos matemáticos



... mas ainda existem tarefas que os **seres humanos** são capazes de executar de forma mais eficiente e eficaz



**Máquinas**



**Humanos**



# Tarefa

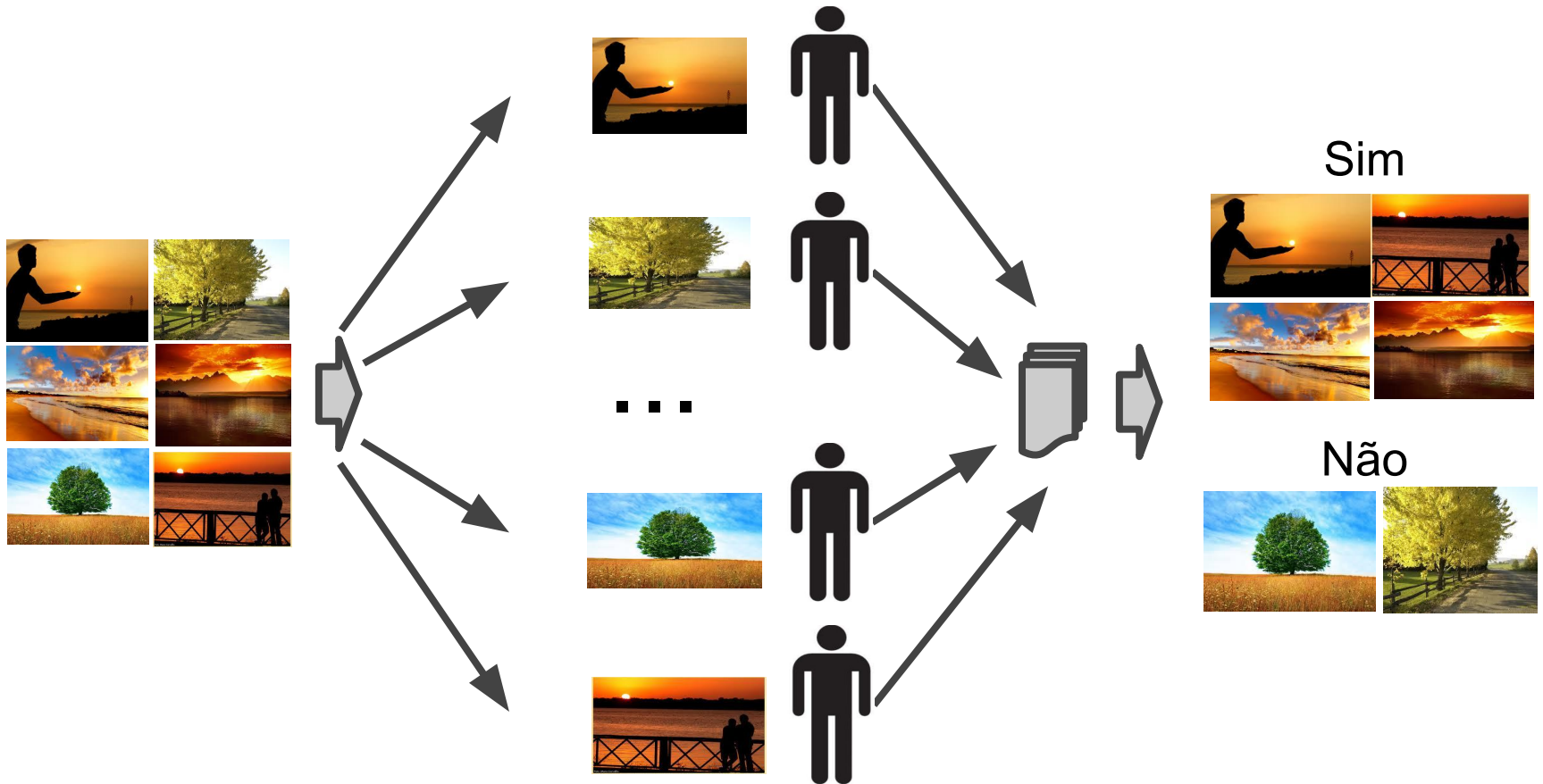
Verifique se a imagem abaixo retrata um pôr-do-sol e marque a resposta mais adequada:



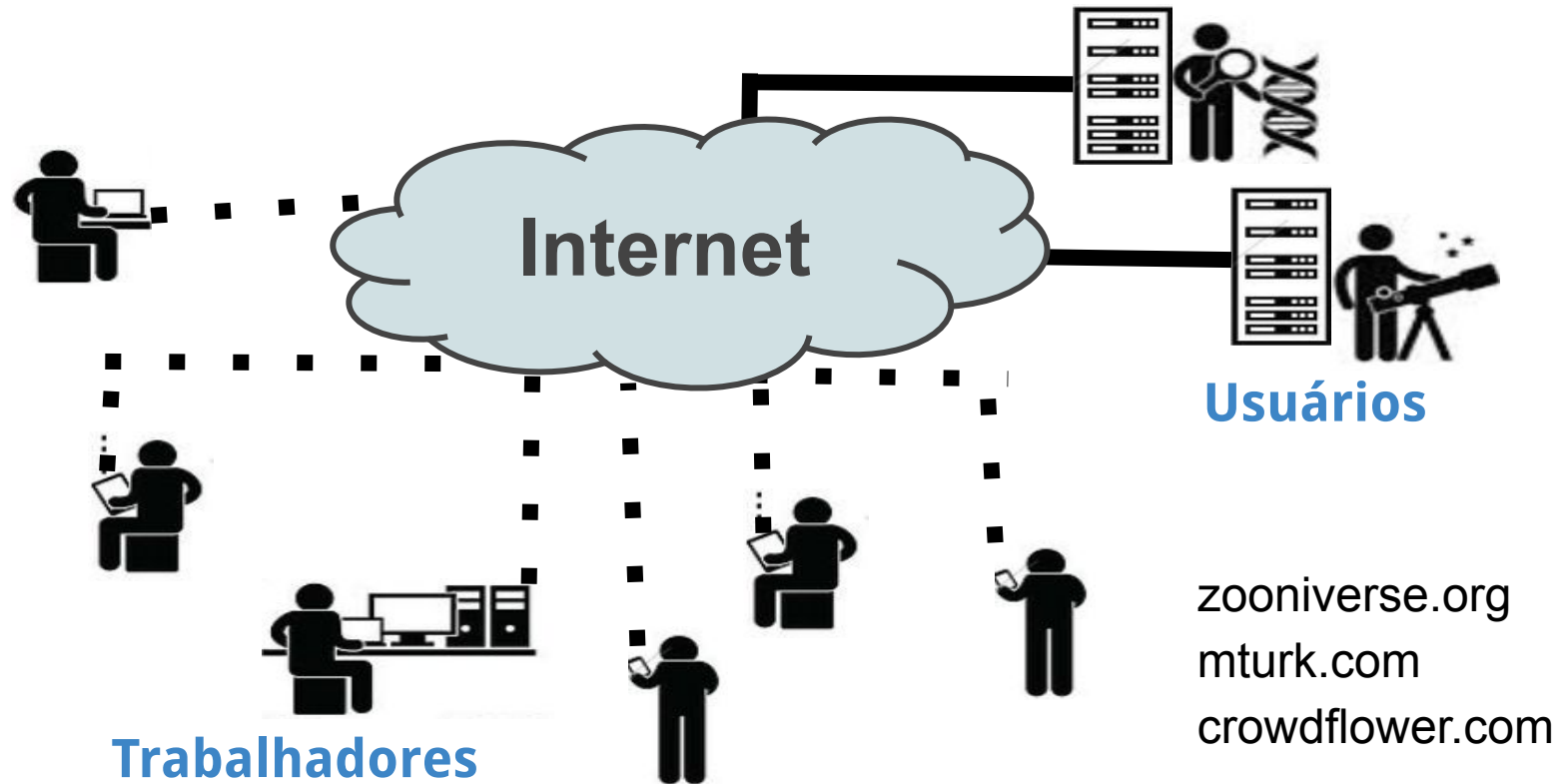
Sim, a imagem retrata um pôr-do-sol

Não, a imagem não retrata um pôr-do-sol

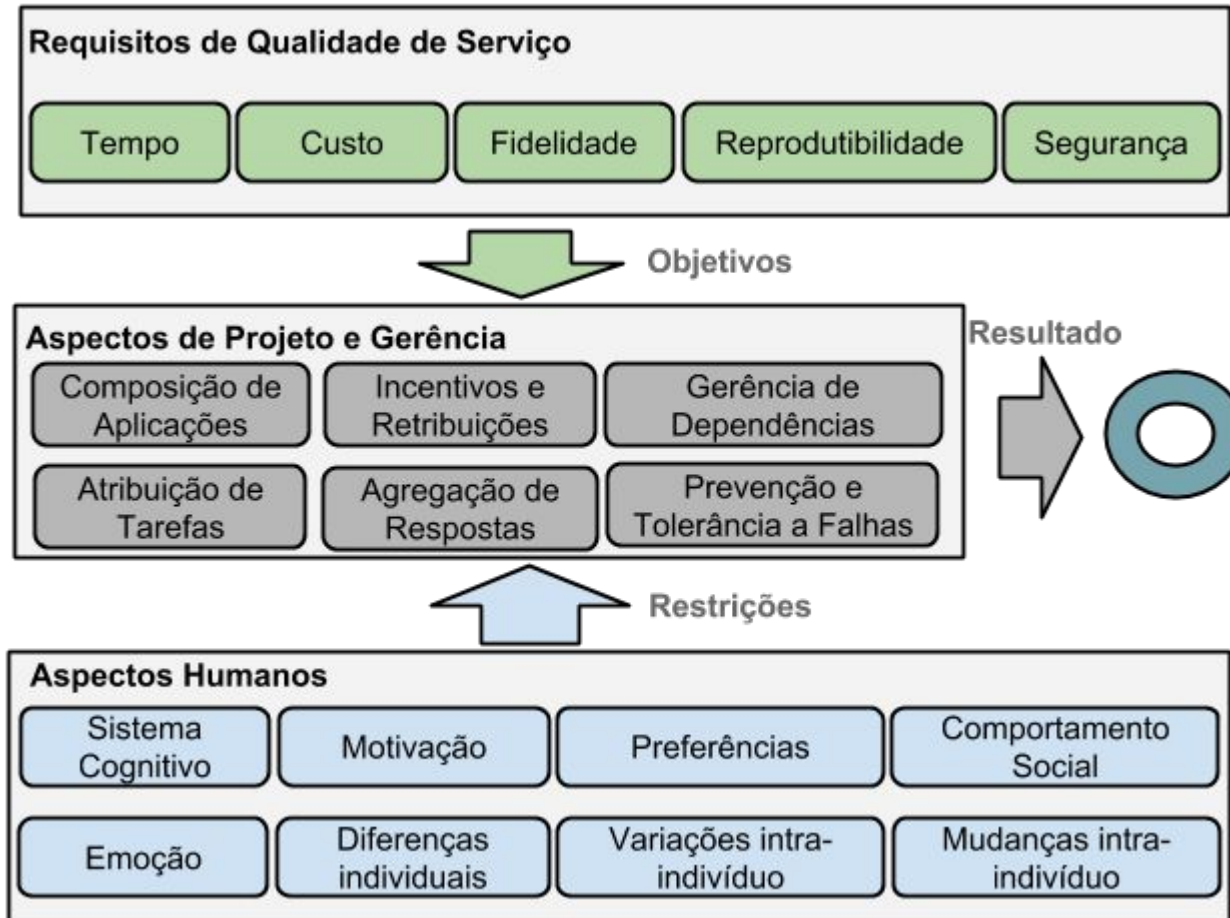
# Aplicação Distribuída



# Sistema Distribuído

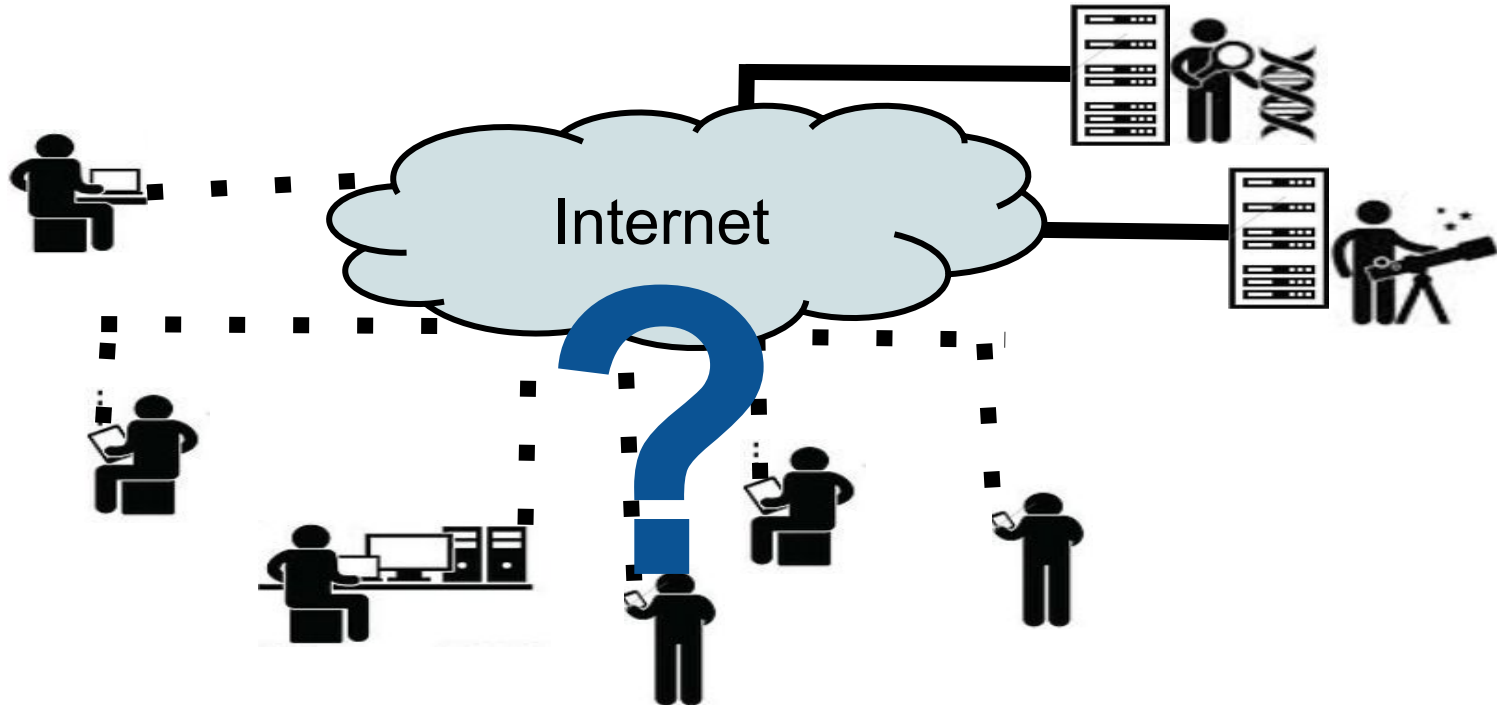


# Funcionamento e Desempenho





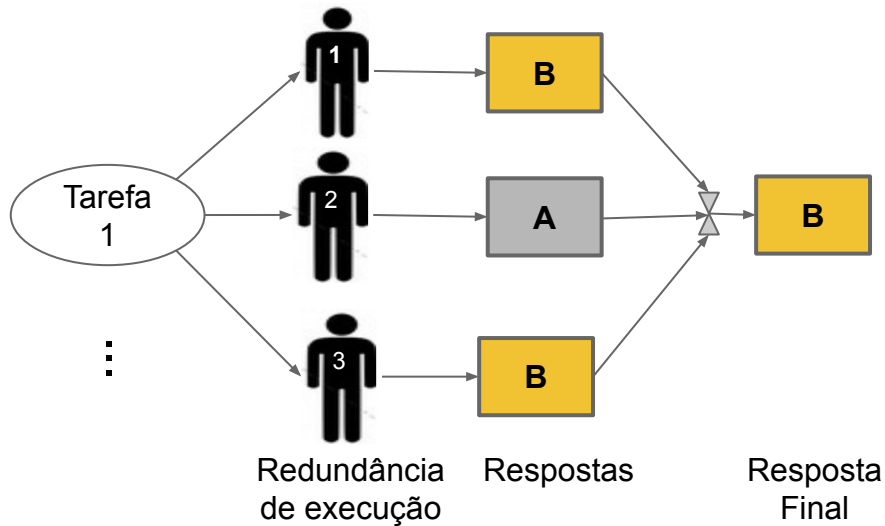
# Problema



Há pouco conhecimento sobre as **características** de oferta de poder cognitivo e de como o sistema pode **usar de forma otimizada** tal oferta

# Uma Questão de Otimização

Redundância para tolerar falhas ou detectar consenso



- Definição da quantidade de redundância
- Atualmente
  - Definido pelo usuário
  - Igual para todas as tarefas
  - Antes da execução
- Subestimar vs superestimar

# Abordagem Proposta

## Caracterização

e

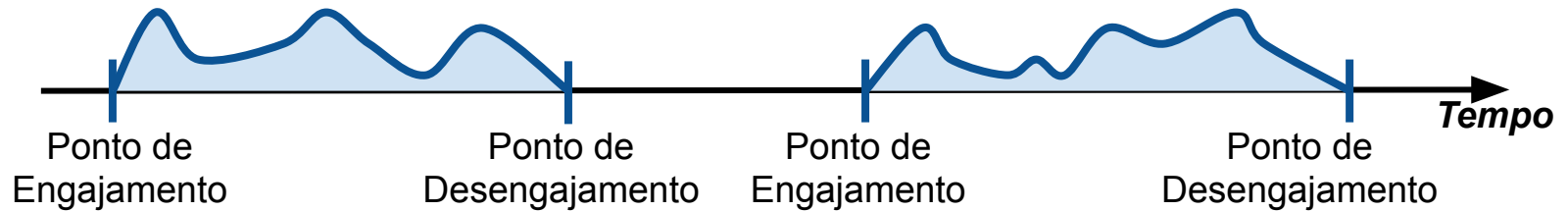
**Engajamento:** Padrões com que os trabalhadores disponibilizam poder cognitivo no sistema

**Credibilidade:** O quão se pode acreditar nas respostas providas pelos trabalhadores

## Otimização

**Replicação adaptativa:** Definição do nível de redundância adequado para cada tarefa de forma automática e durante a execução

# Arcabouço sobre Engajamento



## Avaliação

- Subjetiva/Qualitativa
- Objetiva/Quantitativa

## Dimensões

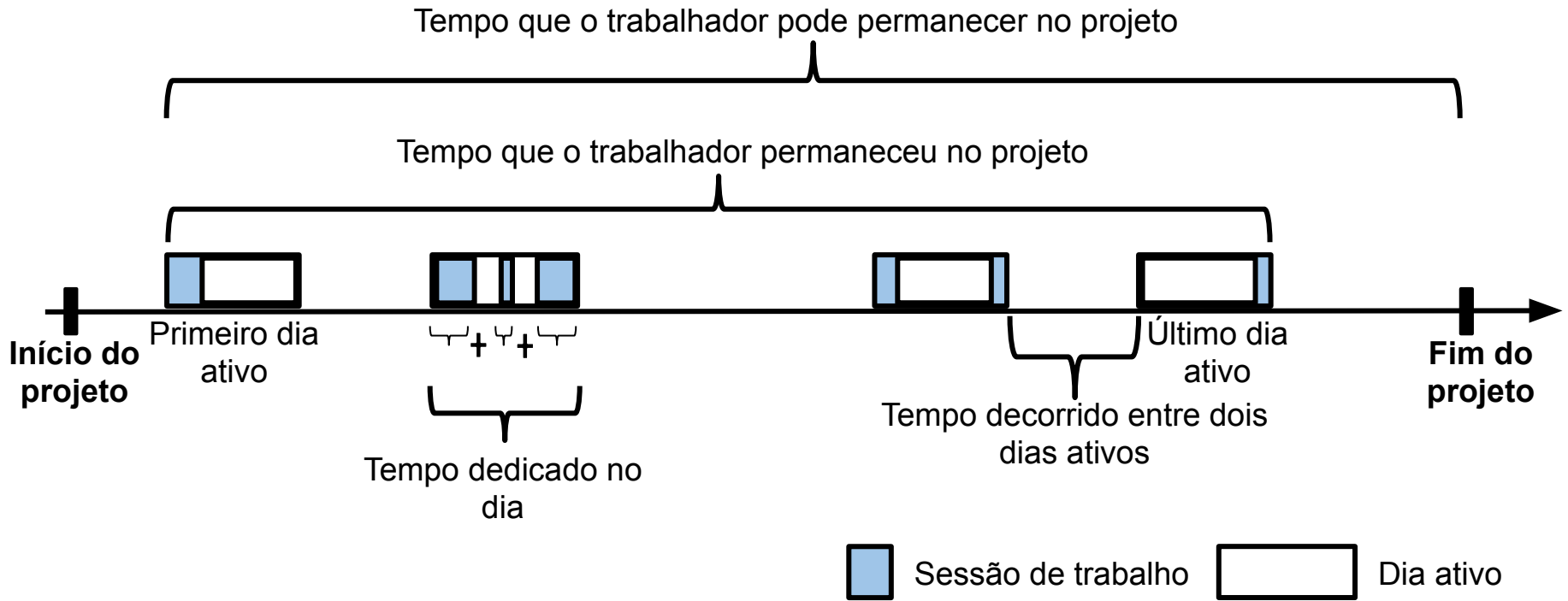
- Tipo
- Grau
- Duração

## Processo

- Curto prazo
- Longo prazo

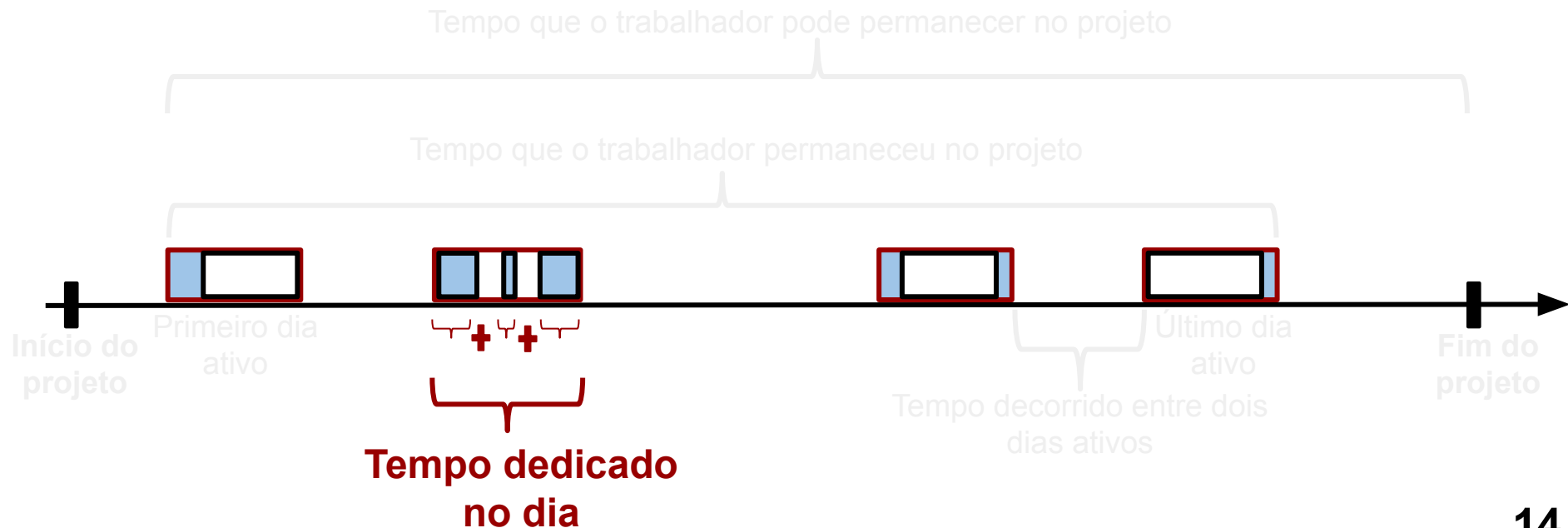
# Métricas de Engajamento

Baseadas em dados da **atuação** dos trabalhadores no projeto



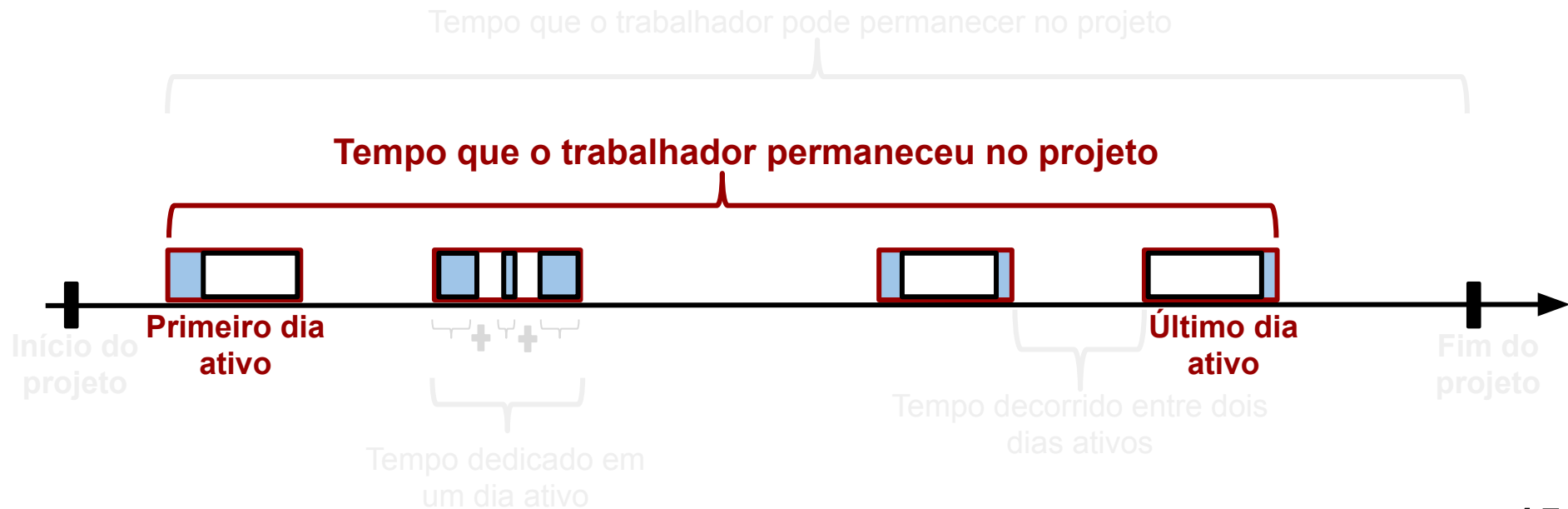
# Tempo Dedicado Diariamente

Média do tempo dedicado nos dias ativos



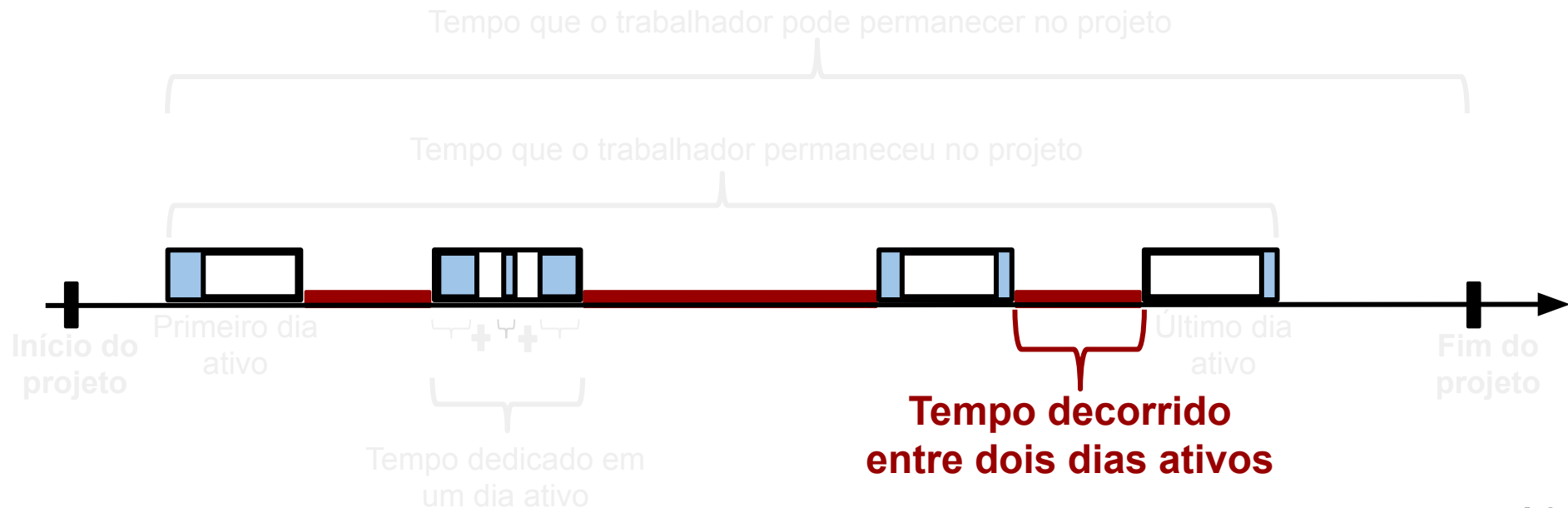
# Taxa de Atividade

Razão entre o número de dias ativos e o número de dias em que o trabalhador permaneceu no projeto



# Variação na Periodicidade

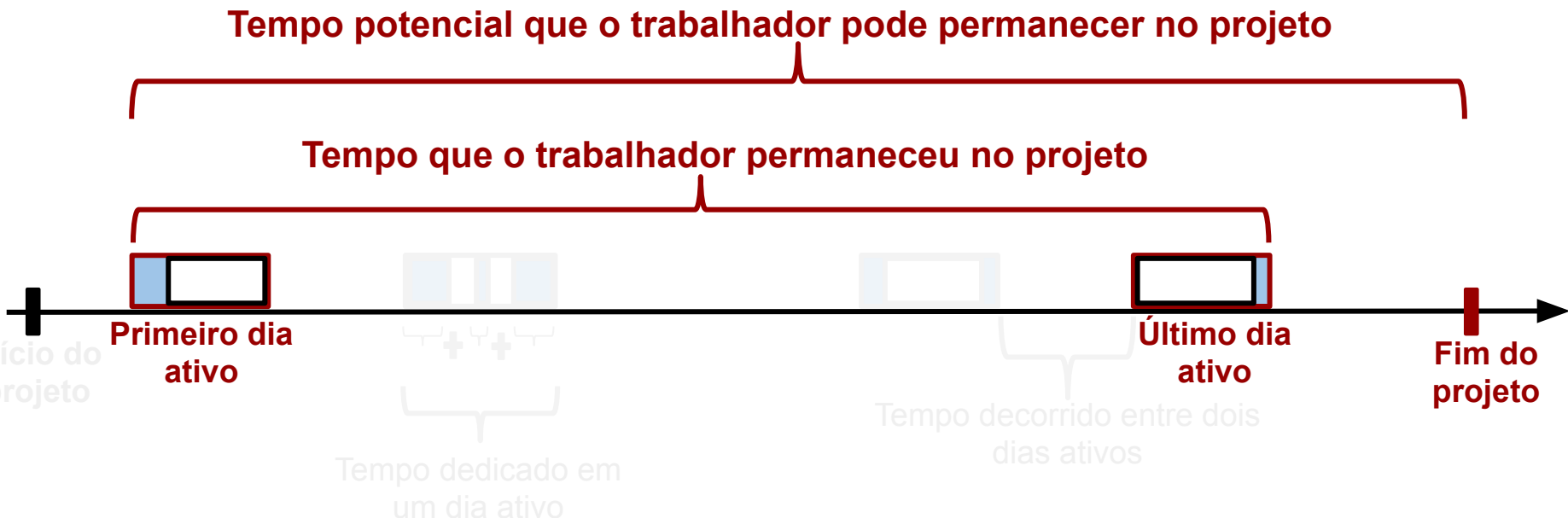
Desvio padrão do tempo decorrido entre dias ativos





# Duração Relativa da Atividade

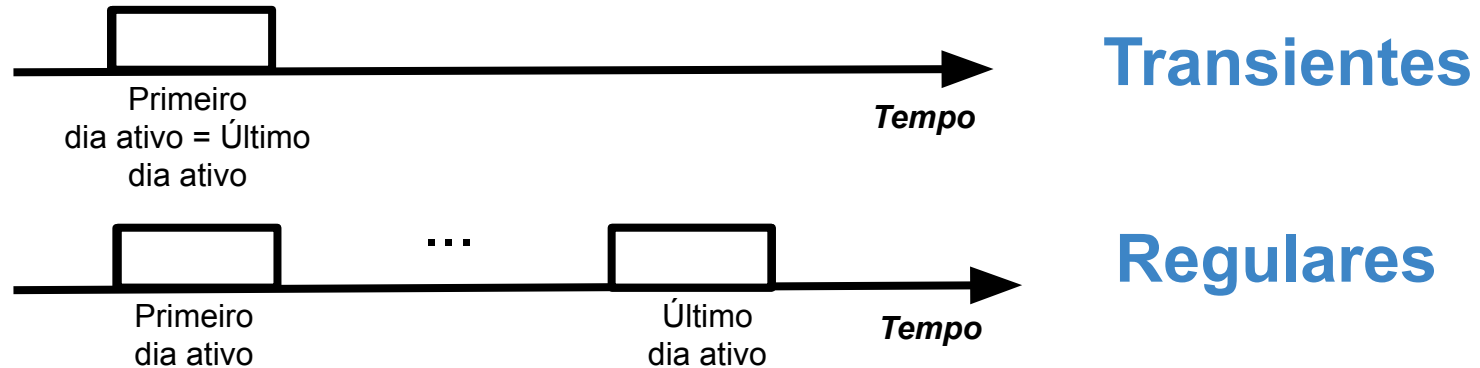
Razão entre o número de dias em que o trabalhador permaneceu no projeto e o tempo potencial de permanência



# Análise em 5 Bases de Dados

	<b>Análise de Sentimentos</b>	<b>Cell Spotting</b>	<b>Galaxy Zoo</b>	<b>Sun4All</b>	<b>The Milky Way Project</b>
Duração (dias)	18	492	840	305	670
#Trabalhadores	1960	1.103	86.413	116	23.889
#Eventos	569.375	94.137	9.667.586	4.328	643.408
Timestamp	●	●	●	●	●

# Classes de Engajamento



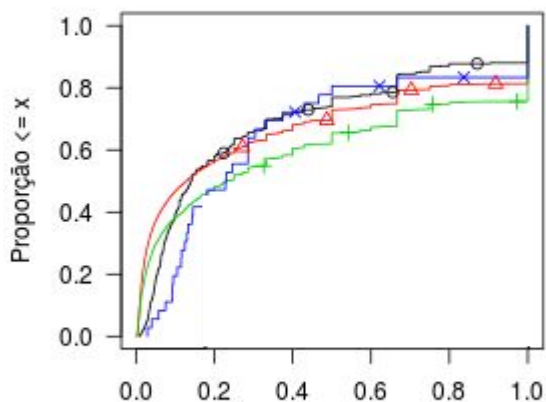
## Regulares

### Minoria

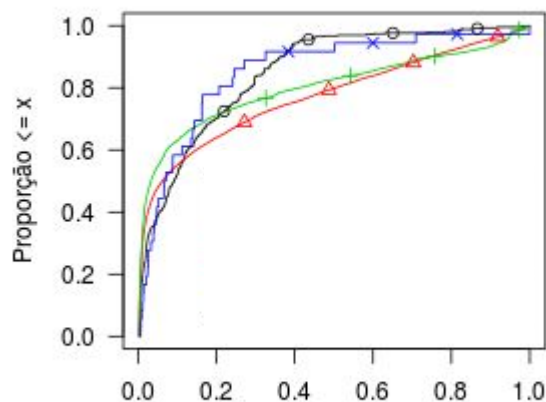
### Importante

Galaxy Zoo	36%	86%
The Milky Way Project	28%	84%
Cell Spotting	42%	90%
Sun4All	35%	70%
Análise de Sentimentos	41% dos trabalhadores	93% do tempo dedicado

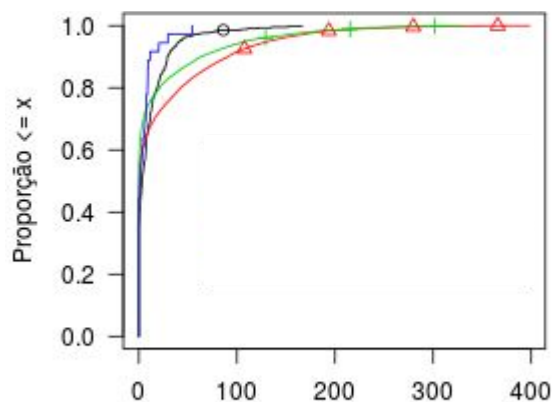
# Distribuições do Engajamento



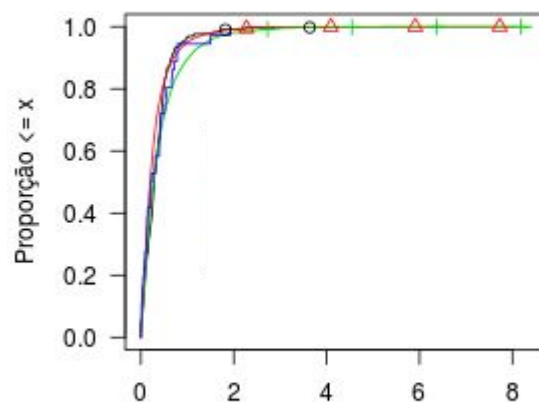
Taxa de atividade



Duração relativa da atividade



Variação na periodicidade (dias)



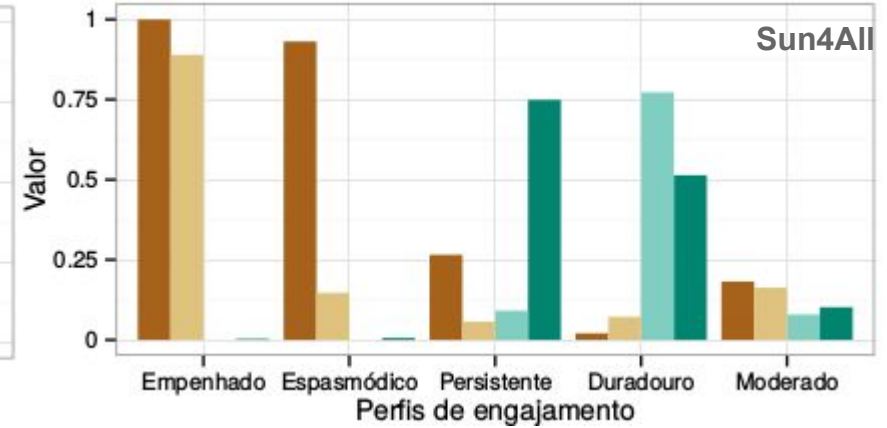
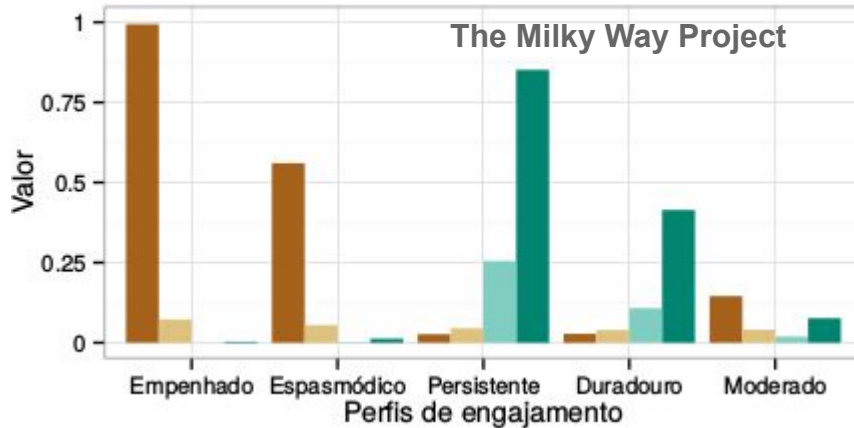
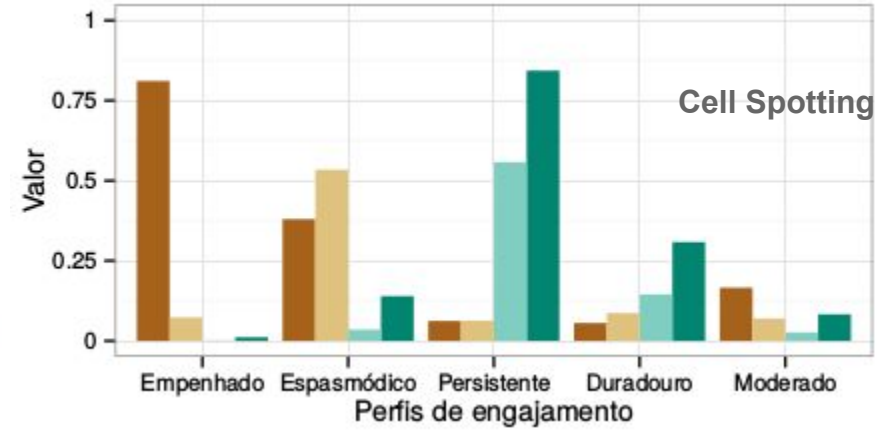
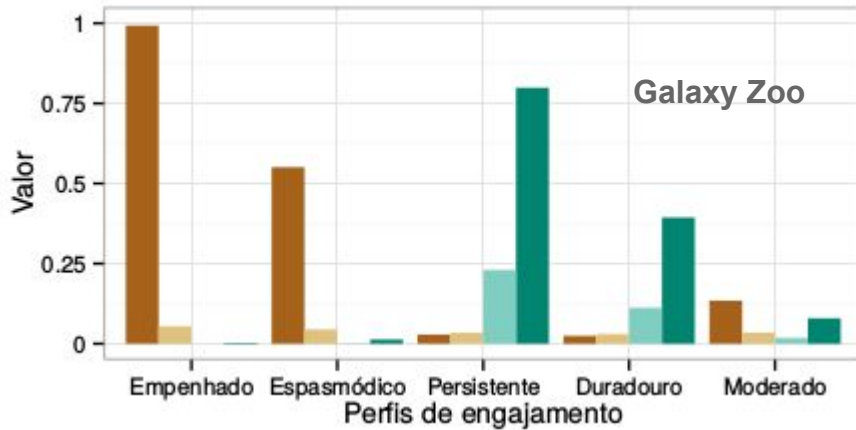
Tempo dedicado diariamente (horas)

- Cell Spotting
- △ Galaxy Zoo
- + The Milky Way Project
- × Sun4All

Semelhanças entre projetos

Diferenças entre métricas

# Perfis de Engajamento



Taxa de atividade
  Tempo dedicado diariamente
  Variação na periodicidade
  Duração relativa da atividade

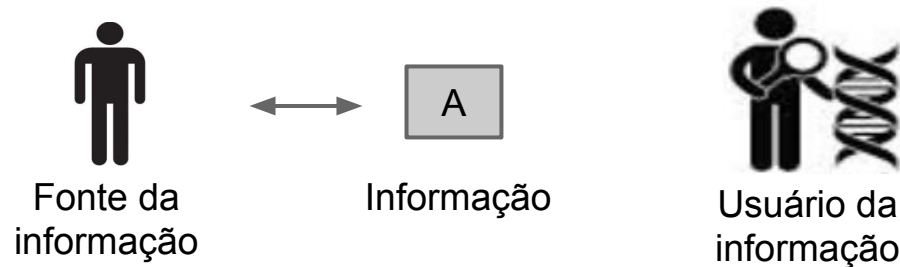
# Importância

	<u>Nº de trabalhadores</u>	<u>Tempo dedicado</u>
Galaxy Zoo	Moderado (31%)	Persistente (41%)
The Milky Way Project	Moderado (30%)	Persistente (40%)
Cell Spotting	Moderado (43%)	Duradouro (32%)
Sun4All	Moderado (69%)	Moderado (51%)

# Destques

- Mapeamos e analisamos o conceito de ‘engajamento’
- Propomos métricas para avaliação quantitativa
  - Grau e duração do engajamento
  - Engajamento de curto prazo e de longo prazo
  - Semelhanças e diferenças entre trabalhadores
  - Importância em termos da oferta de poder cognitivo
- Em dados de 4 projetos reais, observamos
  - Duas classes: regulares e transientes
  - A maioria é transiente e os mais importantes são os regulares
  - Regulares podem ser subdivididos em 5 grupos que diferem em termos do engajamento e da importância

# Arcabouço sobre Credibilidade



## Tipos

- Presumida
- Reputada
- Aparente
- Experimentada

## Modelos

- Binário
- Por limiar
- Espectral

## Fases

- Proeminência
- Interpretação



# Dificuldade da Tarefa

Entropia de Shannon existente no conjunto de respostas

$t_1$	$t_2$	$t_3$
B	B	B
B	B	B
B	A	A
B	A	A
		C
0.0	1.0	1.5

0: quando todas as respostas são iguais

⋮

-> **infinito positivo**: quando aumenta a distribuição dos trabalhadores e a diversidade de respostas

# Métricas de Credibilidade

Baseadas no nível de **concordância** com a maioria

		Tarefas			
		$t_1$	$t_2$	$t_3$	$t_4$
Trabalhadores	$W_1$	A	A	B	A
	$W_2$	B	A	B	A
	$W_3$	B		A	
	$W_4$	B	B	B	B
Respostas majoritárias		B	A	B	A

# Concordância Simples

Proporção de concordância conjunta

	$t_1$	$t_2$	$t_3$	$t_4$
$W_1$	A	A	B	B
$W_2$	B	B	B	A
$W_3$	B	A	A	B
$W_4$	B	A	B	B
	B	A	B	B

Credibilidade Superficial

0: sempre discorda

⋮

1: sempre concorda

# Concordância Ponderada

Média harmônica ponderada entre a credibilidade neutra (0,5) e a proporção de concordância conjunta

	$t_1$	$t_2$	$t_3$	$t_4$
$W_1$	A	A	B	B
$W_2$	B	B	B	A
$W_3$	B	A	A	B
$W_4$	B	A	B	B
	B	A	B	B

## Credibilidade Presumida

-> 0: sempre discorda e o número de tarefas executadas cresce

⋮

-> 0,5: poucas tarefas executadas ou probabilidade de concordância de 0,5

⋮

-> 1: sempre concorda e o número de tarefas executadas cresce

# Concordância Experimentada

Proporção de concordância conjunta, descontada a concordância devido ao acaso (Cohen's kappa)

	$t_1$	$t_2$	$t_3$	$t_4$
$W_1$	A	A	B	B
$W_2$	B	B	B	A
$W_3$	B	A	A	B
$W_4$	B	A	B	B
	B	A	B	A

Credibilidade Experimentada

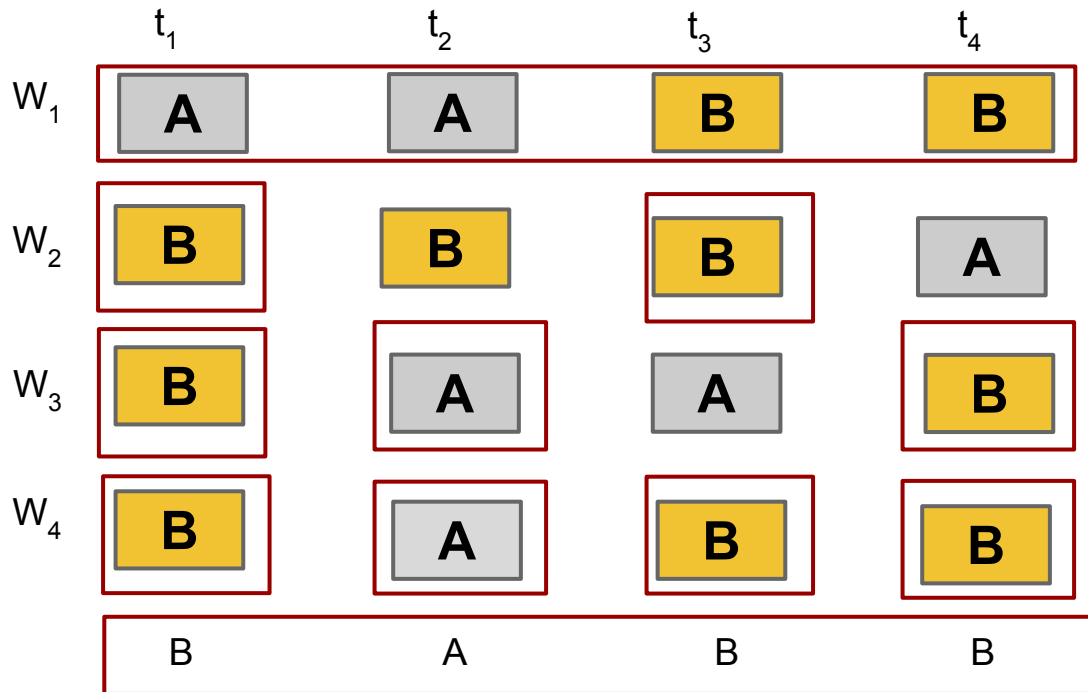
0: discorda mais do que ocorreria de forma aleatória

⋮

1: concorda mais que ocorreria de forma aleatória

# Concordância Reputada

Credibilidade aumenta quando concorda com uma maioria crível e reduz quando discorda de uma maioria crível



Credibilidade Reputada

**0**: sempre discorda e de maioria de credibilidade 1

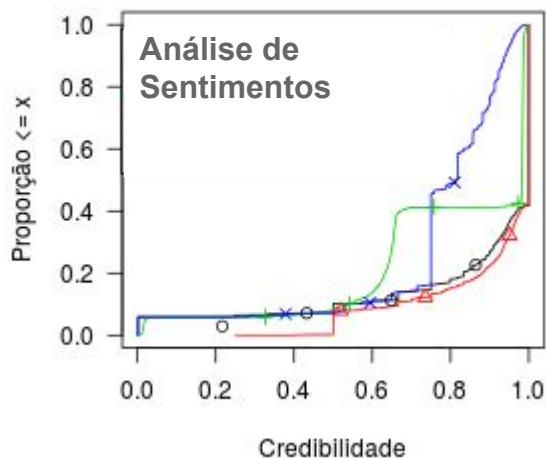
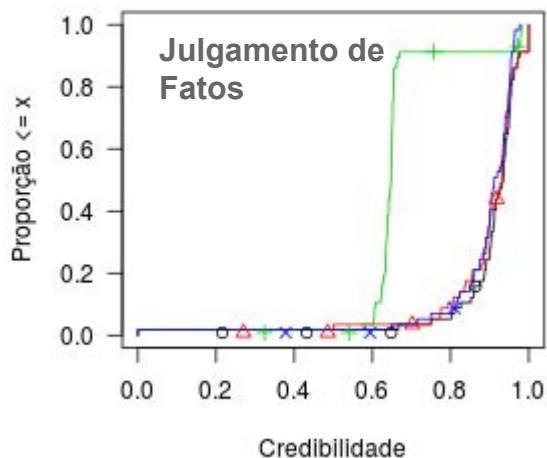
⋮

**1**: sempre concorda e de maioria de credibilidade 1

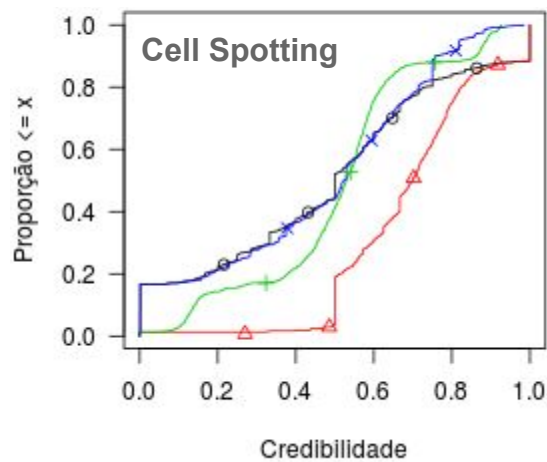
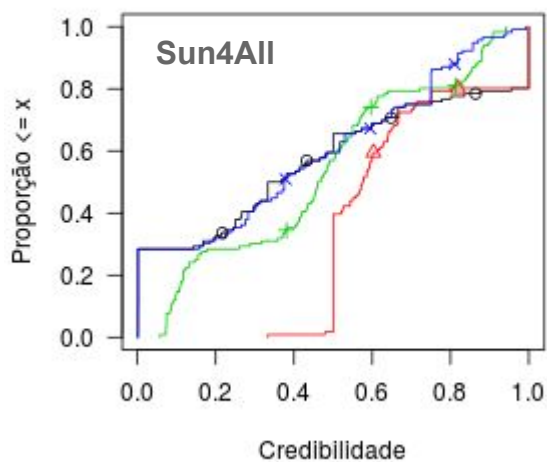
# Análise em 4 Bases de Dados

	<b>Análise de Sentimentos</b>	<b>Cell Spotting</b>	<b>Julgamento de Fatos</b>	<b>Sun4All</b>
#Trabalhadores	1.960	1103	57	116
#Tarefas	98.980	4.067	42.624	417
Respostas	●	●	●	●

# Distribuições de Credibilidade



- Conc. Simples
- △ Conc. Experimentada
- + Conc. Reputada
- × Conc. Ponderada

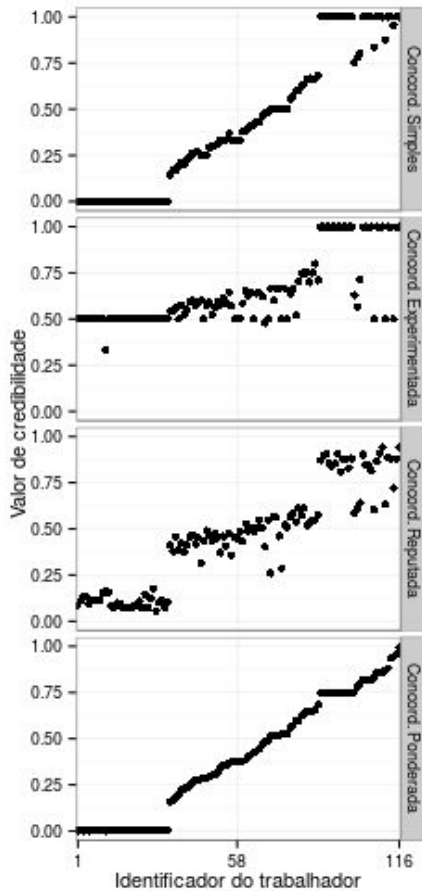


Distribuições diferentes entre métricas

Algumas semelhanças entre projetos



# Distâncias entre Métricas



Sun4All

Distância absoluta média entre os valores de credibilidade

- $d(c,e)=0,41$
- $d(c,r)=0,47$
- $d(c,p)=0,35$
- $d(e,r)=0,48$
- $d(e,p)=0,47$
- $d(r,p)=0,48$

Distância entre os ranques de credibilidade (kendall)

- $\tau(c,e)=0,25$
- $\tau(c,r)=0,12$
- $\tau(c,p)=0,04$
- $\tau(e,r)=0,20$
- $\tau(e,p)=0,28$
- $\tau(r,p)=0,11$

# Distâncias com Dificuldade

## Análise de Sentimentos

		<b>h=0,0</b>	<b>h=0.1</b>	<b>h=0.2</b>	<b>h=0.3</b>	<b>h=0.4</b>
		<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465
Distância entre os valores de credibilidade	$d(c, e)$	<b>0,00</b>	<b>0,00</b>	<b>0,00</b>	<b>0,01</b>	<b>0,02</b>
	$d(c, r)$	<b>0,00</b>	0,01	0,02	0,05	0,05
	$d(c, p)$	0,13	0,12	0,11	0,09	0,11
	$d(e, r)$	<b>0,00</b>	0,02	0,03	0,06	0,06
	$d(e, p)$	0,13	0,12	0,12	0,10	0,13
	$d(r, p)$	0,13	0,12	0,12	0,12	0,13
			<hr/>			
		<b>h=0,0</b>	<b>h=0,1</b>	<b>h=0,2</b>	<b>h=0,3</b>	<b>h=0,4</b>
		<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465	<i>N</i> = 465
Distância entre os ranques de credibilidade	$\tau(c, e)$	<b>0,00</b>	<b>0,04</b>	<b>0,08</b>	<b>0,17</b>	<b>0,17</b>
	$\tau(c, r)$	<b>0,00</b>	0,49	0,50	0,50	0,50
	$\tau(c, p)$	0,45	0,46	0,47	0,46	0,47
	$\tau(e, r)$	<b>0,00</b>	0,49	0,50	0,50	0,51
	$\tau(e, p)$	0,45	0,46	0,47	0,47	0,47
	$\tau(r, p)$	0,45	0,47	0,47	0,49	0,47

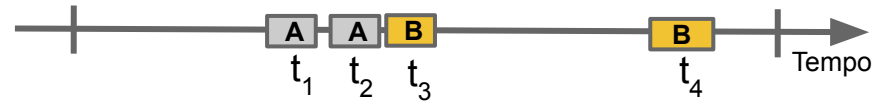
# Destaques

- Mapeamos e analisamos o conceito de ‘credibilidade’
- Propomos quatro métricas de credibilidade espectral
  - Proeminência da credibilidade
  - Concordância entre os trabalhadores
  - Dificuldade das tarefas
- Em dados de 4 projetos, observamos que as métricas:
  - Geram diferentes distribuições dos trabalhadores
  - Estimam valores diferentes de credibilidade
  - Originam diferentes ordem de credibilidade dos trabalhadores

# Relações

## Credibilidade, Engajamento e Dificuldade

	$t_1$	$t_2$	$t_3$	$t_4$
$W_1$	A	A	B	B
$W_2$	B	B	B	A
$W_3$	B	A	A	B
$W_4$	B	A	B	B



Linha do tempo do trabalhador  $W_1$

Concord. Simples

Concord.  
Ponderada

Taxa de atividade

Tempo dedicado  
diariamente

Concord.  
Experimentada

Concord.  
Reputada

Variação na  
Periodicidade

Duração Relativa  
da atividade

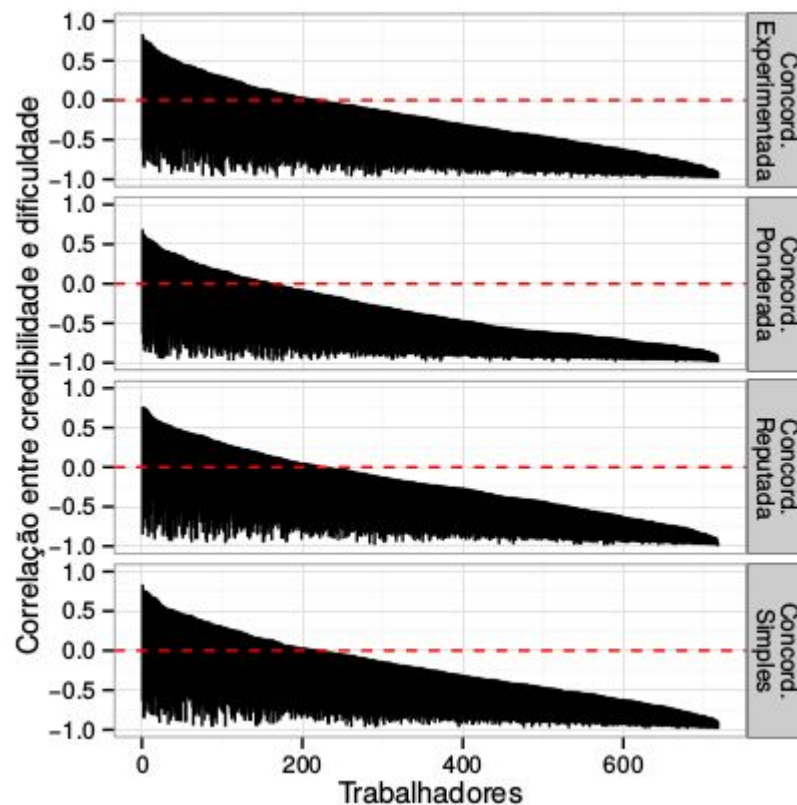
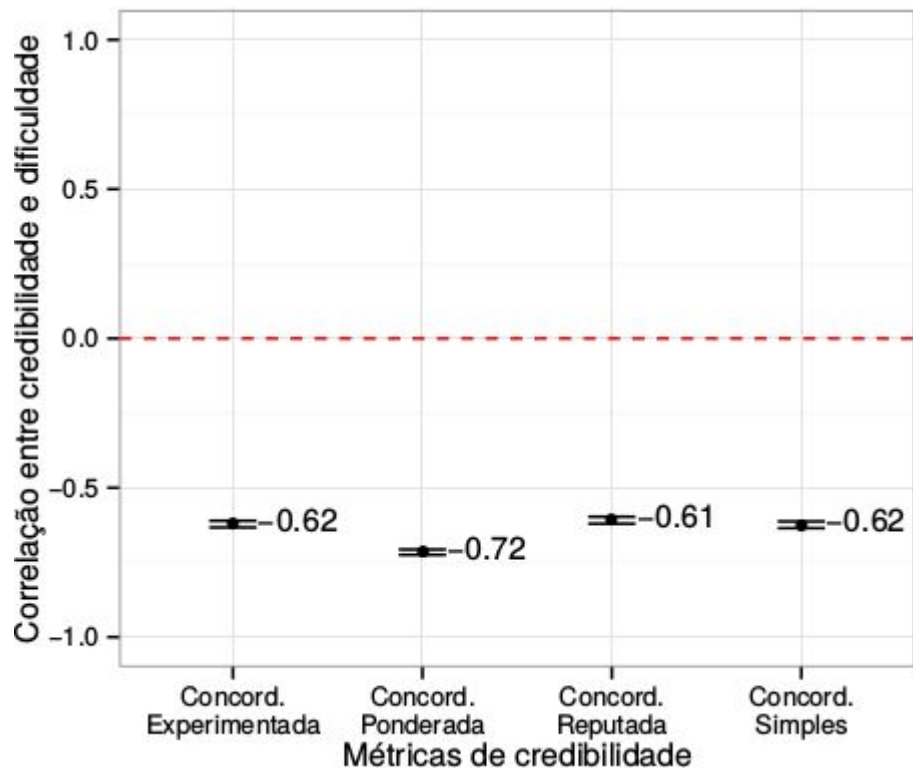
Dificuldade

Dificuldade Média

# Análise em 4 Bases de Dados

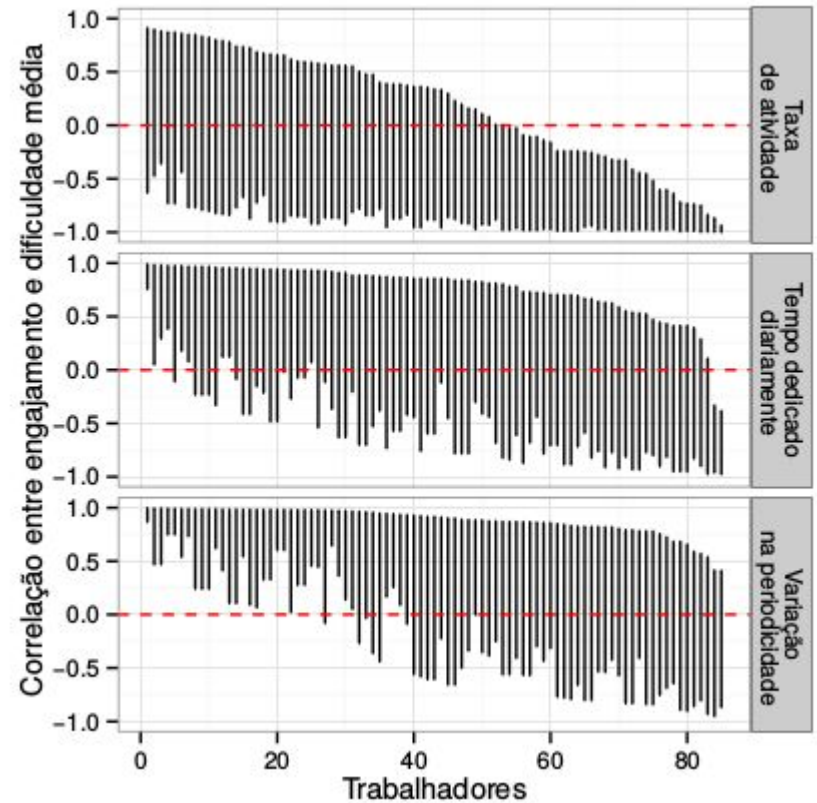
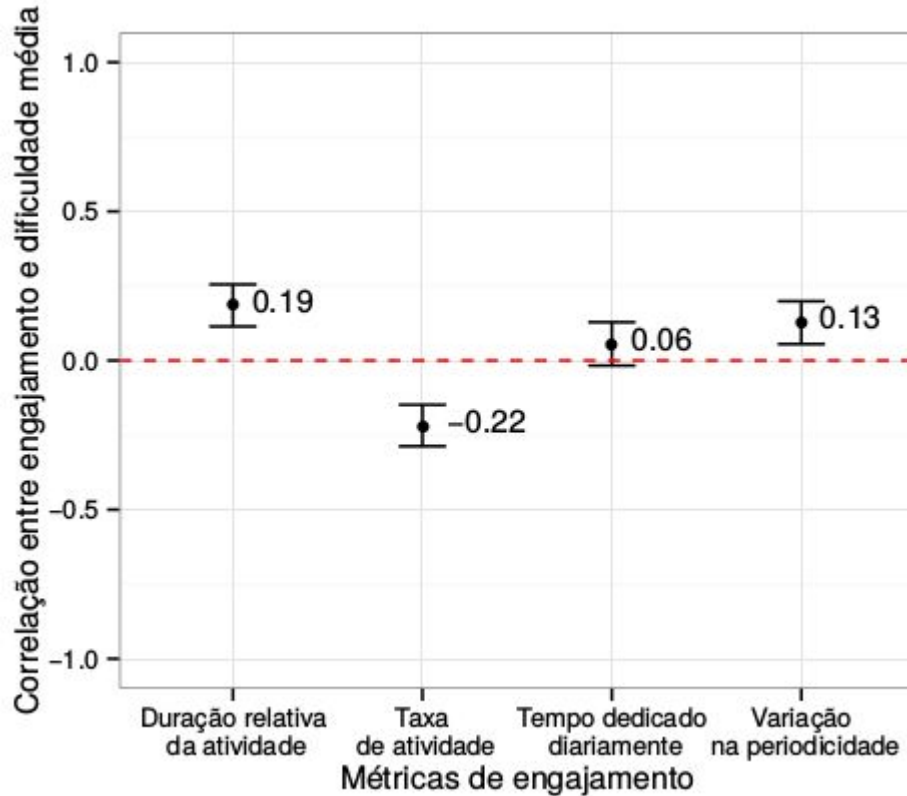
	<b>Análise de Sentimentos</b>	<b>Cell Spotting</b>	<b>Julgamento de Fatos</b>	<b>Sun4All</b>
Duração (dias)	18	492	-	305
#Trabalhadores	1960	1.103	57	116
#Tarefas	98.860	4.067	42.624	417
#Eventos	569.375	94.134	220.000	4.328
Timestamp	●	●		●
Respostas	●	●	●	●

# Credibilidade e Dificuldade



Análise de Sentimentos

# Engajamento e Dificuldade Média

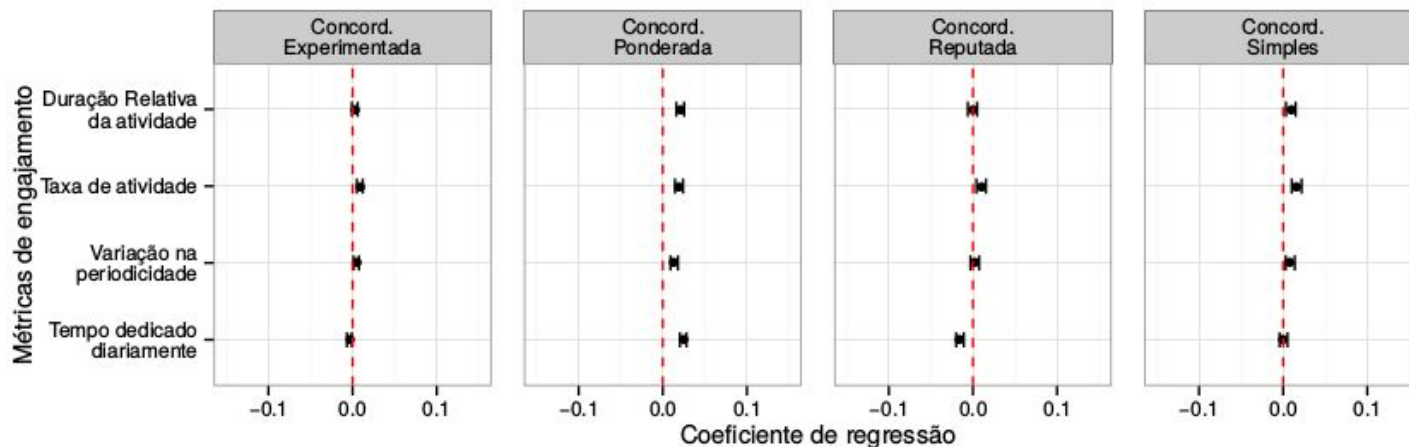


Análise de Sentimentos

# Engajamento e Credibilidade

## Análise de Sentimentos

Métrica	Ganho
Concord. Simples	1.654,47
Concord. Experimentada	1.545,68
Concord. Reputada	1.576,57
Concord. Ponderada	2.050,63

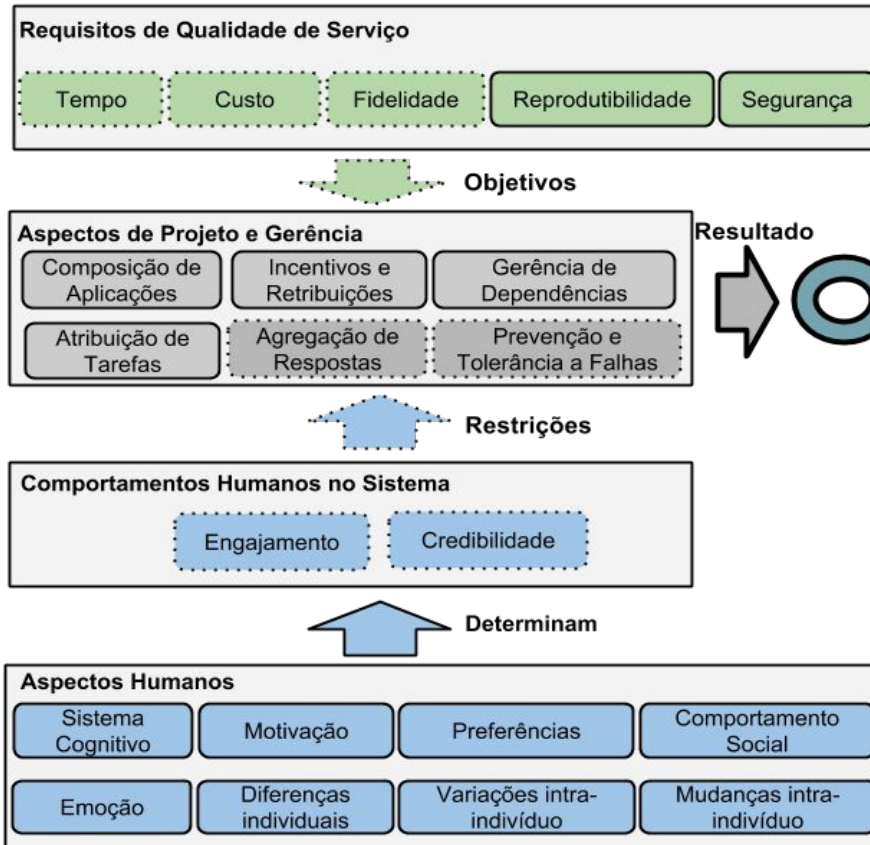




# Destques

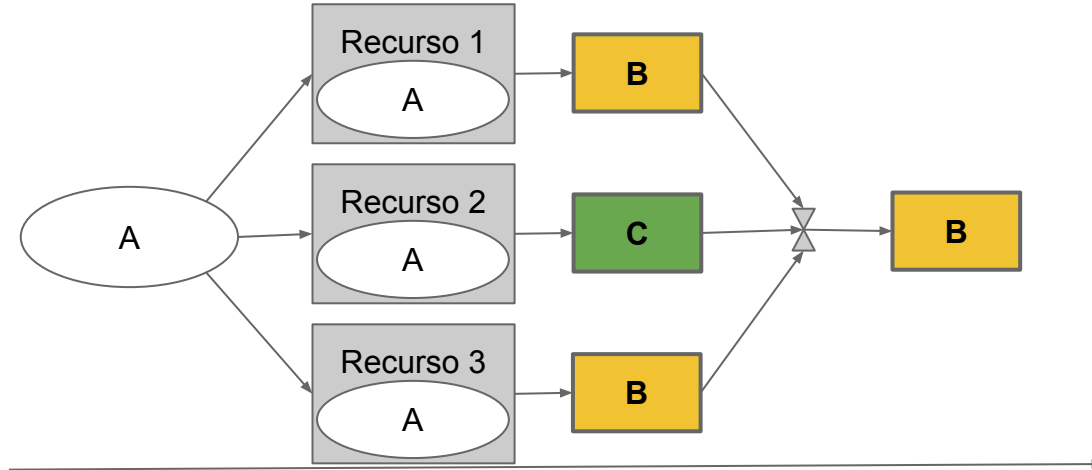
- Discutimos possíveis interdependências entre dimensões da atuação dos trabalhadores
  - engajamento, credibilidade e dificuldade
- Nos projetos estudados, observamos que
  - A credibilidade tem forte relação com a dificuldade das tarefas
  - A relação entre engajamento e dificuldade média varia com o projeto
  - A relação entre engajamento e credibilidade é dependente da dificuldade
  - Engajamento tem baixo poder explicativo da credibilidade

# Estudo de Caso de Otimização



Replicação de tarefas baseada na **credibilidade** dos trabalhadores

# Arcabouço sobre Replicação



## Propósitos

- Desempenho
- Tolerância a falhas

## Decisões

- Grau de replicação
- Agregação de respostas

# Replicação de Tarefas

---

## Algoritmo 1: Replicação Adaptativa Baseada em Credibilidade

---

**Entrada:** Tarefa  $t$ , Credibilidade requerida  $credRequ$ , Máximo de réplicas  $maxRepl$ , Urgência  $urge$

**Saída:** Resposta final para a tarefa  $respGrupo$ , Credibilidade da resposta final  $credGrupo$ ;

```
1  $contRepl \leftarrow 0$ ; // Contador de réplicas
2  $S_t \leftarrow \{\}$ ; /* Mapa cujas chaves são respostas e os valores são listas que
   mantêm os identificadores dos trabalhadores que proveram as respostas
   indicadas nas chaves. */
3  $numReplPorTurno \leftarrow \max(maxRepl * urge, 1)$ ;
4 repita
5    $numReplNesteTurno \leftarrow \min(numReplPorTurno, maxRepl - contRepl)$ ;
6    $atribuiReplicas(numReplNesteTurno, t, S_t)$ ; /* gera  $numReplNesteTurno$ 
   réplicas da tarefa  $t$ , obtém as respostas dos trabalhadores e as
   mantém no mapa  $S_t$  */
7    $G \leftarrow calculaCredibilidadeDosTrabalhadores(S_t)$ ; /* Definido no Algoritmo 2 */
8    $respGrupo, credGrupo \leftarrow pegaGrupoTrabalhadoresMaisCrivel(G)$ ; /* Definido no
   Algoritmo 3 */
9    $contRepl \leftarrow contRepl + numReplNesteTurno$ ;
10 até  $credGrupo \geq credRequ$  or  $contRepl = maxRepl$ ;
11 retorna  $respGrupo, credGrupo$ ;
```

---

## Objetivo

- Obter uma resposta crível enquanto garante:
  - Credibilidade requerida
  - Máximo de réplicas
  - Urgência

## Se $maxRepl$ , mas não $credRequ$

- Perspectiva conservadora
  - Tarefa sem conclusão
- Perspectiva não conservadora
  - Grupo mais crível

# Avaliação em 2 Bases de Dados

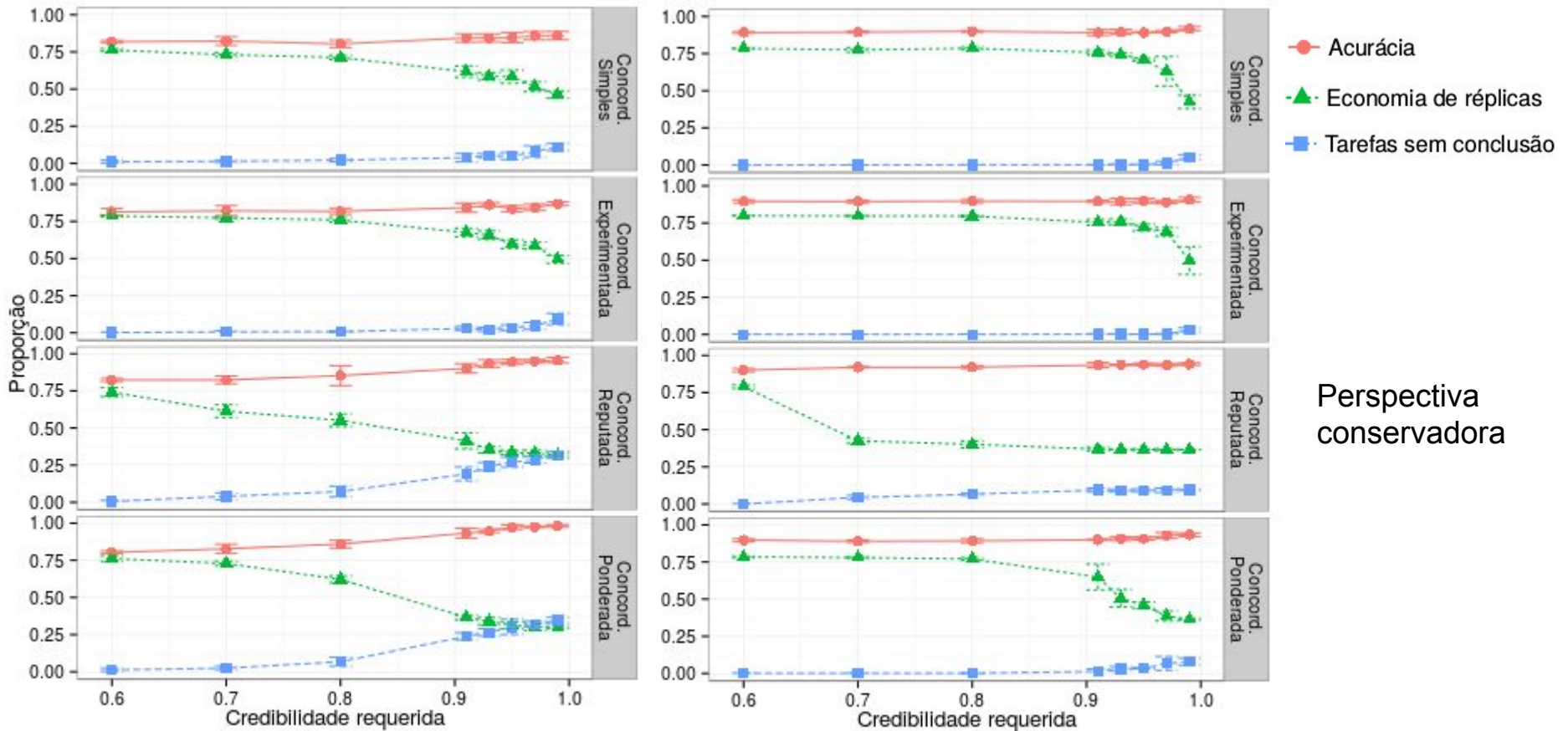
- Simulações usando duas bases de dados como rastros
- Sequência de execução e as respostas providas pelos trabalhadores
- Se a replicação termina antes que todas as réplicas sejam utilizadas

	<b>Análise de Sentimentos</b>	<b>Julgamento de Fatos</b>
Duração (dias)	18	-
#Trabalhadores	1.960	57
#Tarefas	98.980	42.624
#Eventos	569.375	220.000
Respostas	●	●
Ground truth	●	●

# Avaliação

<b>Variáveis Independentes</b>	
Métrica de Credibilidade	Experimentada, Reputada, Ponderada, Simples
Credibilidade Requerida	0,6 - 0,99
Urgência	0;0,25; 0,5; 0,75; 1
<b>Variáveis Dependentes</b>	
Economia de réplicas	
Acurácia	
Proporção de tarefas sem conclusão	
<b>Cenários de Referência</b>	
Valor de referência mínimo	Voto majoritário
Valor de referência máximo	Oráculo

# Efeitos da Credibilidade Requerida

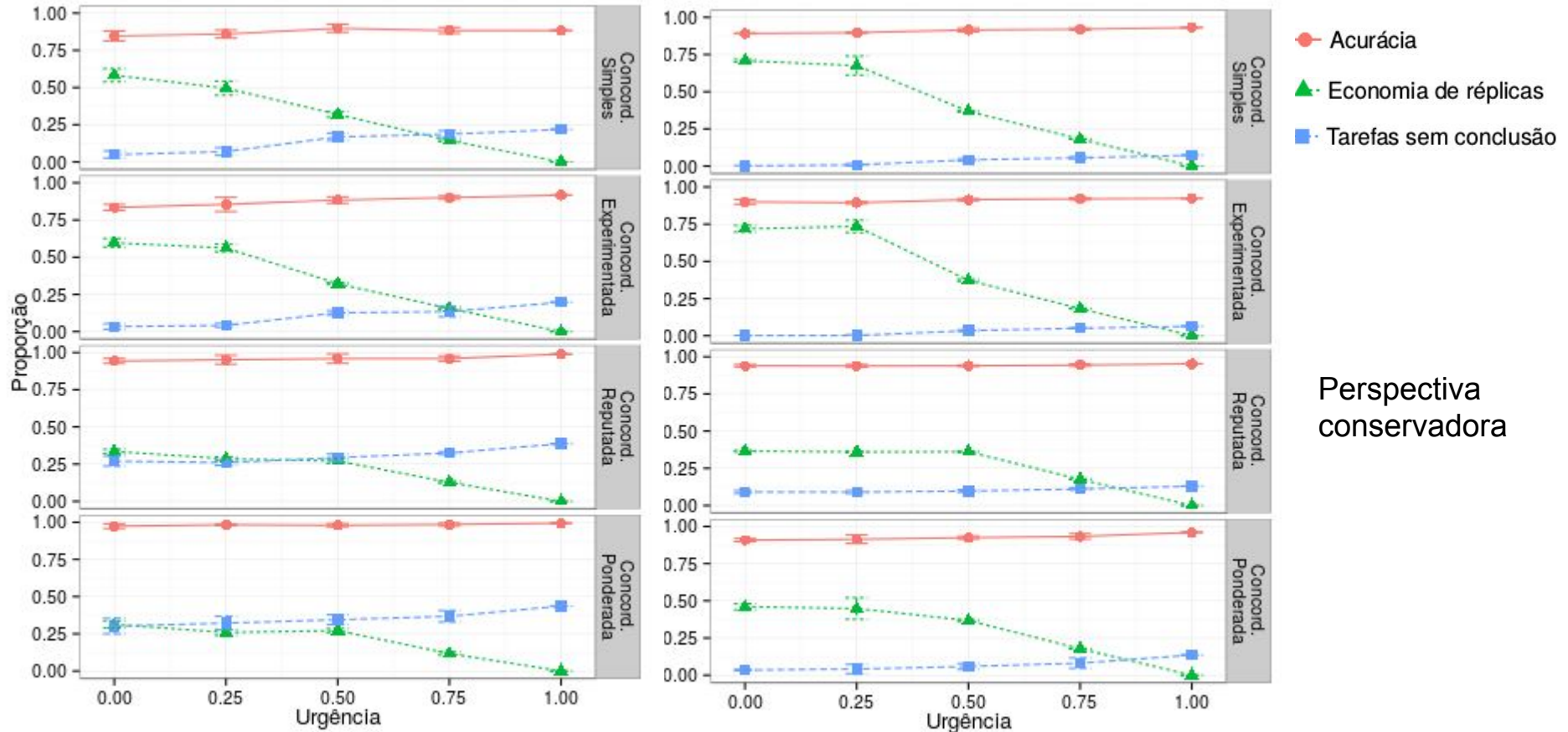


Análise de Sentimentos

Julgamento de Fatos

Perspectiva conservadora

# Efeitos da Urgência



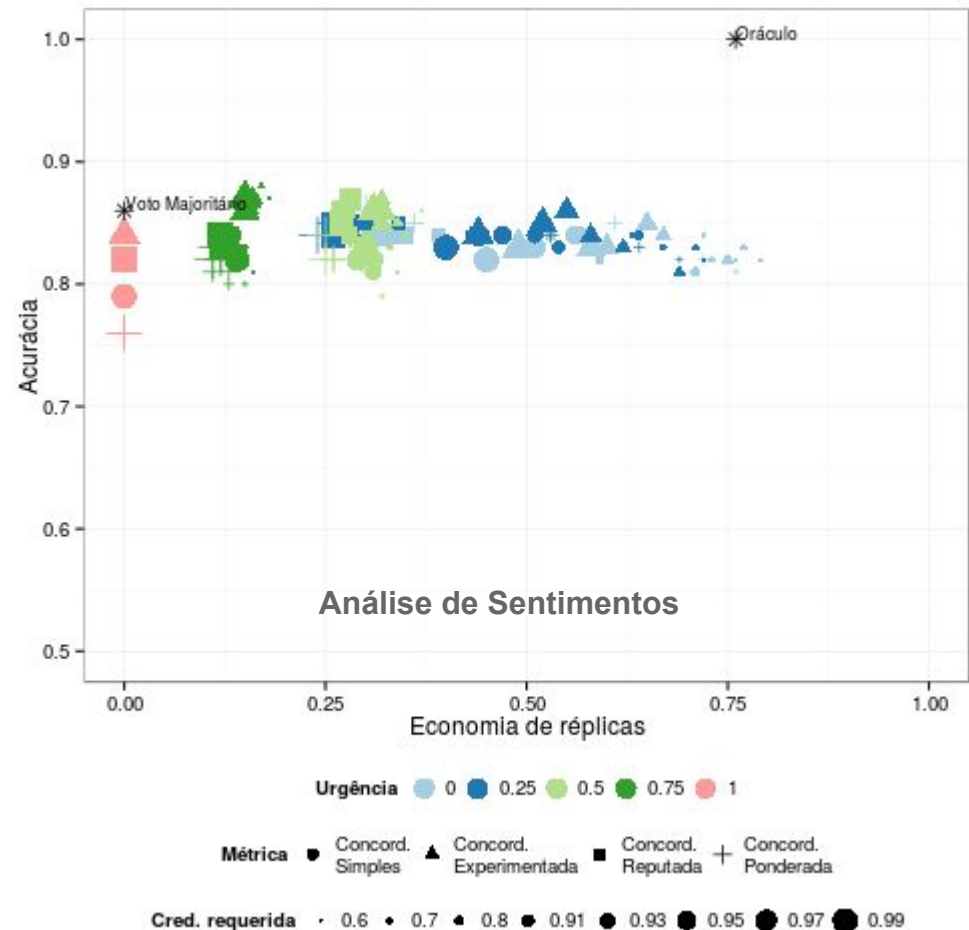
Análise de Sentimentos

Julgamento de Fatos



# Cenários de Referência

- Interesses do usuário
  - Configurações
- Oráculo
  - Referência para economia de réplicas
- Voto majoritário
  - Referência para a Acurácia



# Destques

- Mapeamos e analisamos o conceito de replicação
- Propomos um algoritmo de replicação baseado em
  - Credibilidade espectral e binária
  - Proeminência e interpretação
- Em simulações, observamos que
  - O desempenho do algoritmo depende dos parâmetros e das tarefas
  - Pode superar o oráculo e o voto majoritário
  - Concordância experimentada e concordância reputada sempre estão entre as melhores métricas

# Contribuições

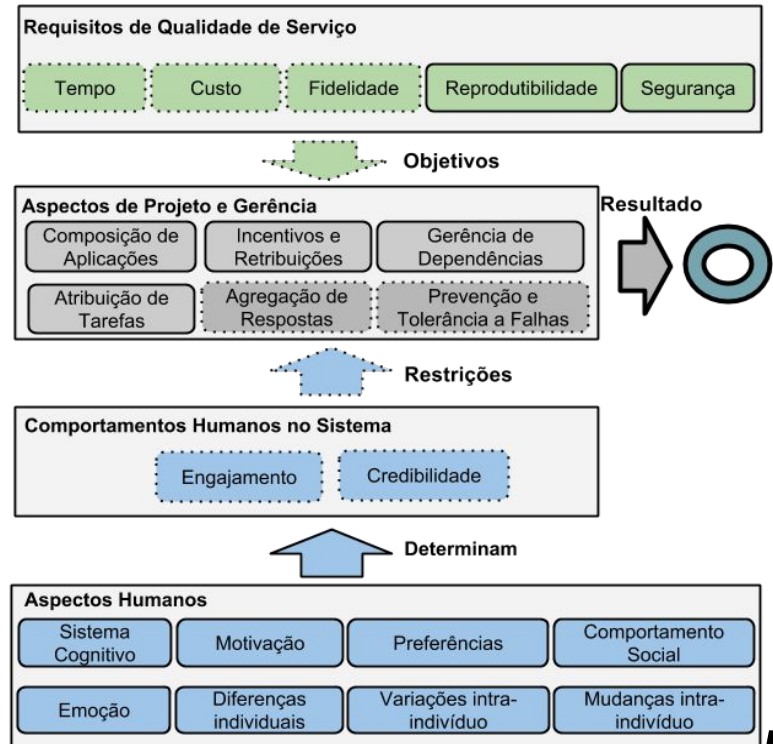
Problema: Pouco conhecimento sobre características da oferta de poder cognitivo e de como o sistema pode usar tal oferta de forma otimizada

Analizamos computação por humanos na perspectiva de um sistema distribuído

Analizamos a oferta de poder cognitivo no sistema usando as lentes do engajamento e da credibilidade

Avaliamos o emprego de replicação adaptativa para melhorar o uso do poder cognitivo

Obtivemos resultados sobre engajamento, credibilidade e replicação em 6 sistemas



# Muito Obrigado

por sua presença e atenção!