

2019年10月24日

第1回 SPARC Japan セミナー2019

「人文社会系分野におけるオープンサイエンス ～実践に向けて～」

# オープンサイエンス的 市民協働のために 大学ができること

京都大学情報環境機構

小野 英理

京都大学



# 自己紹介

- 専門： 霊長類学、学位：博士(理学)
- 経験： 研究とデザイン(Web)
- 2015～2017年 京都大学・URA(若手研究者支援)
- 2018年～現在 京都大学Web戦略室・特定講師
- 2016年よりKYOTOオープンサイエンス勉強会を主催
- 主に市民参加型研究プロジェクトの調査
- 月1で演者を招待、現在30回以上開催

## ボランティア

Wikipedia

オープンデータで  
ハッカソン

手動発電機

みんなで翻刻

歴史学カフェ

魚類

歴史教育

ナメクジ

雷雲γ線

マルハナバチ

ゲーム

ヒアリ

地下探査

## アカデミア

KYOTO UNIVERSITY

オープンサイエンス政策

URA

シビックテック

放射線測定

ゴジラ

・災害対策政策

・細胞生物学

関東雪結晶

シチズン  
サイエンスの倫理

サイエンスの意味

学校健診

研究データマネジメント

クラウドファンディング

ほとんど0円大学

いきものコレクションアプリ

ウェザーニュース

ハマるゲームの作り方

科学雑誌ニュートン

プロジェクト管理

デバイス開発

## 政策・体制 インフラ

## 企業

# オープンサイエンスを 分かりやすく捉えるなら…

情報通信技術 (ICT) の発展  
で可能になる、

- オープンアクセス
- オープンデータ
- オープンコラボレーション

によって、

学術研究の透明性、協働、  
イノベーションを促進する。

} 情報基盤

→ 多くの協働タイプのうち  
変化の好例として  
「シチズンサイエンス」

# 市民が学術研究に関わる機会が拡大



# オープンサイエンス的市民参加型研究 (シチズンサイエンス(CS)、市民科学)



eBird



# Shakespeare's World

2015年12月8日～2019年10月4日

- シェイクスピアと同時期の文書を多数集め、アノテーションおよびテキスト起こしを行う。シェイクスピアとその朋輩の人生について新しい発見があれば、オックスフォード英語辞典に反映される。(Wikipedia)

- コラボレーション

- Folger Shakespeare Library
- Zooniverse.org at Oxford University
- Oxford English Dictionary of Oxford University Press



<https://www.zooniverse.org/projects/zooniverse/shakespeares-world>

# CSに係る政策

日本: 第5期科学技術基本計画(2016-2020年度)

- (2)③オープンサイエンスの推進  
「市民参画型のサイエンス(シチズンサイエンス)が拡大する兆しにある。」

G7/日本学術会議(2019年)

- Gサイエンス学術会議共同声明:「インターネット時代のシチズンサイエンス」が含まれ,今後日本においてCSを推進するための環境整備の必要性を指摘

欧州: ホライズン2020

- 総額800億ユーロ(約10兆円)の一部をCSの助成に充てる\*

# 本日の発表について

- (人社系に限らず) 多分野の市民協働プロジェクトを調査してきた経験に基づく現状(関係資料)の紹介
  - 特に日米のプロジェクト比較や参加者の調査
  - 活動はURAの業務ではない
- 私見を交えた将来像の検討
  - 特に大学関係者が市民協働型研究にどう関わるのか

# 本日の主題

Sparc Japan セミナーウェブサイトより

- 「人文社会系分野におけるオープンサイエンス ～実践に向けて～」
- 「研究者によるデータ構築と市民科学との間を繋ぐ媒介者としての役割を担いうるURA (リサーチ・アドミニストレーター) の取組についても論じていただきます。」

## 論点

1. Why: 大学としてCSに取り組むべき？
2. Who: 誰が担うのか？
3. How: どうやって実践する？

# Why: 大学としてCSに取り組むべき？

「誰が担うのか」に関係するが、URAをはじめとした研究推進/支援人材が業務として取り組むためには、CSが「組織の方針」に合致するべき

そもそもシチズンサイエンスは何のために行うのか？



Designed by studiogstock / Freepik

# シチズンサイエンスの目的

| 目的                              | 内容                                                                                                                                                            |
|---------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 研究                              | <ul style="list-style-type: none"><li>• データを得る</li><li>• データを処理する 他</li></ul>                                                                                 |
| 教育                              | <ul style="list-style-type: none"><li>• 非公式科学教育 (Informal Science Education) <sup>*1</sup></li><li>• 生涯学習</li></ul>                                           |
| パブリック<br>エンゲージメント <sup>*2</sup> | <ul style="list-style-type: none"><li>• 包括的かつ参加型の研究テーマの設定</li><li>• 社会課題への対処に関して異なる視点を得る</li><li>• 政策立案者、市民、研究者間の相互理解を促進し、科学政策における社会的なコンセンサスを強化する 他</li></ul> |

1. National Research Council. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. Washington, D. C.: National Academies Press.

2. OECD (2017), *Open research agenda setting*, Science, Technology and Industry Policy Papers, No. 50, OECD Publishing, Paris, <https://doi.org/10.1787/74edb6a8-en>.

# CSにおける市民の関与

|            | 貢献型<br>Contributory | 協働型<br>Collaborative | 共創型<br>Co-created |
|------------|---------------------|----------------------|-------------------|
| 課題設定       |                     |                      | ○                 |
| 情報収集       |                     |                      | ○                 |
| 仮説立案       |                     | ○                    | ○                 |
| データ収集の方法設計 |                     | ○                    | ○                 |
| サンプル収集     | ○                   | ○                    | ○                 |
| サンプル分析     | ○                   | ○                    | ○                 |
| データ分析      |                     | ○                    | ○                 |
| データ/結果を説明  |                     |                      | ○                 |
| 成果を広報      |                     |                      | ○                 |
| さらなる研究への議論 |                     |                      | ○                 |

Phillips, T. B., Ferguson, M., Minarchek, M., Porticella, N., and Bonney, R.  
 2014. *User's Guide for Evaluating Learning Outcomes in Citizen Science*.  
 Ithaca, NY: Cornell Lab of Ornithology. のTable 1を改変(翻訳)

# 市民参加の類型：Haklayの分類

1. Crowdsourcing: センサーの役割
2. Distributed intelligence: 基礎的な解釈
3. Participatory science: 問題設定やデータ収集
4. Extream/Up-science: 問題設定やデータの収集・分析

注) 参加程度が高いほど良いという意味ではない

Haklay M. Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: Crowdsourcing Geographic Knowledge. Dordrecht: Springer Netherlands; 2013:105-122. doi:10.1007/978-94-007-4587-2\_7

# パブリックエンゲージメント：CSとSDGs



nature  
sustainability

Perspective

| Published: 09 October 2019

## Citizen science and the United Nations Sustainable Development Goals

Steffen Fritz , Linda See, [...] Sarah West

*Nature Sustainability* **2**, 922–930 (2019) | [Download Citation](#) 

<https://www.nature.com/articles/s41893-019-0390-3>

# アイデア:「市民協働」を学術研究のユニバーサルデザインとして考える

誰もが研究に参加しやすい  
環境を目指す

||

従来のステークホルダーに  
とってより活動しやすい  
環境となる

著作権保護のため公開資料から割愛

ここまでの概括

# Why: 大学としてCSに取り組むべき？

現状

- CSの目的は研究、教育、パブリックエンゲージメントなど、多様

将来像

- 研究者(運営者)は計画時に目的を見極めることが望ましい
- それが各機関の目的に合致するならば、組織として推進/支援することは妥当

# Who: URAがシチズンサイエンスの媒介者？

## 京都大学の状況

- URAは「研究力強化」のため(研究大学強化促進事業)
  - CSが「研究力強化」につながるか自明でない(という空気感)
  - 現在のURA業務に市民協働の構築/実践は含まれない  
(個別ケースとして「みんなで翻刻」の広報を支援した事例あり)
- 「現在のURA」が業務としてCSを実施することは考えにくい

そもそもオープンサイエンスの担い手は誰か？

# オープンサイエンスの担い手

- 研究者
- 省庁
- 助成機関
- 大学等教育研究機関
- 図書館/リポジトリ/データセンター
- NPO/NGO
- 出版社
- 産業界



Designed by studiostock / Freepik

# 大学等教育研究機関の役割

- オープンサイエンスに関するポリシー策定
- 必要なスキルについて研究者をトレーニング
  - リポジトリの使い方
  - データの扱い(クリーニング、キュレーション、管理)



# 図書館/リポジトリ/データセンターの役割

- オープンサイエンスに向けて重要かつ実現に不可欠なアクター
- デジタルな研究資源、つまり出版物やデータ等の研究関連のコンテンツを保存し、キュレーションし、公開し、広める
- 研究者が成果をシェア、利用、再利用する為のインフラを構築し、オープンサイエンスを実現させる



# CSの体制：日本(1) - ナメクジ捜査網

- 2016年から開始
- 運営者：宇高寛子助教
  - Webサイト(Googleサイト)を制作
  - 全ての写真をチェック
  - Twitter、メール等でメッセージに返信

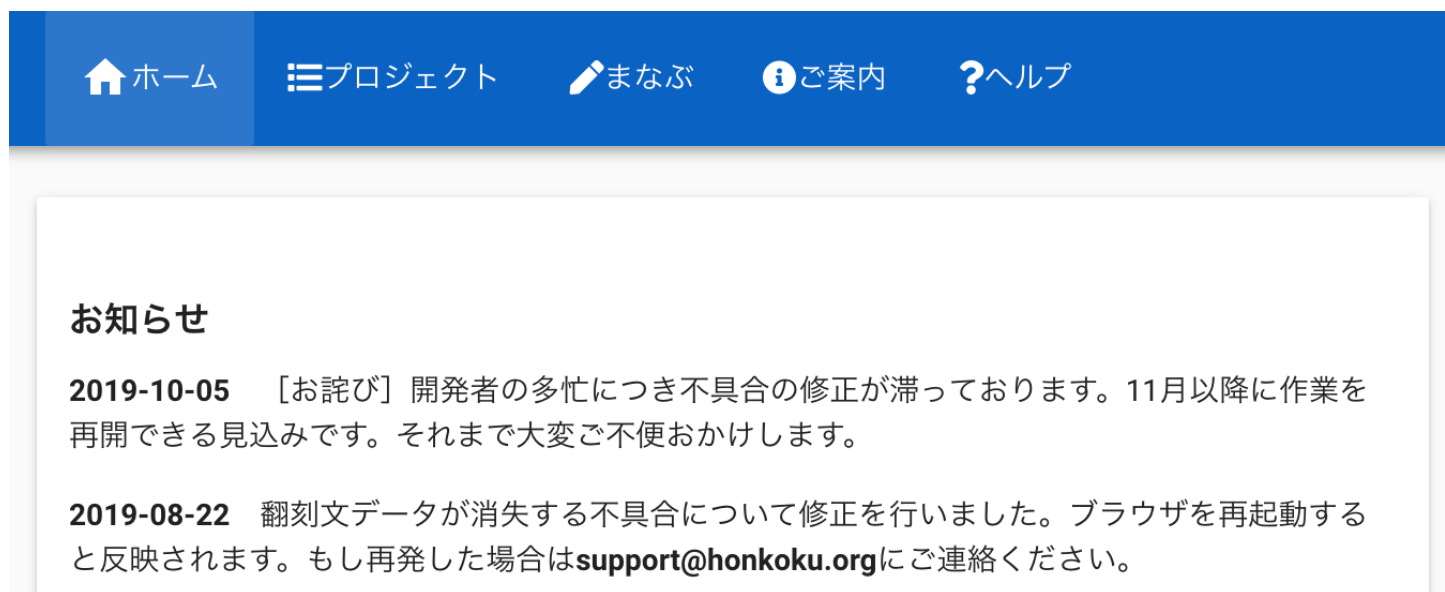
ひとりで可能な範囲で実施



<https://sites.google.com/site/udakawebbsite/madarakouranamekuji-limax-maximus>

# CSの体制：日本(2) - みんなで翻刻

- (少なくとも最初のバージョンは)ひとりの研究者が開発



The screenshot shows the top navigation bar of the Honkoku Project website. It has a blue background with white text and icons. The navigation items are: ホーム (Home) with a house icon, プロジェクト (Project) with a list icon, まなぶ (Learn) with a pencil icon, ご案内 (About) with an information icon, and ヘルプ (Help) with a question mark icon. Below the navigation bar is a white box with a grey border containing an announcement section titled 'お知らせ' (Notice). The announcement contains two entries: one dated 2019-10-05 about a delay in corrections, and another dated 2019-08-22 about a fix for missing data and a contact email.

ホーム プロジェクト まなぶ ご案内 ヘルプ

## お知らせ

**2019-10-05** 〔お詫び〕 開発者の多忙につき不具合の修正が滞っております。11月以降に作業を再開できる見込みです。それまで大変ご不便おかけします。

**2019-08-22** 翻刻文データが消失する不具合について修正を行いました。ブラウザを再起動すると反映されます。もし再発した場合は[support@honkoku.org](mailto:support@honkoku.org)にご連絡ください。

ひとりではプロジェクト運営に影響が出ることも

# CSの体制：日本(3) - 花まるマルハナバチ国勢調査

- 富士通の携帯フォト・クラウドシステム 無料提供期間終了
- 携帯アプリ「ここピン!」活用 サービス終了
- 写真の同定は研究者1名

研究者が何でもやる  
体制には限界がある

大野ゆかり, 第20回KYOTOオープン  
サイエンス・ミートアップ, 2018

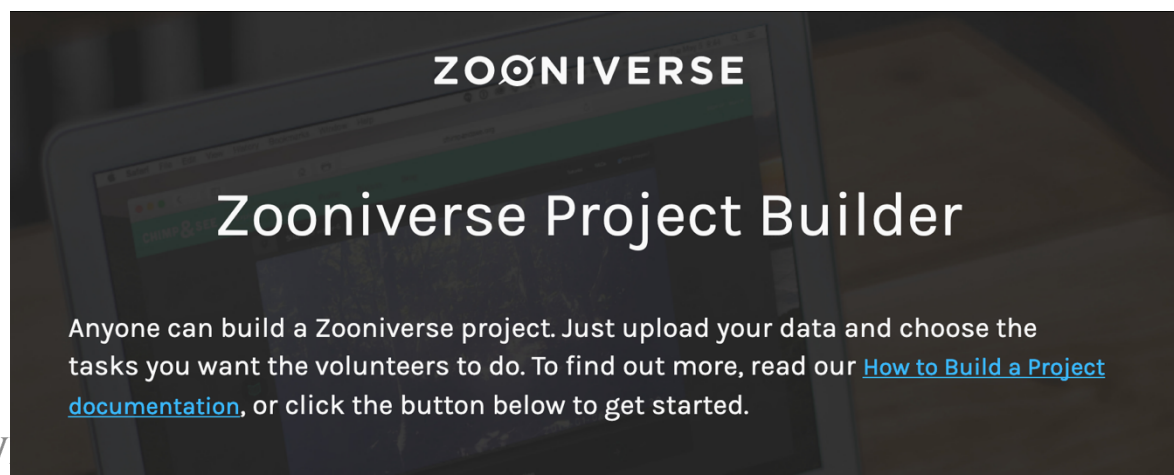
研究者だって、市民の皆さんの期待に応えたい！  
でも、できない！



研究者は、研究に特化した人間です。マネジメントや普及活動に長けているわけではありません。決して怠けているわけではありません。  
努力しますので、どうぞ暖かく見守ってください…。

# CSの体制：米国(1) - Zooniverse

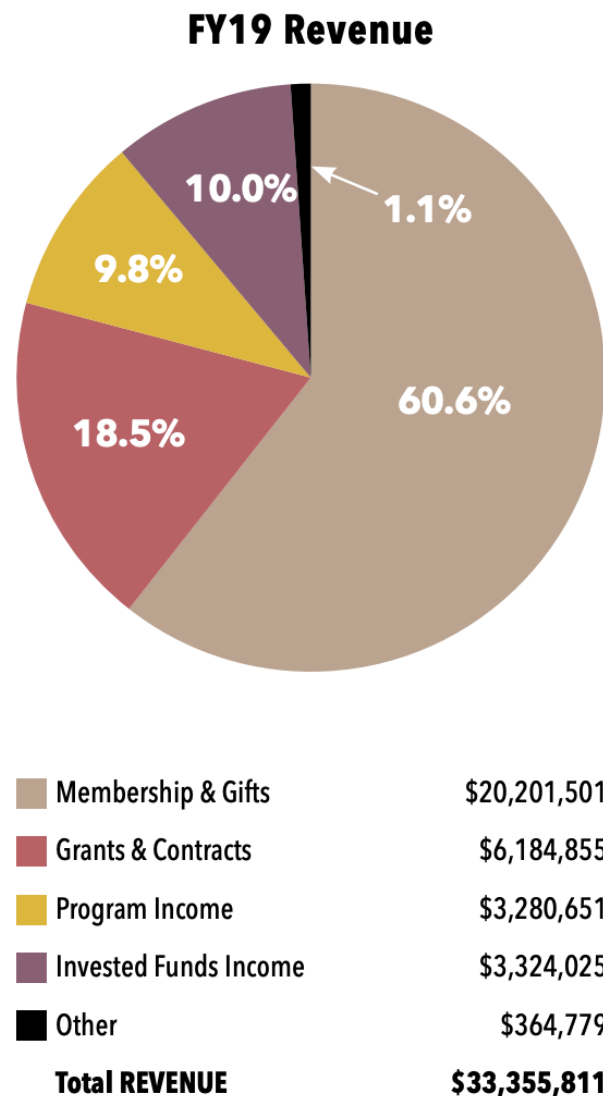
- 複数のCSを統合するプラットフォーム(e.g. Galaxy Zoo)
- 運営：Citizen Science Alliance(オクスフォード大など9つの機関が参画)
- 簡単な画像分析であれば分野を問わず構築可能
- スタッフ:34名 2019年10月現在



# CSの体制：米国(2) - Cornell Lab of Ornithology

- 市民による世界中の鳥類の観察記録を集約(eBird)
- コーネル大学, 全米オーディュボン協会によって2002年に設立<sup>\*1</sup>
- 年間予算：約35億円<sup>\*2</sup>  
収入：約36億円
- スタッフ：200名以上  
2019年10月現在

1. Wikipedia  
2. The Cornell Lab of Ornithology. "2019 Annual Impact Report". p. 8. 2019.



# CSの体制：米国(2) - Cornell Lab of Ornithology

## スタッフの職種

- Director/Manager/Leader
- Developer/Engineer/Programmer
- Designer
- Programs Coordinator
- Analyst
- Researchers/Postdoctoral Fellow
- Services Associate 等々

## 図書館/博物館との連携

- データ(音声、画像、動画)をマコーレー図書館\*にアーカイブ  
\*市民科学者 マコーレー夫妻に因んで命名
- 標本をコーネル大学脊椎動物博物館に保存

# Who: 誰がCSを担うのか

## 現状

- 日本では研究者の個人的努力に委ねられている
- 海外の成功事例は豊富な資金と組織的な活動が基盤

## 将来像

- 大学においてURA(のような組織/部署間連携を推進する者)が組織的基盤を構築することは妥当かもしれない
- 情報インフラやウェブサイトの構築・維持において図書館/リポジトリ/データセンターの役割は重要
- 資金の問題はどうやって解決するのか…

# How: どうやって実践する？

1. 参加者を知る
  - 属性
  - 動機
2. データの質を担保する
3. プロジェクト設計のために評価を行う



Designed by macrovector / Freepik

# 参加者の属性

年齢、性別 (Zooniverse: 生態学5プロジェクト、天文学5プロジェクト)

| Age range (years) | Ecology page views | %     | Astronomy page views | %     | Odds ratio | Lower 95% CI | Upper 95% CI | p value   |
|-------------------|--------------------|-------|----------------------|-------|------------|--------------|--------------|-----------|
| 18–24             | 1242               | 13.11 | 2215                 | 22.15 | 1.89       | 1.75         | 2.04         | 8.86E-62  |
| 25–34             | 1971               | 20.82 | 1677                 | 16.78 | 0.77       | 0.71         | 0.82         | 5.67E-13  |
| 35–44             | 1504               | 15.89 | 1566                 | 15.66 | 0.98       | 0.91         | 1.06         | 6.80E-01  |
| 45–54             | 1160               | 12.25 | 1096                 | 10.96 | 0.88       | 0.81         | 0.96         | 5.11E-03  |
| 55–64             | 2826               | 29.85 | 1418                 | 14.18 | 0.39       | 0.36         | 0.42         | 4.77E-156 |
| 65+               | 765                | 8.08  | 2025                 | 20.26 | 2.89       | 2.64         | 3.16         | 4.16E-134 |

| Sex    | Ecology project views | %     | Astronomy project views | %     | Odds ratio | Lower 95% CI | Upper 95% CI | p value |
|--------|-----------------------|-------|-------------------------|-------|------------|--------------|--------------|---------|
| Male   | 4469                  | 43.93 | 8127                    | 75.45 | 3.92       | 3.7          | 4.16         | 0       |
| Female | 5703                  | 56.07 | 2644                    | 24.55 | 0.25       | 0.24         | 0.27         | 0       |

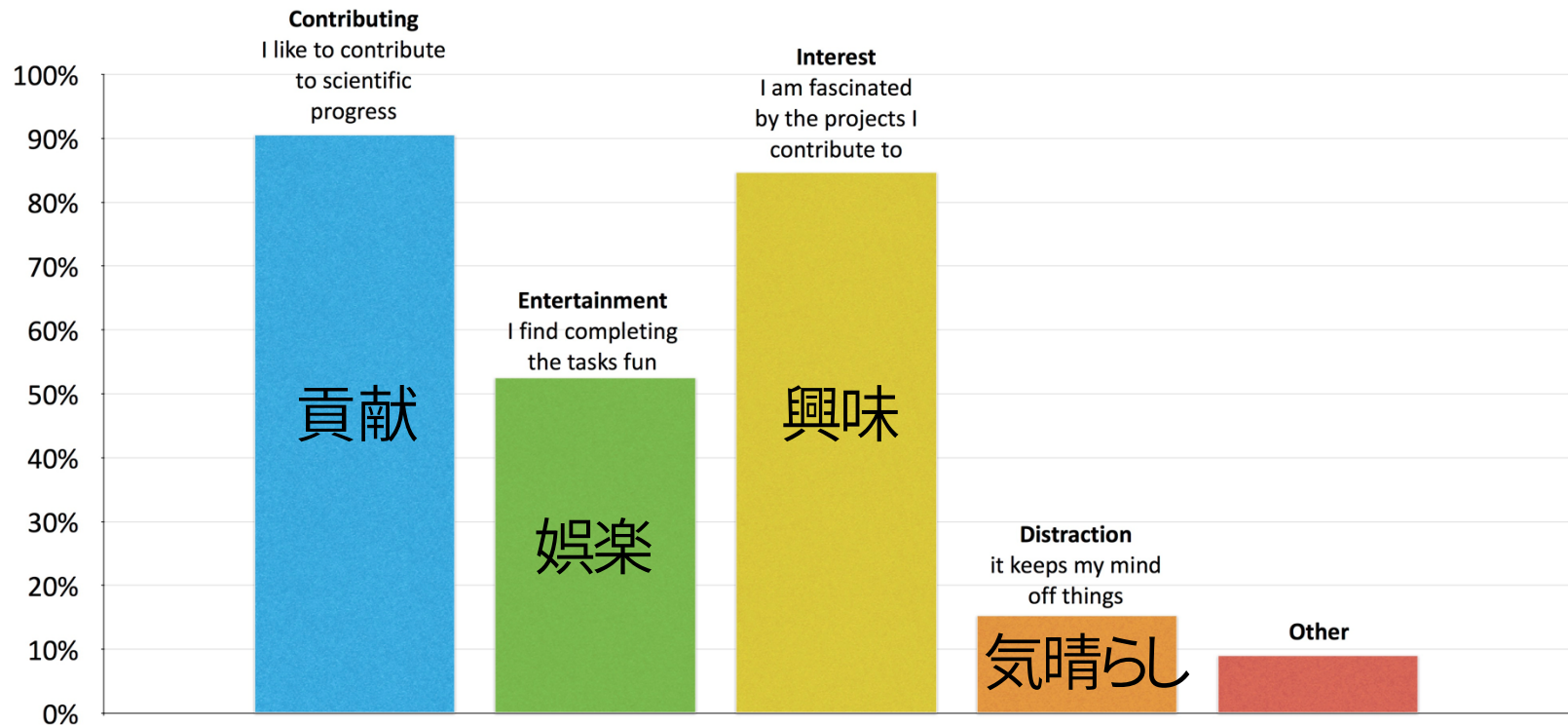
分野によって参加者の属性は異なる

Spiers H, Swanson A, Fortson L, et al. Everyone counts? Design considerations in online citizen science. J Sci Commun. 2019;18(01):A04. doi:10.22323/2.18010204

# 参加者の動機

Zooniverseオンラインアンケート(回答数300)

WHO ARE THE ZOONIVERSE  
COMMUNITY? WE ASKED THEM...,  
2015, <https://blog.zooniverse.org/>



# スーパーボランティア super-volunteer

- 通常の参加者よりも熱心に参加

- Galaxy Zoo

4-7%のユーザが全体の85%のタスクを実施

Sauermann H, Franzoni C. Crowd science user contribution patterns and their implications. Proc Natl Acad Sci U S A. 2015;112(3):679-684. doi:10.1073/pnas.1408907112

- みんなで翻刻

約3%のユーザが全体の最大88%を翻刻(2017年10月時点)

- 初心者への解説・ワークショップ・入門講座

Ono E, Ikkatai Y, Enoto T. Review of six Japanese online citizen science projects with main indicators to encourage citizen motivation. Int J Institutional Res Manag. 2018;2(1):19-34.

# データの質を担保する

データ収集型の60プロジェクトにおける質を担保する取り組み

Table I  
VALIDATION METHODS REPORTED

## 上位5項目

1. 専門家のチェック
2. テキストだけでなく  
写真を提出
3. オンラインだけでなく  
紙でもデータを提出
4. 他の参加者による評価
5. トレーニングプログラム

| Method                                              | <i>n</i> | Percentage |
|-----------------------------------------------------|----------|------------|
| Expert review                                       | 46       | 77%        |
| Photo submissions                                   | 24       | 40%        |
| Paper data sheets submitted along with online entry | 20       | 33%        |
| Replication or rating, by multiple participants     | 14       | 23%        |
| QA/QC training program                              | 13       | 22%        |
| Automatic filtering of unusual reports              | 11       | 18%        |
| Uniform equipment                                   | 9        | 15%        |
| Validation planned but not yet implemented          | 5        | 8%         |
| Replication or rating, by the same participant      | 2        | 3%         |
| Rating of established control items                 | 2        | 3%         |
| None                                                | 2        | 3%         |
| Not sure/don't know                                 | 2        | 3%         |

Wiggins A, Newman G, Stevenson RD, Crowston K. Mechanisms for Data Quality and Validation in Citizen Science. In: 2011 IEEE Seventh International Conference on E-Science Workshops. IEEE; 2011:14-19. doi:10.1109/eScienceW.2011.27

# データの質を担保する

データ分析型 (Galaxy Zooにおける例)

1. 明らかな偽データを削除し、重みづけされていないデータを作成
2. 多数派と同じ分析をするユーザに重み付け
3. 10回以上分析されたデータのうち、80%または95%のユーザが同じ回答をしたデータを抽出 (最終的には80%を解析に使用)
4. 先行研究にてすでに分析された結果と今回の分析結果を照合 (銀河のタイプによって85~99.9%の正答率)
5. 銀河の形態と色の関係で生じる可能性のあるバイアスを調査

# プロジェクト設計のための評価：目的

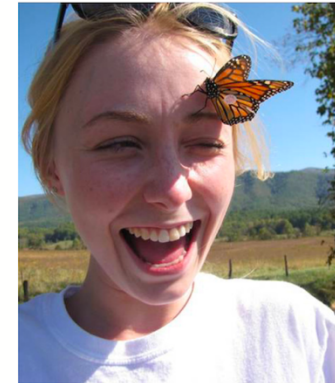
- 多くのプロジェクトで評価が行われていない

## 評価の目的

- プロジェクトの強み、弱みを見つける
- 参加者のニーズを知る
- プロジェクトの成功をステークホルダーに示す
- さらなる資金獲得につなげる

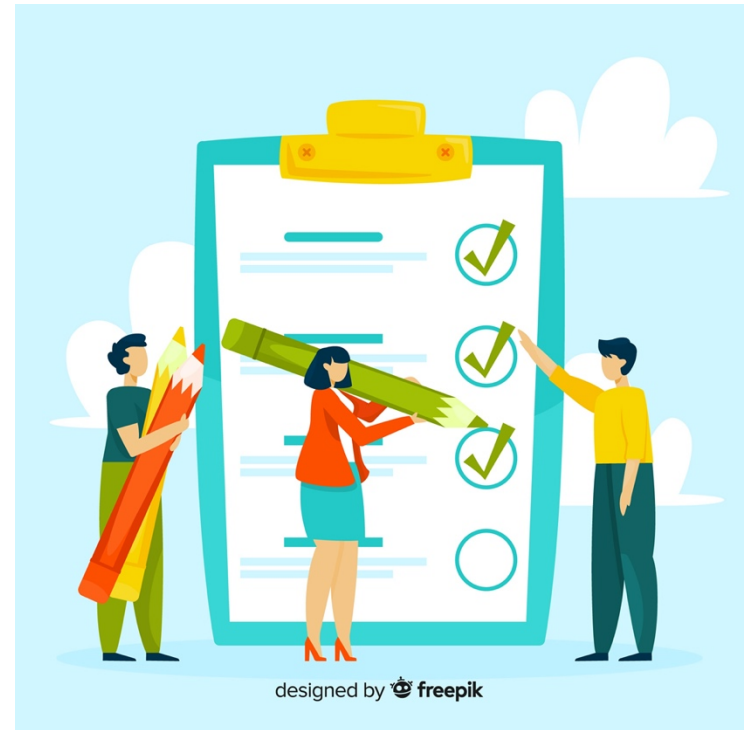
評価を研究推進/支援職が担うことは可能かもしれない

## 参考資料



# プロジェクト設計のための評価：タイミング

- 初期  
参加者の属性や知識レベルを知る
- 実施中  
参加の障壁を明らかにしたり、計画通りに進んでいるかについて知る
- 事後  
成果、出来なかった点、応用可能性などを知る



Designed by background / Freepik

# プロジェクト設計のための評価：流れ

1. 評価手法を選択

2. プレテスト

3. 手法を洗練

4. データマネジメントプランを作成

5. 参加者を集める

6. 評価を実施

<実施時の注意点>

- 自発的な参加
- インフォームドコンセント
- リスクの説明
- 機密保持
- 個人情報保護

# まとめ：実践に向けて(将来像)

1. Why: 大学として取り組むべき？
  - 多様な目的が組織の方針に合致すれば業務として推進は妥当
2. Who: 誰が担うのか？
  - 組織間連携にURA(のような人材)が貢献できそう
  - インフラ/ウェブサイトの構築/維持は図書館等が重要な役割
3. How: どうやって実践する？
  - 参加者の属性や動機を見極める
  - データの質を担保する
  - プロジェクト設計のために評価を行う(研究推進/支援人材?)