Agenda and Meeting Record

Who	W3C Credible Web CG
When	28 January 2020 1pm ET
Video	https://zoom.us/j/706868147 Please try to be on camera
Telephone	+1 669 900 6833 or any dialin, code 706 868 147
Chair	Sandro Hawke
Scribe	JC Goldenstein
Present	Listed in "Introductions" below
At Risk	(add your name here, if applicable)
Regrets	(add your name here, if applicable)

Summary

The current series of meetings is focused on creating a short document specifying a small set of credibility signals where we have consensus they're "promising". This was our second meeting in this series. First half of the meeting was about generalities and process, then we got into the details of today's proposed signal. We ended up with consensus to approve Date Website First Archived, and at the end looked at who could volunteer to write up some signals for next week.

1. Introductions

Please put your name (affiliation) here. Extra information and links welcome.

1. Sandro Hawke (W3C/MIT, HawkeWorks)
2. Subbu Vincent (Markkula Center/SCU)
3. JC Goldenstein (CREOpoint) who took notes of the call below
4. Jesse Kranzler (Trust Metrics)

5. Anand Upender (TED)
6. Tzviya Siegman (Wiley)
7. Annette Greiner (NERSC, Berkeley Lab)
8. Connie Moon Sehat (Credibility Coalition, NewsQ)
9. Tanu Mitra (Virginia Tech)
10. Scott Yates (CCC and under contract to JTI)

2. Announcements

- Misinfocon
  https://misinfocon.com/meet-the-speakers-at-misinfocon-nasem-3ca77ca188ab
- CREOpoint patent application "Containing the spread of disinformation" is now public (search for other relevant ones by clicking here. Contact: jc@creopoint.com. Sandro notes W3C practice that people shouldn't say anything about a patent that could be used to suggest a listener knew about the patent, and then might be accused of knowingly infringing, if they were infringing.

3. Group Meeting Schedule

New year, maybe new meeting schedule. Poll at
https://doodle.com/poll/c3v8nthtakru6a8m

Result is Wednesdays 2pm ET starting next week (8 days from now)

4. Decision Process

Plan: publish a Community Group report which highlights a handful of signals, before the end of February, to revise regularly thereafter if it seems useful.

- W3C Consensus Process: see
  https://www.w3.org/2019/Process-20190301/#Consensus
- Aim for general agreement during meetings
- Allow review by absent group members for 72 hours, after which decisions are complete, unless an objection is raised in group email.
- Decisions may be re-opened in light of new information
- We'll have separate decisions to publish docs, to review how results of decision are presented

Sandro about decision-making process above:

1. Objective: which signal and what to say. See green in 5 below for the proposed decision we will discuss today.
2. Hopefully the justification will be self-explanatory in the link so readers don't have to track down our archives.
3. 1st time making decision during a meeting in over 2 years.
4. We'r don't generally vote, but instead we minimize objections by working them out
5. People who can't make the weekly meeting will have 72 hours after the call to object to a decisionl
6. New info ---> new decision. We don't reopen things that were known
7. Signal approval by our group means it goes to draft for iterations = it's not final

JC asked what happens standards and communications wise when our group agrees on a signal

Sandro replied:

- We would need to be a W3C Working Group to be empowered to create standards. This is more like pre-standard work. To create a standards working group requires approval from about 20 w3c member companies. If our work goes well, it might lead to that.
- W3C has some channels for getting us publicity, especially since our work is of such clear importance to society. We might want to wait for a second release before trying to get the full attention, with Twitter and w3.org news

Tzviya: as a W3C Advisory Board member, I'm looking at improving this type of incubation process. See https://www.w3.org/2001/tag/doc/distributed-content/ explains this issue well

5. Proposed Signal: Date Website First Archived

PROPOSED: We endorse as "promising" the signal currently defined and described at Signal: Date Website First Archived (version 3)

Group discussion before endorsement (see vote at the end)

- Sandro: The word "promising" may evolve
- Annette: we need to consider context eg an image just posted on FB. Sandro recommend to start with easier stuff like websites. Clarify this is applicable for syndicated content.
- Subbu wanted clarification if it's about when is signal important to be used or only explanation of use case?
- Tzviya: Useful metadata to connected to original URL like timestamps? it's not obvious If the original URL is changed. Sandro responds that a modified image would be harder. NewsQ signals about domains, Trust Project also about domains and web pages, JTI about news organizations.
- Connie: NewQ changing data model new concept of entity to represent relationships between URLs and publishers so use concrete cases in decision making. Sandro wonders about complexity given users like/need to keep things simple
- Jesse has examples Buzzfeed from clickbait to BF News ½ legitimate.
- Scott mentioned IPTC (click here). Sandro could not schedule our meetings to work for MD https://iptc.org/news/tag/iptc-managing-director Brendan Quinn, who lives in Eastern Europe
- Sandro: Let's not consider change of ownership, yet, since it could be too complicated. Let's start with simpler signals.
- Connie: More easily gleaned from quick observation of the site. Describing more than content leads to other entities. Is is the same quotation? Publisher? Owner? Is there a site at this URL? Wants cleaner signal in the long run. Also when publication or owner change there's rebranding so seeing same potential site now in a different URL not uncommon.
- Jesse: this domain exists at this point, this signal on his own is not the end-all but when combined with the publisher and context it would be. Warned there are 100s of signals so important to make it clean and as clear as possible.
- Subbu: Keep one signal as a signal since when taken in a group at a time (analogy with atoms and molecules!). This is the date this domain was first archived add which publisher the domain
- Connie: how someone signal may cluster together. Site 1st archived URL. Publication started before URL X years before and started accrued many outlets. Query is to know what the data is and what relationships and history exist.

About the value of the signal to an algorithm

- Sandro: Back to simplicity. Flagging newly created hoax site would be a pretty simple and useful signal.
- Sandro: An algorithm could add 1 point for every year of history of the site

- Jesse value history of a publication but note it could be positive point or neutral. However watch giving negative points to a new site that could be very good. Sandro replied by switching from chemistry to math where <2 yrs old not good etc

Importance of context and content rather than only the source

- Subbu agreed with Annette's initial point about complexity given various contexts. For example students interested in digital literacy may rate more positively a new media born yesterday from a new NGO
- Annette: yes so much is content dependent. Signal is promising but "I wouldn't want it to be used alone."
- JC agreed
- Sandro also agreed it's important to know also how the signal will be used.
- JC expressed concern about all the time it takes for just one signal related to URL that may end up being useless anyway. Even the BBC makes mistakes (e.g., recently showed footage of Lebron James instead of Kobe Bryant).
- JC proposed to consider focusing on damaging breaking news content: rating URLs is necessary but far from sufficient if we don't take into account the context such as a new doctored video content of a political or business leader. What democracy, brands and hedge funds could use is a simple and quick signal that would help before the damaging content spreads and the harm is done.
- On the issue of this signal should count towards your credibility but not against it if missing, there was agreement that only makes sense in some deployment models, like if issuing a warning to the user. If you're calculating a credibility score, there's probably no difference between a downgrade and the absence of an upgrade.

Things take time so focus on the right objectives is key

- Sandro: We could spend a month talking just about one signal. We've got a lot more signals we want to look at soon, and ten more minutes this meeting. Do we have enough agreement to go forward with this one?

Vote

- Sandro: The group reached a consensus after participants voted yes or no in chat
- Subbu took a group pix since this is a first!!!

RESOLVED: We endorse as "promising" the signal currently defined and described at [Signal: Date Website First Archived](#) (version 3)

6. Next Signals to Endorse

Discussion about some ideas below:

- Pick one to write up for next week? Quick straw poll?
- Participants to add per below or agree which signal they volunteer to review
- JC will also propose a signal about the trustworthiness of impactful breaking news (content - not sources) to contain the spread disinformation. Sandro agree to review :-)
- Subbu will work corrections policy with Connie helping per list below

Possible signals to discuss

1. Valid markup
2. https://credweb.org/signals-nq87/#edward-r-murrow-prizes Connie, award
3. https://credweb.org/signals-nq87/#corrections-policy   Subbu to write, Connie to review
4. https://credweb.org/signals-nq87/#about-us
5. https://credweb.org/signals-nq87/#nobias-slant
6. https://credweb.org/signals-nq87/#contact-info-mailing-address
7. https://credweb.org/signals-nq87/#article-attribution
8. Includes citations
9. Popularity/Traffic
10. Puts clear dates on posts
11. Title Representativeness, CONNIE to write
12. Lacks necessary citations
13. Engages in harmful amplification
14. Spreads disinformation
15. Makes claims that have been deemed false by IFCN

7. Adjourn

8. Post-Meeting Discussion Notes

```
Zoom chat log:
13:09:12      From connie :
https://misinfocon.com/meet-the-speakers-at-misinfocon-nasem-3ca77ca188ab?source=colle
ction_home---4------0----------------------
13:27:26      From Tzviya Siegman :
https://www.w3.org/2001/tag/doc/distributed-content/ explains this issue well
13:35:12      From Tanu Mitra To Sandro Hawke(Privately) : So sorry to leave the call
now (conflict with teaching time). Hopefully from next week I can be present for the
entire duration.
13:51:24      From Scott Yates : I vote YES
13:52:29      From Annette Greiner : so far, I think it's promising, but I wouldn't
want it to be used alone.
13:52:33      From Jean-Claude Goldenstein : JC: concerned about time it takes we could
spend a month talking about one signal. Would prefer to agree signals about breaking
disinformation/content
13:52:55      From connie : 1+ annette
13:53:10      From Scott Yates : I voted "GO FASTER" which means yes+
13:53:15      From Tzviya Siegman : +1
```