Global Health Data Science Hub working group meeting notes – 6 June 2023

1) Review Hub priorities (based on discussions at the May meeting)

Notes
3 main priority areas for the rest of the year are: **Issues around data collection, management, and reuse:** looking at data management plans, but also some of the tools and terminology used around data. Would also touch on data reuse.   **Data curation and analysis:** a focus on skills, tools, and techniques (and the reasons behind those techniques)   **Governance, trust and ethics:** lots of issues around privacy and governance. Trust is in terms of working in communities and working with community data. Data access is another issue that falls under this – who owns the data and who benefits from research?   - We can address priorities in different ways, like articles, blogs, webinars. Tableau is another important tool to integrate into the skills learning package – Tableau Public is the free version of this tool.   - Should add federated data analysis under the governance theme but also related to cloud computing.

2) Review content and curation log

Notes
- Ideas for the "spotlight on" and webinars have also been added   - The content and curation log populates the Resources page on the Hub – some new resources have been added since the last meeting. In particular, some resources around R training have been added (Applied Epi). Some information around AWS has also been added; other cloud provider resources will be added too.   - The things that we're missing at the moment are good example projects or use cases – that's something we could target using the priority areas from the previous agenda item.   - We are encouraging members to promote their work through this platform. The tools offer great functionality for highlighting work. We can mint DOIs as well. A good tool is Discourse (a forum of sorts) – can link this to the Hub.

3) Review Hub metrics

Notes
- Consistently getting >200 new members every month.   - We are getting returning visitors too, but not as many as new members.   - UK, Brazil, India are the top 3 countries of visitors (March), UK, US and Bangladesh are the top 3 in April   - English speaking countries in SSA are the ones that engaging in that region, would be good to broaden out

4) Discuss first 'Spotlight on' (suggested focus on R and Data Science Clubs and Clinics

Notes
'Spotlight On' might be a good way to drive more traffic to the Hub – we will publish info on a particular specific topic – can tie it to training, webinars, resources

- The first one could be on R as a tool for data science – Global Surg and Applied Epi have free R training resources that could be tied into this. Data clinics & clubs can tie into this too.
- R has a lots of applications, so an overview of that would be really helpful as one of the first webinars.
- Participants will likely be those who are new to research or intermediate researchers to build expertise/skill in data
- Work with key stakeholders like Africa CDC to help us reach our target audience

5) Share plans for an AI webinar and seek input/suggestions

Notes
- TGHN AI Hub has launched and we are discussing joint webinar series on AI and machine learning as it links well with data science work.   - Focusing on a number of themes (potentially intro to AI & data science, data collection, responsible AI, etc)   - We will add a brief example from a Pathfinder Project to the Hub

Next meeting: Tuesday 4 July 12.00-13.00 BST

If you are interested in joining the Global Health Data Science Hub working group please contact Sally Boylan on Sally.Boylan@hdruk.ac.uk