

50+ Reasons Not to Release Open Data

For the background on this document, [see this blog post](#).

Update: We have rolled out these rebuttals as blog posts, which you can view here:
<https://sunlightfoundation.com/taxonomy/term/why-open-data/>

Thanks so much to all of you who contributed ideas for this project throughout its evolution! We couldn't have done it without the help of so many different perspectives, and we feel the end result is richer for it. Join the continuing conversation with [#whyopendata](#).

Index of categories (click to jump to a topic):

[Apathy/Confusion: The Problem of Newness](#)

- a. [No one cares](#) [1 - 3]
- b. [Why do this?](#) [4 - 11]

[The Challenges of Integrating New Systems \(Or, Rage Against the Bureaucracy\)](#)

- c. [It's hard](#) [12 - 18]
- d. [It's expensive](#) [19 - 22]

[People Power: Staff\(ing\) Concerns](#)

- e. [It's not my job](#) [23 - 25]
- f. [We don't have the people to do that](#) [26 - 27]

[Heavy Hitters: Legality, Security, Accuracy, and Privacy](#)

- g. [We can't legally do that](#) [28 - 31]
- h. [It will be misunderstood / misused / it's not accurate](#) [32 - 37]
- i. [Privacy](#) [38]

[Open Data as a Work \(Already?\) in Progress \(and Other Excuses\)](#)

- j. [It's already public](#) [39 - 41]
- k. [Say what? / Misc.](#) [42 - ##]

Apathy/Confusion: The Problem of Newness

No one cares

- 1. Why would we do this?**
 - a. You mean people are actually interested in this?
- 2. No one cares**
 - a. There is a lack of public interest
 - b. There are few public requests for data to be open/for data in general
 - c. Government is boring
 - d. No one would be interested in it
- 3. We don't have community support**
 - a. We don't have civic hackers

Why do this?

- 4. There's no clear use for {insert specific type of data here}**
- 5. I don't know what you could possibly need that for**
 - a. How would someone even use that kind of data?
- 6. What kind of data do you want?**
- 7. We need more proven results before we can open data**
- 8. The old system works**
- 9. We don't want to have you (academic researcher) doing research for research sake. We want your research to benefit us directly, otherwise we don't want to give you our data.**
- 10. It would just confuse people.**
- 11. We should really be selling this.**

The Challenges of Integrating New Systems (Or, Rage Against the Bureaucracy)

It's hard

- 12. It's hard**

- a. I don't know how to start
- b. I don't know how to engage my government/my boss on this issue
- c. I don't know who to talk to about it
- d. I don't know how to organize on this issue

13. We don't know how

- a. I honestly haven't ever heard of an API or (insert other data term here)
- b. [You'll need to write custom software / our system can't do that](#) (Getting access to Police incident logs from a computer aided dispatch system)

14. We don't know how to address our backlog

- a. We don't know how to address our archives

15. We won't be able to keep up! / We don't have the time

- a. I don't have time/no one will help me
- b. There will be unending data release
- c. ... unending cost
- d. ... exhaustion
- e. Our staff can't handle new software/training/processes
- f. We don't have the technology
- g. We don't have a website
- h. We don't have the servers for this
- i. We would have to take our entire production system offline for a couple of days to program the system to get the data you want.

16. There's too much data

17. It's unfamiliar

- a. I honestly don't know what open data is
- b. We don't have data
- c. I don't know what data is available that can be opened

18. It will take too much of our time.

It's expensive

19. It's expensive / We don't have the budget

- a. It requires expensive software
- b. It would be costly to provide copies of data in open formats
- c. It would require a lot of staff time
- d. It would require new processes and staff training
- e. Keeping everything updated would be costly

20. We're a small town/city/county/state/country

21. We have a tight budget right now

22. My Professional Incentives are to NOT Share

- a. When I share my work or products publically, I either have to take an additional time-consuming step or potentially end up creating new work that doesn't actually lead to any personal rewards. Will I get paid more if I open government data? Will I get promoted faster? Will I be able to leave work earlier or receive good recognition from my supervisors? The answer to all of these questions is usually "No". If open data is not part of a compensated job function, there shouldn't be any reasonable expectation that it will get done.

People Power: Staff(ing) Concerns

It's not my job

23. It's not my job

- a. That's someone else's job

24. My boss hasn't told me to and you're not my boss

25. My bosses/the politicians don't like it

We don't have the people to do that

26. We don't have/can't afford a CTO/CDO/CIO/other staffer to do that

27. We don't have anyone who knows how to do that

Heavy Hitters: Legality, Security, Accuracy, and Privacy

We can't legally do that

28. We don't have the legal authority to make open data/records

29. It's illegal

30. It's proprietary

- a. We don't have the rights to do that

31. It's not classified, but we don't think it would be good PR to open this

It will be misunderstood/misused

32. If we put the data out there in bulk, people will alter it

- a. Hackers will do stuff with it
- b. People can steal it
- c. The data will be [mishandled](#) or misinterpreted
- d. We can't make sure the government will still own the information if it's open
- e. People might misuse it

33. If we share our data/code, we'll be hacked

34. It might be presented in ways that result in people misunderstanding it

- a. The media will misreport

35. People don't understand my data.

- a. It's complex/magical/for experts

36. [The data source is a mess](#)

- a. It will be hard to parse and put online
- b. Releasing it like this will make us look bad

37. The data might have errors or mistakes and could misinform the public

- a. We're not even sure how accurate the data is, so we don't see the point in sharing it
- b. There are problems with the data

Privacy

38. PRIVACY!

- a. It's classified
- b. We're worried about the mosaic effect.
- c. We don't want any perception at all that the government is collecting information about the American people

Open Data as a Work (Already?) in Progress (and Other Excuses)
(Or, [I don't think that word means what you think it means](#))

It's already public

39. It's already public

- a. Anyone can access records during working hours at our government building
- b. It's available on LexisNexis or through another provider / You can buy it from X company for a fee

40. That's what FOIA is for

41. We make enough available now

- a. We are an open government

Say what? / Misc.

42. But we released APIs!

- a. We're more interested in APIs

43. Oh, you mean make a data portal? We have one of those.

44. How can we verify the authenticity of the data?

45. We spent a lot of [time, staff-power, resources, etc.] collecting this data

46. [Our agency is making money from selling the data](#)

47. Someone will repackage and sell the information

48. That would really piss off industry

49. We want to control who uses it and what gets done with it (even if it means nothing is done with it)

50. Our competition will be able to do more with it than we do with it

51. The federal government/other overlapping agency already has our data

- a. Another layer of government is already collecting this data -- get it from them, not us

52. That would really piss off Congress

53. It's not a good time

- 54. Moving too quickly may make us go too far and then be unable to return
- 55. Data being used by privileged technologists and educated communities creates a further divide to access to information
- 56. This just isn't a priority now
- 57. It's pending/not final/in progress
- 58. We lost it (the data) / we don't have it

Additional reasons to potentially address:

From the document "**Concerns about opening up data, and responses which have proved effective**," available here:

https://docs.google.com/document/d/1nDtHpnIDTY_G32EMJniXaOGBufjHCCK4VC9WGO7jK4/mobilebasic?pli=1

We'll get spam

- We don't have to include email addresses in the data
- You already have those email addresses on the website
- We have a spam filter!
- We can make including your email opt in or opt out.

Our lawyers want to make a custom license^[d]

- This usually comes from not explaining to the lawyers what you're trying to do
- Often there's chinese whispers via levels of management between techies and lawyers, see if you can talk with the lawyer directly, and may not be familiar with open licenses.
- Precedents can help. If you can show several peer organisations that have published under this open license you wish to use, people will be more confident.

If we publish this data, people might sue us

- If the concern is about information about people, make a policy and stick to it. eg. we'll correct/remove things in 7 days. (This is already embodied in one of the [Data Protection Principles](#) in the UK)
- If the concern is about errors which impact business, case law may help, but there's not much yet. (let us know when there is some!)
- Ensure you can state what data you would have been providing on any given

day, eg. someone claims they couldn't find the nearest hospital in your data and someone died. Very helpful to be able to prove/show what data you were providing at they day/time.

If the data is already available on your website or via an app, then are you really increasing risk by opening the data?

We want people to come direct to us so we know why they want the data^[g]

This is similar to "People might misinterpret the data" if the intention is to help people satisfy their information need.

Well, OK, can the *metadata* or some minimal subset at least be made open? If there's the option to open up the data, chances are that the information would already be available under Freedom of Information rules, wherein requestors aren't required to justify or explain themselves. This still allows the owner to find out on a "best efforts" basis, but it's important to remember that they won't be able to find out on every occasion.

An example case is a comms department "find an expert" service for arranging media interviews with your staff. In this case, rather than provide direct contact information you can offer to make the contact information via a logged web-form or specific email address. This will be subverted if the people in the dataset is linked with your open phonebook dataset.

This can be a tricky one as once you open your data you can't track every instance of use. This can be very threatening to people who produce this data and need to be able to provide documented proof that they/their team are doing something of value, especially when staffing levels are being reduced. One option is to find someone in power to reassure them that their work is valued and that their management understand & accept that when the data is opened they are trading control for utility.