

Information Security Engineering

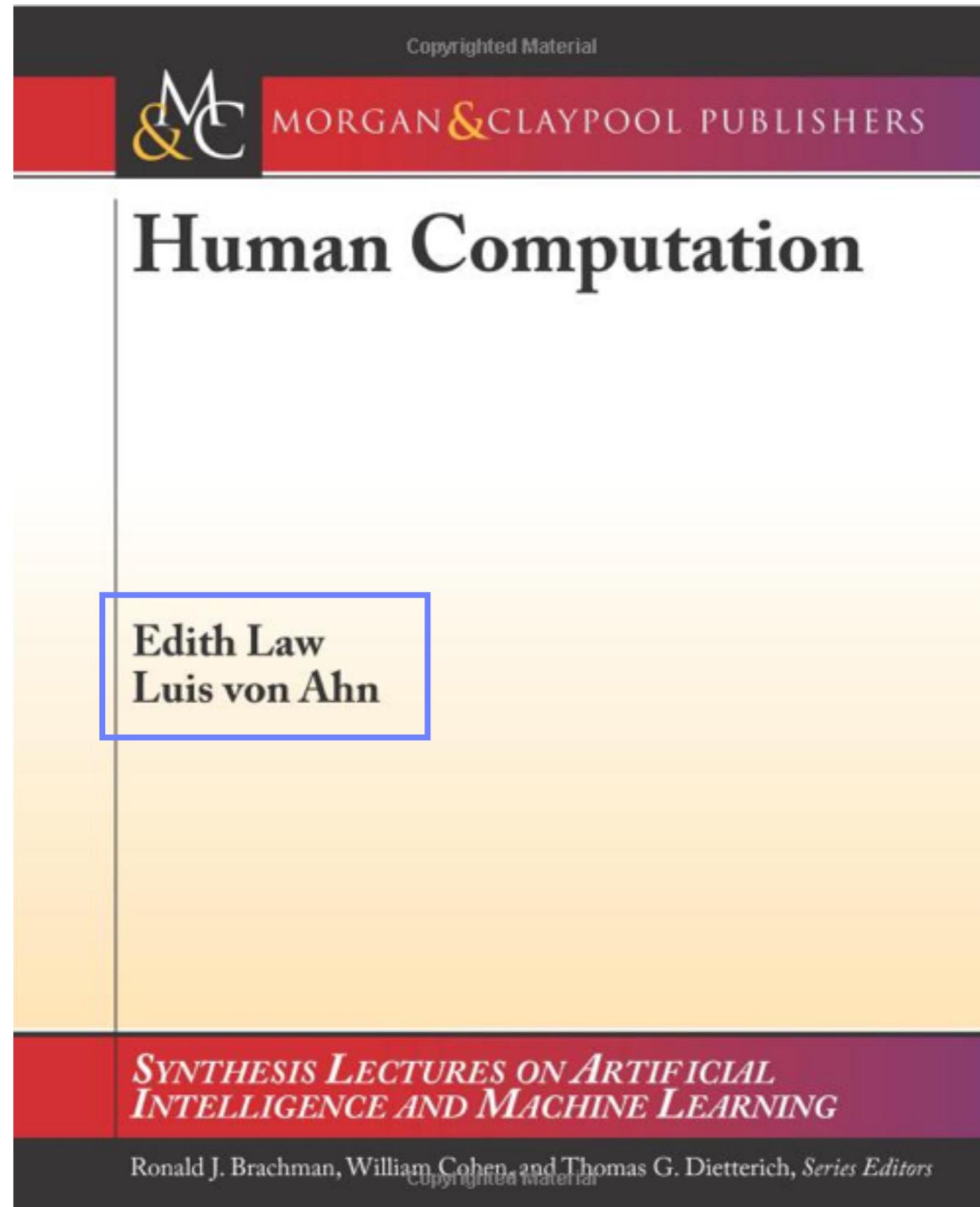
2017.11.08

Human Computation



Huiping Sun(孙惠平)
sunhp@ss.pku.edu.cn

参考书



参考书章节

- *Introduction*
- *Human Computation Algorithms*
- *Aggregating Outputs*
- *Task Routing*
- *Understanding Workers and Requesters*
- *The Arts of Asking Questions*
- *Conclusions*

旗舰会议

HCOMP 2017

The fifth AAAI Conference on Human Computation and Crowdsourcing



October 24-26, 2017



Quebec City, Canada



主要内容

- *Human Computation*概念
- *Human Computation*算法
- *Human Computation*例子

Human Computation概念

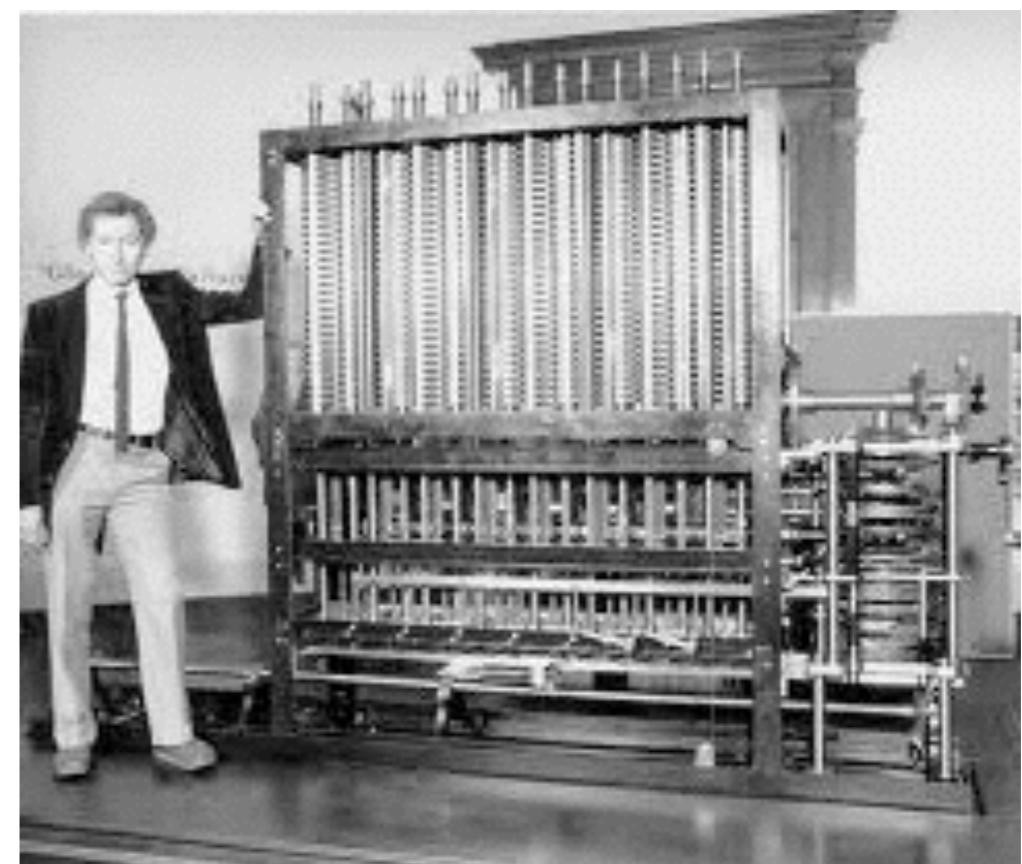
Computation 历史



计算公式 → 任务分解



操作指南 → 结果合并



Human Computation

- 计算
 - * 使用算法映射输入到输出的过程
- Human Computation 2005
 - * 人来执行的计算 → 人工智能难题

思考

现在有哪些Human Computation应用？

Human Computation产生需要什么基础？

人工智能难题

- 现在依然存在许多人工智能难题
 - * 人很容易解决
 - * 但是复杂的计算机算法很难解决
- 常见的人工智能难题
 - * 感知（目标识别、分类）
 - * 自然语言分析（观点分析、翻译）
 - * 认知（计划、推理）

人工智能难题 - OCR

The Breckinridge and Lane Democrats, having taken courage at the recent eastern advices, are organizing energetically for the campaign. Several prominent Democrats who at first favored DOUGLAS, are coming out for the other side, apparently under the pressure of Federal influence. An address to the National Democracy of California, urging the party to support BRECKINRIDGE, has recently been published, which manifestly has strengthened that side of the question. It is signed by 65 Democrats, many of whom occupy respectable and prominent positions in the party, 22 of them are Federal office-holders, eight more are recipients of Federal patronage, and the others represent a mass of politicians giving the document most weight. The Douglas Democrats are also active. The Irish and German vote will mostly go with that branch of the party, but it is difficult to estimate which wing is the stronger. Thus far 17 Democratic newspapers have declared for DOUGLAS, 13 for BRECKINRIDGE, and 9 remain non-committal, with even chances of going either way. Under these circumstances the Republicans entertain not unjustifiable hopes that the Democratic divisions may be so equally balanced as to give the State to LINCOLN. Some very respectable Bell and Everett meetings have been held in different parts of the State, but thus far that party does not exhibit much rank and file strength.

The New-York State Yacht Squadron, on its annual cruise to Newport, came into the harbor yesterday afternoon. The following are the names of the boats that came to anchor here: *Jessie*, *Geraldine*, *Evelyn*, *Annie*, *Manning*, *Julia*, *Bonita*, *Magie*, *Widgeon*, *Rambler*, *Fleur-de-Lis*, *Henrietta*, *Sea-Drift* and *Maria*, with the steamer *America* as a tender. On anchoring, each boat fired a gun, according to custom. The reports were heard distinctly in the city, causing considerable inquiry as to "what was up," and quite a number of sanguine individuals came into our office to inquire if the guns were not annunciations signals of the successful laying of the Atlantic Cable. We invariably replied in the negative. The squadron will leave to-day for Newport. The yachts *Washington* and *Rattler*, of this city, start with it, with parties of New-Haven people.

人工智能难题 - 图像识别



Tag系统 Q&A系统

相关概念

动物
机器

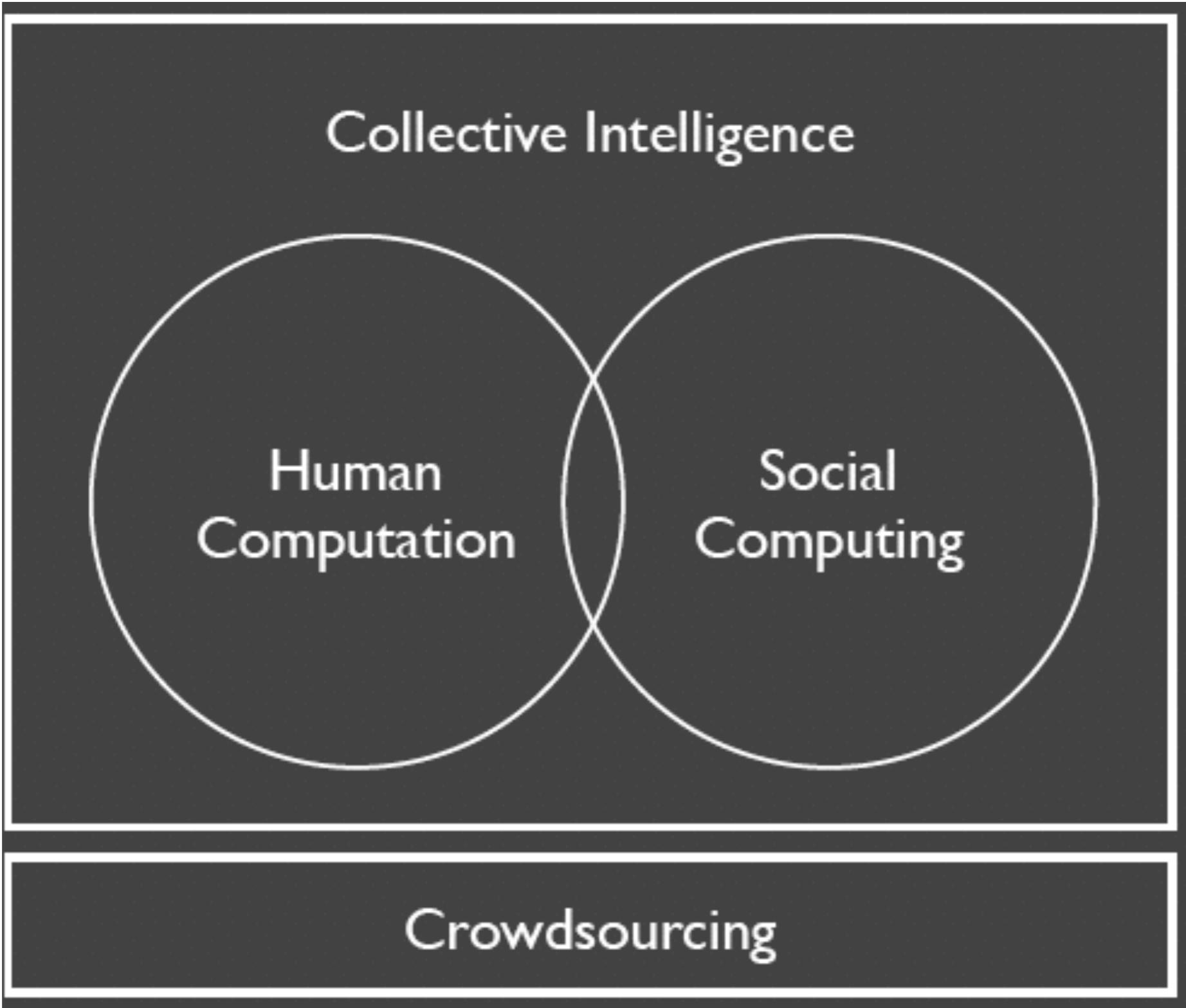
社会
行为

工具
方法

群体
智能

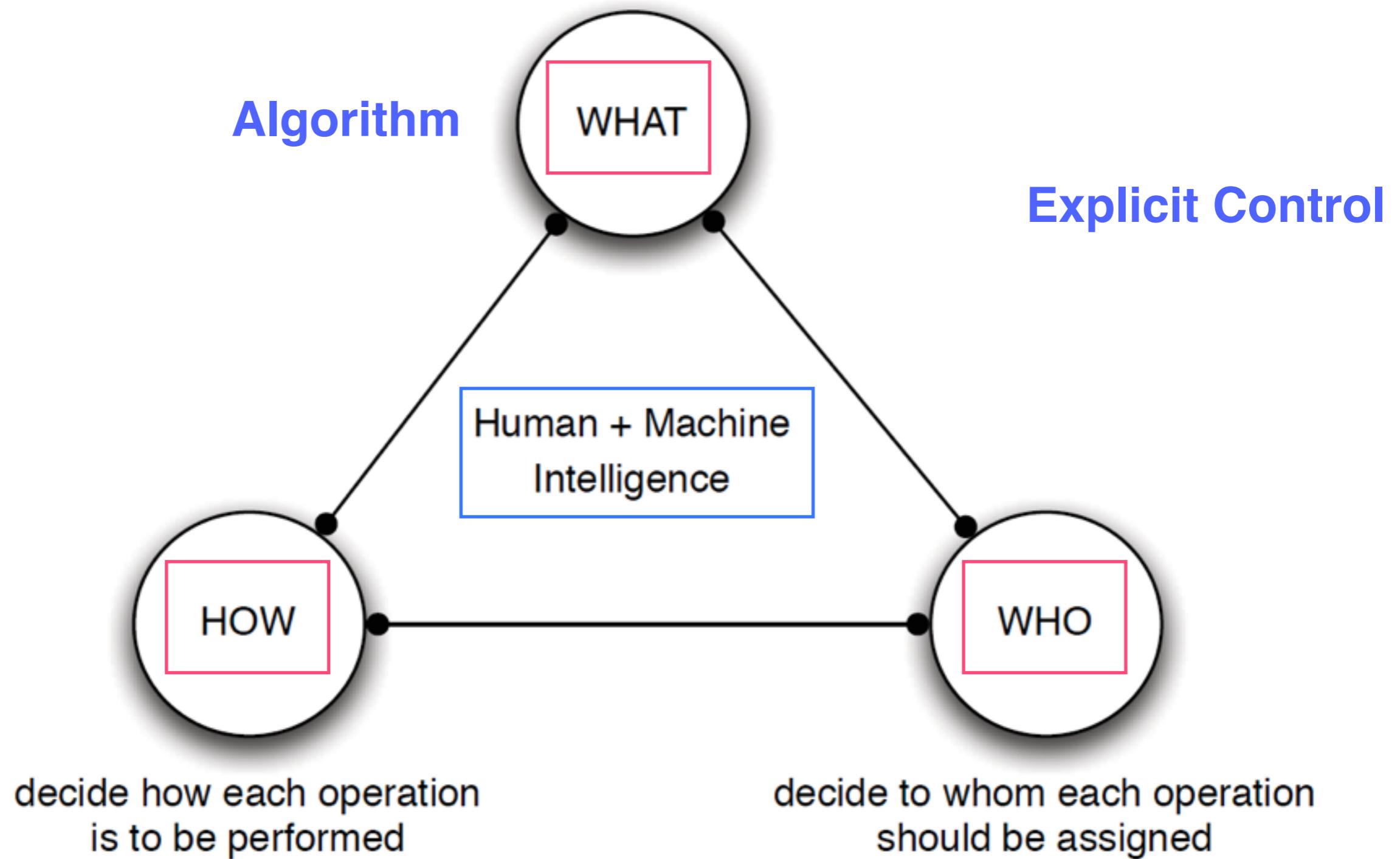
社会
计算

众包



三个层面

decide what operations need to be performed in what order



AI难题

- 识别： 图像识别、语音识别, ...
- 语言： NLP、翻译, ...
- 计算密集问题： 象棋、围棋、NP问题、生物学基因测序, ...
- 自愿计算： SETI@Home, ...
- 其余： 医疗诊断、文本编辑、计划...

Human Computation 算法

算法

Termination → Condition
Outputs →
Sufficiently → Basic
Operations →

Inputs

```
function quicksort(A)
    initialize empty lists L and G
    if (length(A) ≤ 1)
        return A
    pivot = A.remove(find_pivot(A));
    for x in A
        if compare(x, pivot)
            L.add(x)
        else
            G.add(x)
    return concatenate(quicksort(L), pivot, quicksort(G))

function pivot(A)
    return randomIndex(A);

function compare(x, pivot)
    return (x < pivot)
```

Precisely
← Defined
Steps

人计算

```
function quicksort(A)
    initialize empty lists L and G
    if (length(A) ≤ 1)
        return A
    pivot = A.remove(find_pivot(A));
    for x in A
        if compare(x, pivot)
            L.add(x)
        else
            G.add(x)
    return concatenate(quicksort(L), pivot, quicksort(G))

function pivot(A)
    return randomIndex(A);

function compare(x, pivot)
    return human_compare(x, pivot)
```

Mechanical Turk Task

Instructions

You are shown two images. You must select the image that is more indicative of suspicious activities.

Task

Imagine that you are a security guard and you are monitoring two places. Someone informed you that there are suspicious activities in one of the places, but you were not told which one. Which place will you attend to?



Submit

人计算

```
function quicksort(A)
    initialize empty lists L and G
    if (length(A) ≤ 1)
        return A
    pivot = A.remove(find_pivot(A));
    for x in A
        if compare(x, pivot)
            L.add(x)
        else
            G.add(x)
    return concatenate(quicksort(L), pivot, quicksort(G))

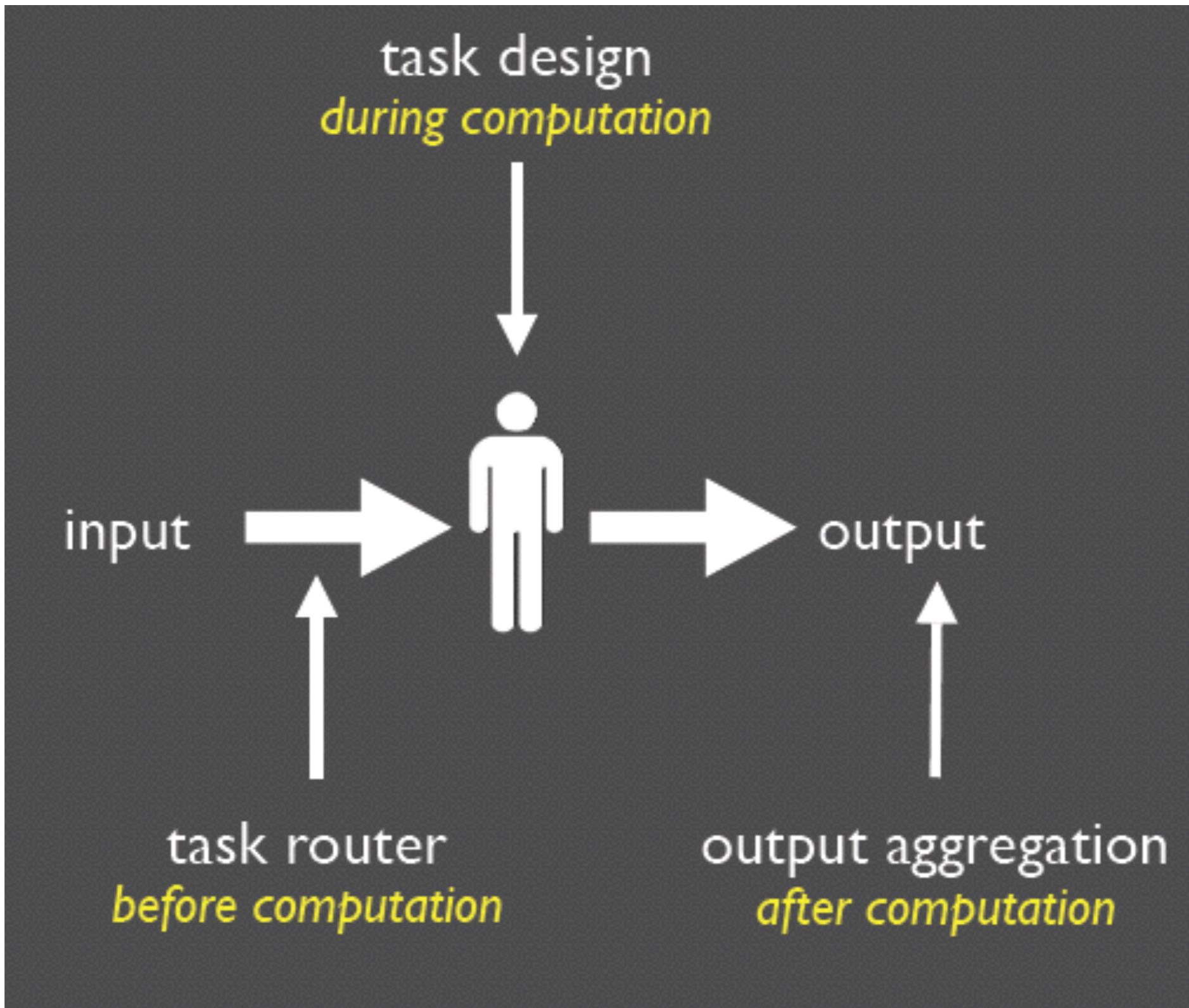
function pivot(A)
    return randomIndex(A);

function compare(x, pivot)
    return human_compare(x, pivot)
```

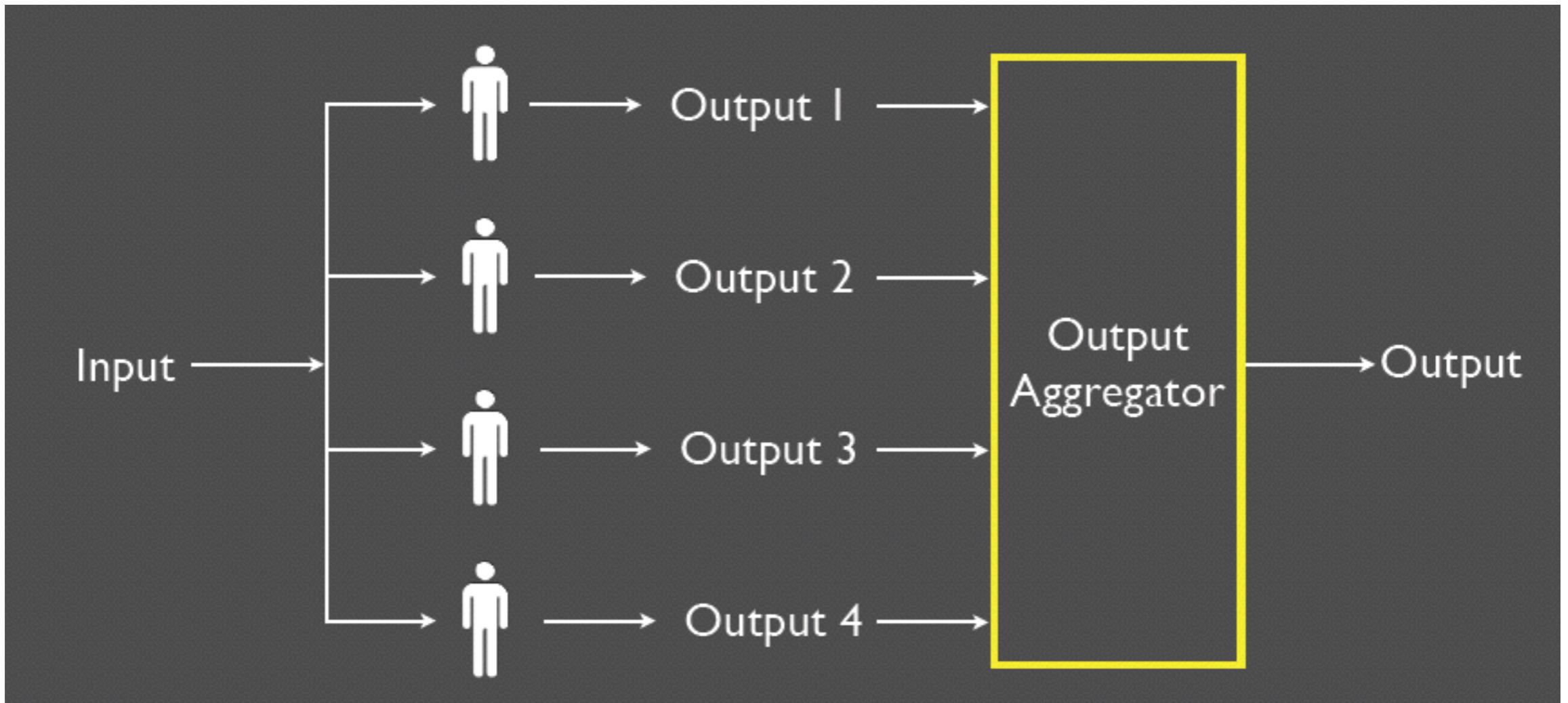
Games with a Purpose



算法正确性



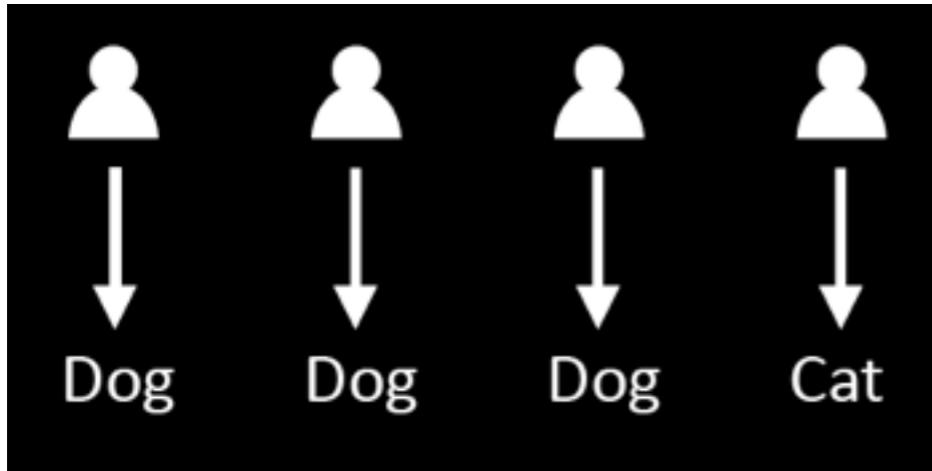
结果汇集



动机

- Money
- Access
- Game
- Volunteer
- Learning

正确性



The King himself directed the operations

Tbe Kiag himself directed tlie operations

The King birnsell directed tlie operations

Tbe King bimself directed tlie operations

The King lala the

himself the

Human Computation例子

ESP: Web图像识别

- Web上的图像识别是一个主要的技术挑战
- 大量图片存在，但是文本描述很少，自动识别很不准确
- 人来做标记是一个无奈的选择

游戏

- 每周现在有20亿用户玩在线游戏
- 21岁美国人万游戏时间，平均一生5年



ESP游戏

- 两个用户同时独立的标记一个图片
- 如果标记一致会得到奖励



Player 1 guesses: purse
Player 1 guesses: bag
Player 1 guesses: brown

Success! Agreement on “purse”



Player 2 guesses: handbag

Player 2 guesses: purse
Success! Agreement on “purse”

Citizen Science

- 科学是数据密集型的
 - * 气候
 - * 物种分布
 - * 星体轨迹
- 科学家数量少
- 非科学家来收集和解释数据

Galaxy Zoo

Galaxy Zoo: Hubble [EN](#)

www.galaxyzoo.org

Galaxy Zoo is a ZOO NIVERSE project ...just like MOON ZOO

GALAXY ZOO HUBBLE

[Home](#) [The Story So Far](#) [How To Take Part](#) [Classify Galaxies](#) [Explore Galaxies](#) [The Science](#) [FAQ](#) [Forum](#) [Blog](#) [Contact Us](#)

Pictures



You aren't logged in so we can't record your results!

Welcome to Galaxy Zoo, where you can help astronomers explore the Universe

Galaxy Zoo: Hubble uses gorgeous imagery of hundreds of thousands of galaxies drawn from NASA's Hubble Space Telescope archive. To understand how these galaxies, and our own, formed we need your help to classify them according to their shapes — a task at which your brain is better than even the most advanced computer. If you're quick, you may even be the first person in history to see each of the galaxies you're asked to classify.

More than 250,000 people have taken part in Galaxy Zoo so far, producing a wealth of valuable data and sending telescopes on Earth and in space

Classifier Log In

[Click here to log in](#)

[Register](#) [Forgotten Password?](#)

Explore galaxies

Enter a search term

eBird

eBird News and Features — [X](#)

ebird.org/content/ebird/ [star](#) [RSS](#) [e-mail](#)

eBird



[Home](#) [About eBird](#) [Submit Observations](#) [View and Explore Data](#) [My eBird](#)

[Sign In](#) | [Register as a New User](#)

Translate to: [English](#)

Welcome to eBird

eBird is as birder does!

eBird News and Features

[eBirder of the Month -- Mary Gustafson](#)
November 12, 2011

The [Rio Grande Valley Birding Festival](#) is taking place this weekend, which means that our current November eBirder of the Month will be buzzing around Harlingen corralling teams of leaders, making sure field trips get off on time, and generally making sure that the 2011 festival comes off as another huge success. Sadly, all this means that Mary won't be getting out in the field as much as she would like. So it seems like an appropriate time to recognize her not only for her hard work as field trip chair for this festival, but also for her unflagging commitment to eBird, both as a user and a reviewer. Moreover, Mary has solidified the Rio Grande Valley Birding festival's commitment to eBird by ensuring that all field trip data are entered into the system, and that all festival participants become aware of the project. This adds up to hundreds of new eBirders each year thanks to this festival! Thanks Mary, and be sure to get out (e)Birding again when things settle down on Monday!

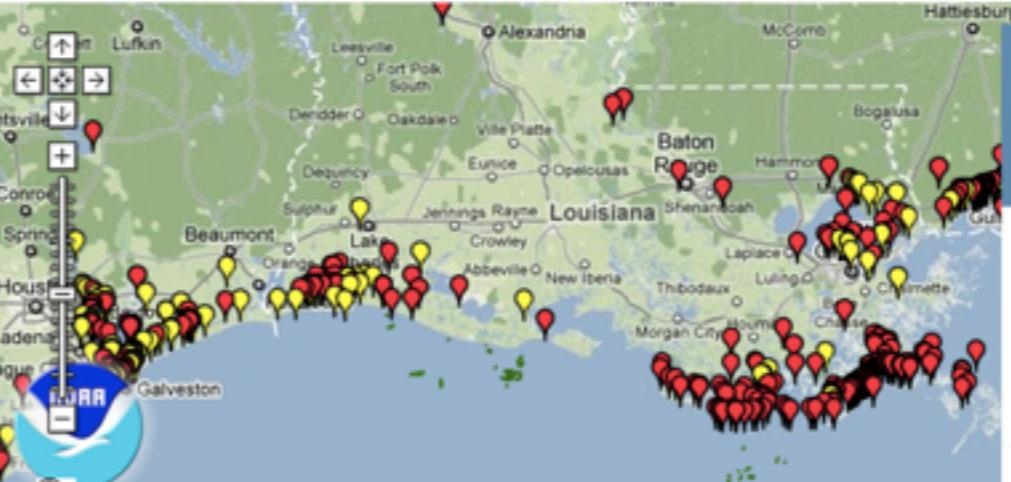


[Possible hybrid Harlan's Red-tailed Hawk x Rough-legged Hawk in Kansas](#)
November 07, 2011

In North America, Red-tailed and Rough-legged Hawks share highly variable plumage, often causing

Gulf Spill Bird Tracker

Oil Slick Forecast: Now 24 hr 48 hr 72 hr



Potomac Oil Slick Footprint - NOAA/NESDIS (Green) observed on 2010-08-05

POWERED BY Google 50 mi 100 km

eBird DataONE [NSF](#)

Maps and data: [Google](#) Current spill location: [NOAA](#) Forecasts: [NOAA-NESDIS](#)

SIGHTINGS

-  Sightings Cluster
-  Within last 30 days
-  Older than 30 days

eBird Rarity Photos Pool



Amazon Mechanical Turk

The screenshot shows the homepage of the Amazon Mechanical Turk website. At the top, there's a browser header with tabs, back/forward buttons, and a URL bar showing <https://www.mturk.com/mturk/welcome>. Below the header is the Amazon Mechanical Turk logo and navigation tabs for 'Your Account' (selected), 'HITs', and 'Qualifications'. To the right, there's a link for existing users to sign in as a 'Worker' or 'Requester'. A blue navigation bar at the bottom of this section includes links for 'Introduction', 'Dashboard', 'Status', and 'Account Settings'. The main content area features a yellow gradient background. It starts with a bold statement: 'Mechanical Turk is a marketplace for work.' followed by a subtext: 'We give businesses and developers access to an on-demand, scalable workforce. Workers select from thousands of tasks and work whenever it's convenient.' Below this, a red call-to-action button says '185,774 HITs available. [View them now.](#)'.

Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

As a Mechanical Turk Worker you:

- Can work from home
- Choose your own work hours
- Get paid for doing good work



or [learn more about being a Worker](#)

Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Register Now](#)

As a Mechanical Turk Requester you:

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results



Amazon Mechanical Turk

[Your Account](#)[HITs](#)[Qualifications](#)**253,412 HITs**
available now[All HITs](#) | [HITs Available To You](#) | [HITs Assigned To You](#)

Find containing that pay at least \$ for which you require Mast

All HITs

1-10 of 32004 Results

Sort by:

GO![Show all details](#) | [Hide all details](#)

Classify Arabic Tweets Dialects SEE REVISED HIT

Requester: [Chris Callison-Burch](#)

HIT Expiration Date: Aug 4, 2013 (9 weeks 6 days) **Reward:** \$0.00

Time Allotted: 60 minutes **HITs Available:** 23530

Identify whether the phrase/keywords belongs to the category provided?

Requester: [InterestProfiler](#)

HIT Expiration Date: Jun 21, 2013 (3 weeks 4 days) **Reward:** \$0.01

Time Allotted: 10 minutes **HITs Available:** 17969

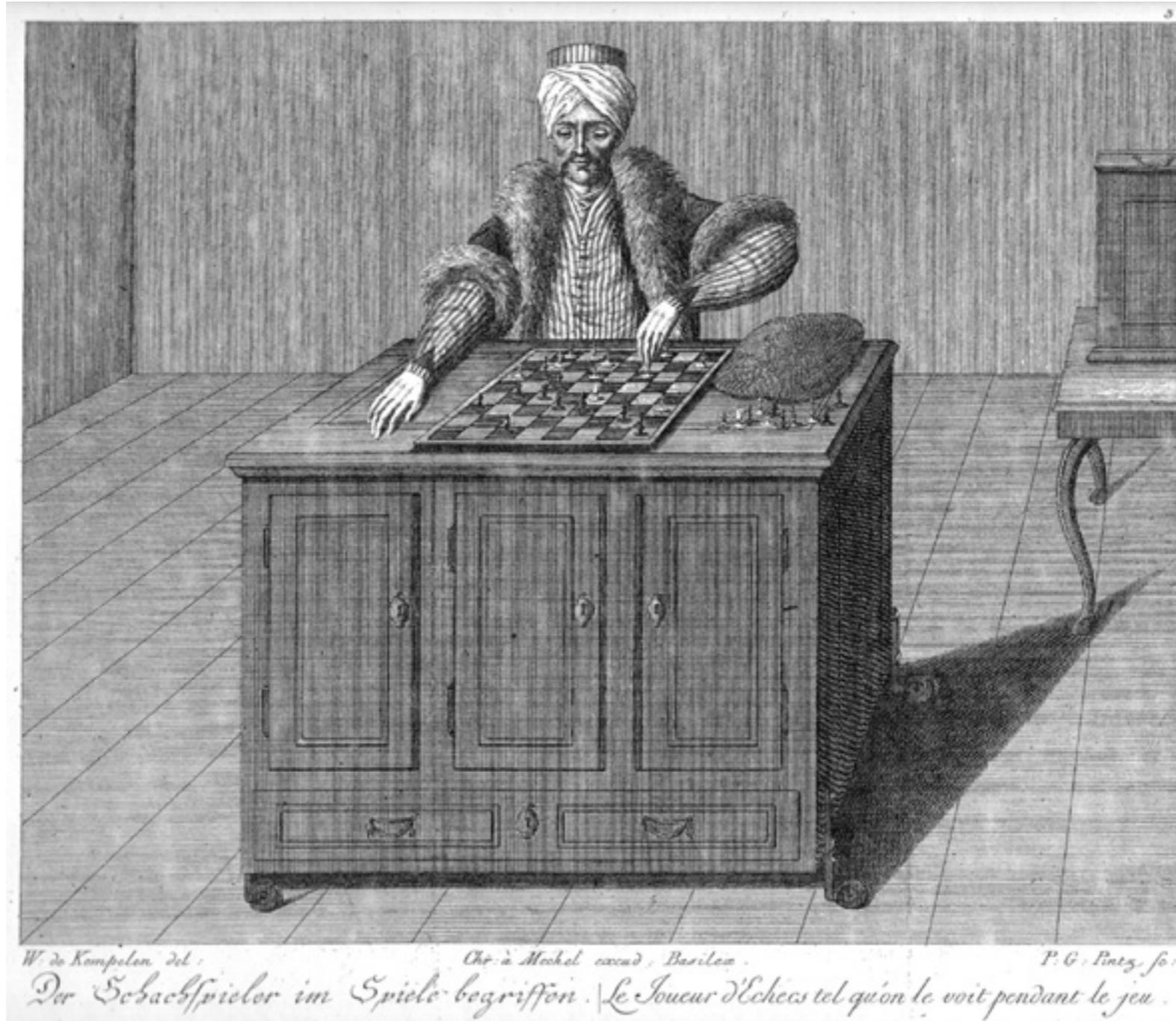
Search: Keywords on Google.com (US)

Requester: [CrowdSource](#)

HIT Expiration Date: May 27, 2014 (52 weeks) **Reward:** \$0.08

Time Allotted: 16 minutes **HITs Available:** 14969

Mechanical Turk



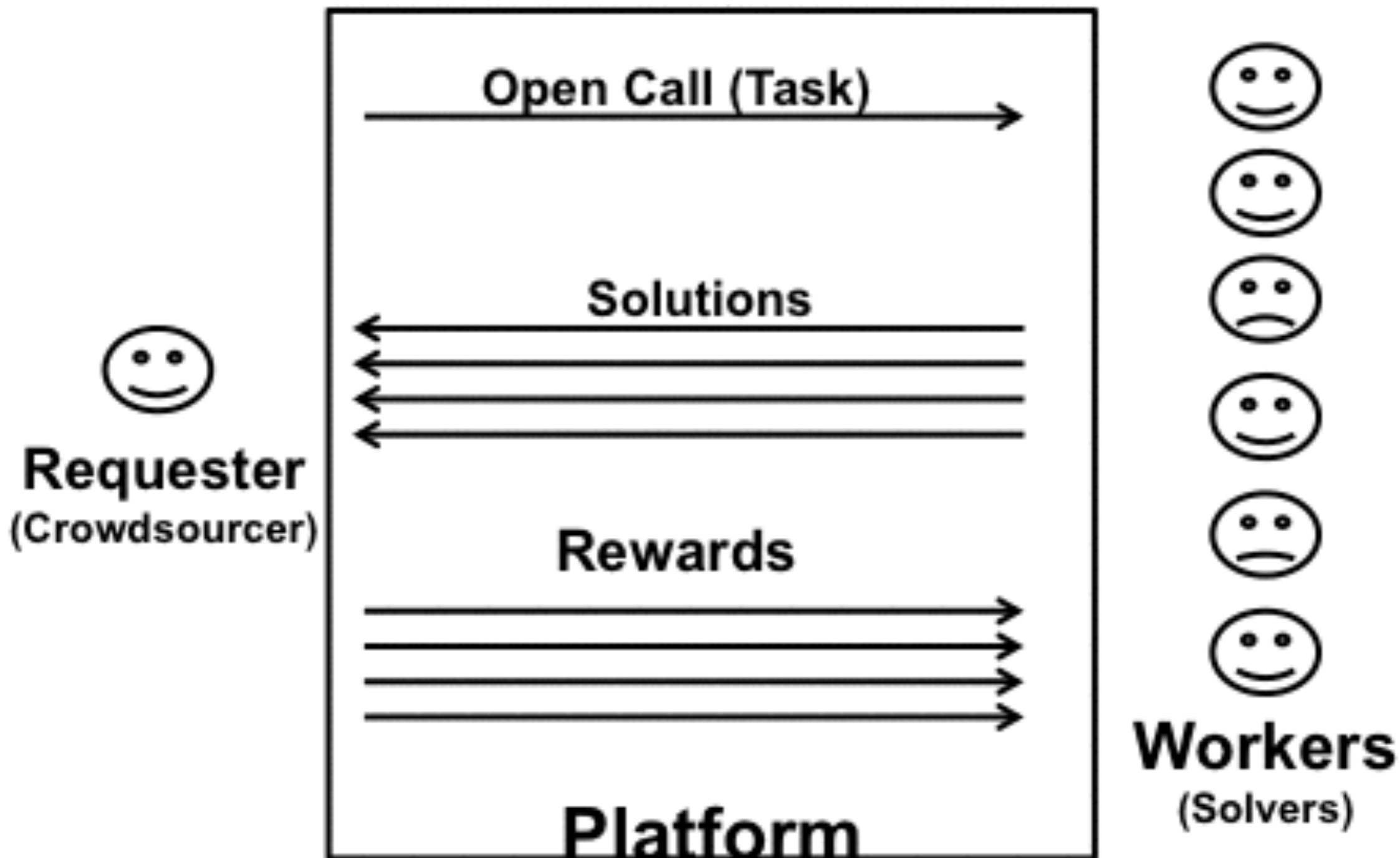
W. de Kompolen del.

Ohr à Mechel excud. Basilea.

P. G. Pintz. sc.

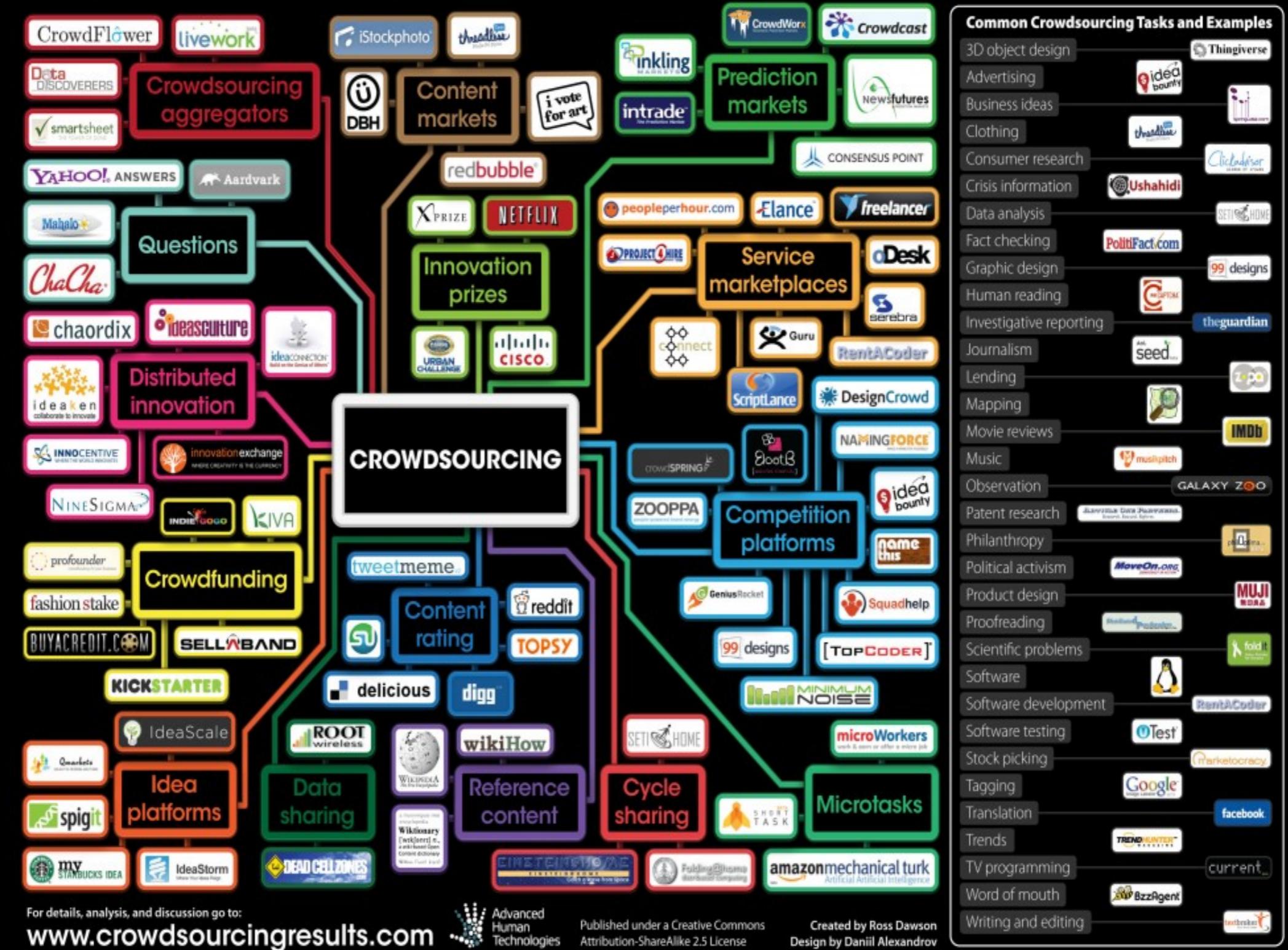
Der Schachspieler im Spiele begriffen. | Le Jouer d'Echecs tel qu'on le voit pendant le jeu.

众包



众包

CROWDSOURCING LANDSCAPE Beta v1

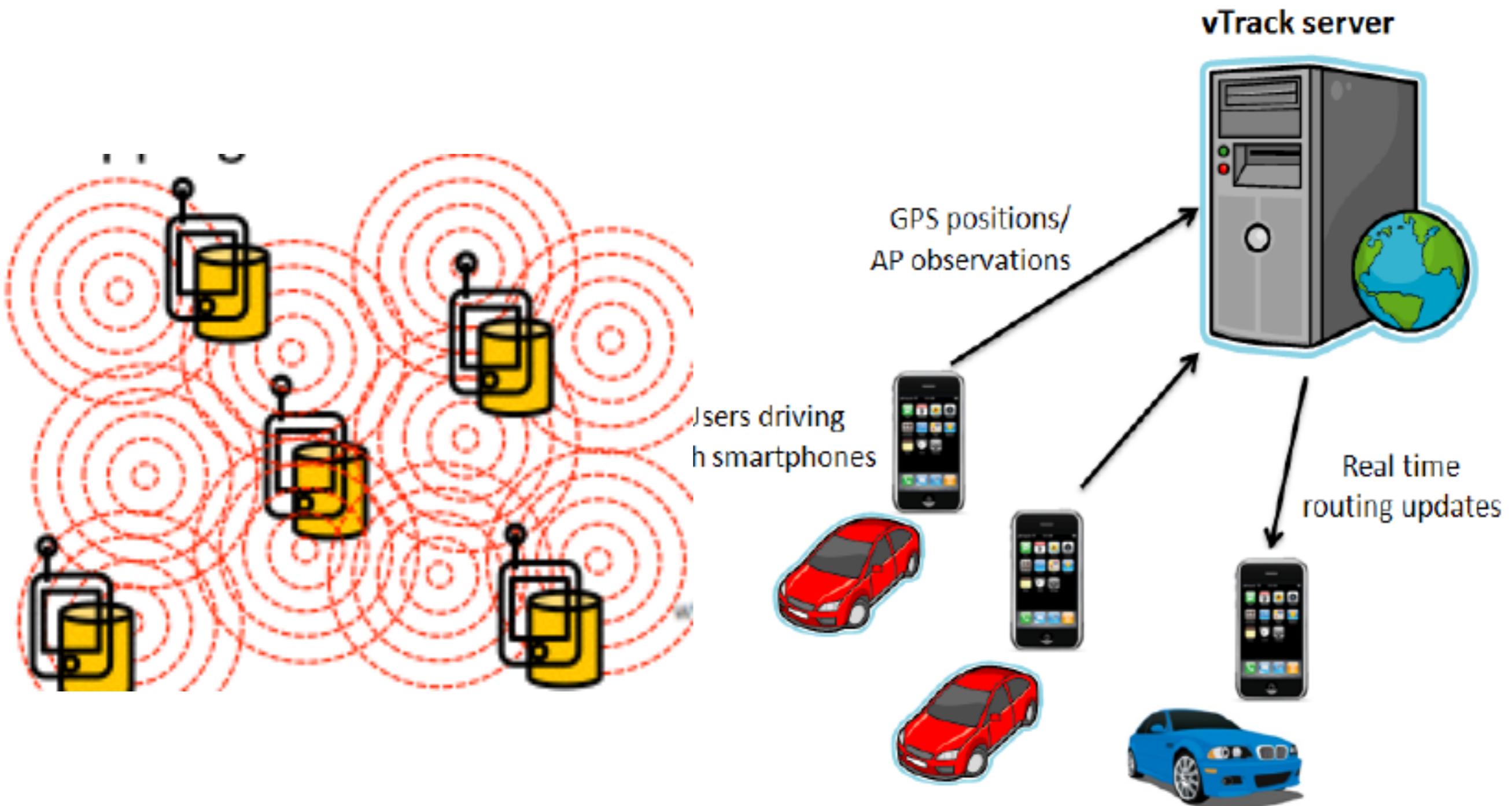


道路监控

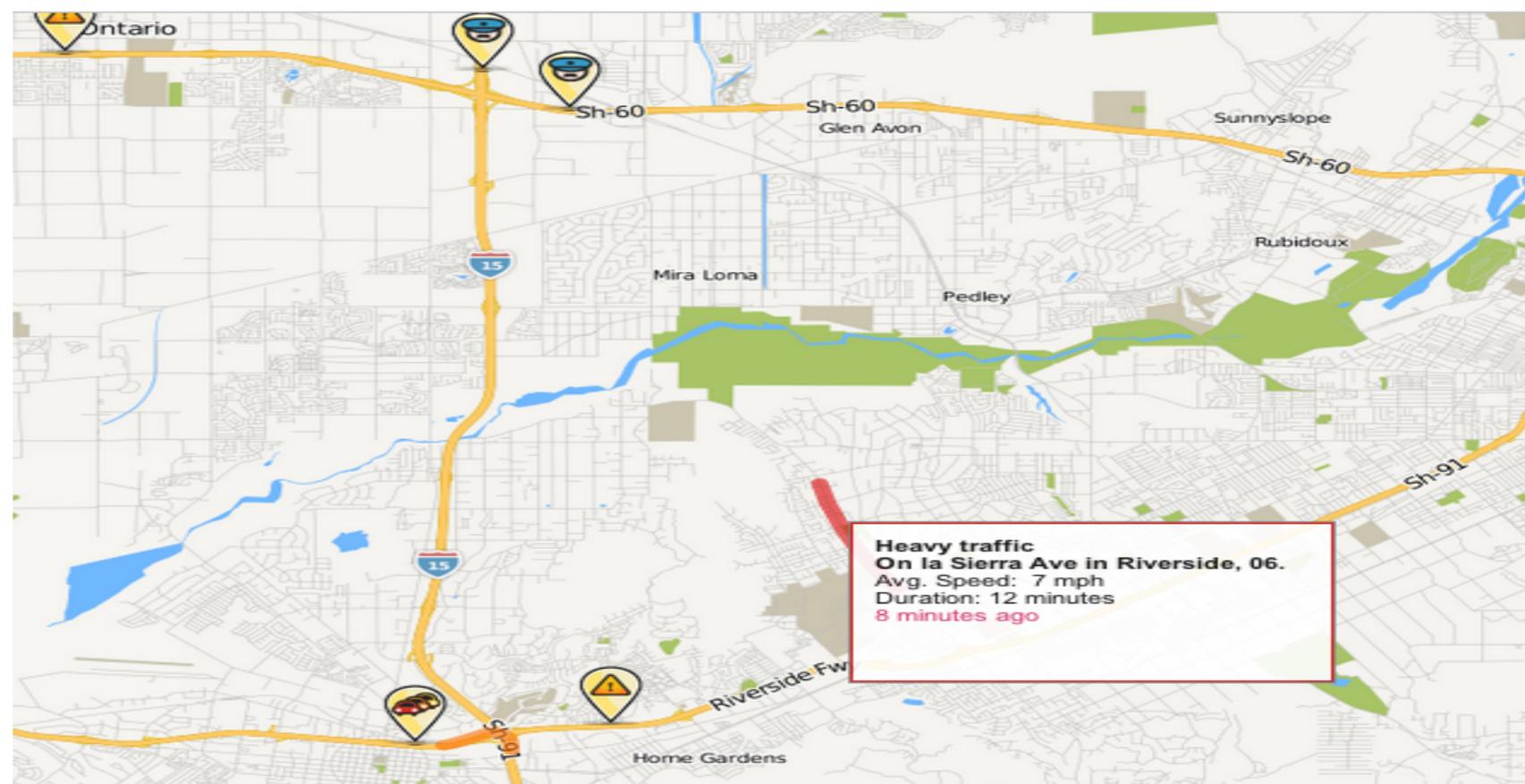


This screenshot shows the Transport Roads & Traffic Authority (NSW) traffic monitoring interface. At the top, there's a browser toolbar with tabs like LOGI, MSDN, BIBLOG, BANNER, FinanceApp, KEDIMA, EPSS, BILLS, and Other Bookmarks. The main header features the NSW Government logo and the Transport Roads & Traffic Authority branding. Below the header is a map of the Sydney metropolitan area, with major roads and local streets labeled. Numerous black camera icons are placed along the roads, indicating active traffic surveillance. A legend on the right side, titled 'Traffic hazards', lists various types of incidents: Incidents (6), Scheduled road work (116), Major events (4), Fire (0), Flood (5), and Alpine conditions (0). The 'F3 critical alerts' and 'RSS' buttons are also visible at the top right. The bottom of the page includes standard footer links for Privacy, Copyright, Terms, and Sitemap, along with a 'Print map' button.

道路监控



道路监控



- 要求阅读Human Computation. Part II Design:

→ *Understanding Workers and Requesters*

→ *The Arts of Asking Questions*

下次上课测试!

谢谢！

孙惠平

sunhp@ss.pku.edu.cn