# Integration of Administrative Data for Science and Public Policy: An Experience in Brazilian Public Health

Authors:
Bethania de Araújo Almeida, Paula Xavier and Mauricio Barreto.

Contact: Bethania Almeida. Center of Data and Knowledge Integration for Health (CIDACS), Fundação Oswaldo Cruz. Email: bethania.almeida@fiocruz.br

## Background and Content

The deployment of administrative databases for research purposes holds great potential. Administrative data becomes even more useful when linked to other datasets, making it possible to elucidate the effects of combined factors that could potentially impact the health of populations. On the other hand, the access, use and reuse of these administrative datasets, principally those containing personally identifiable information, are all topics being widely discussed nowadays in an effort to maintain confidentiality, protect individuals' privacy, respect terms of consent and manage security or other risks. For these reasons, we believe that our data management experience at the Fiocruz Center of Data and Knowledge Integration for Health (CIDACS) in Brazil can make a significant contribution by providing high-quality linked data for research purposes and generating evidence for public health policy.

The usage of administrative data containing personal information for linkage purposes requires a legal basis for access, in addition to considering the provision of sound security arrangements, exclusive usage for previously specified purposes, appropriate credentials from the requesting institution and justification regarding the ethical basis of the proposed study. Furthermore, it is very important that researchers understand their responsibility to use data for bona fide purposes only, as well as be aware that legal action will be taken if/when data are used inappropriately or without due care (Harron et al., 2017).

Since linked administrative datasets constitute a valuable tool for population-based research applications and have implications for public policy-making, this poster presents the experience of a Brazilian data center that uses large volumes of administrative data for linkage purposes in the area of public health, taking into account privacy, ethical and information security concerns, as well as technical and scientific considerations, which may eventually become aligned with the principles of open science.

## Method
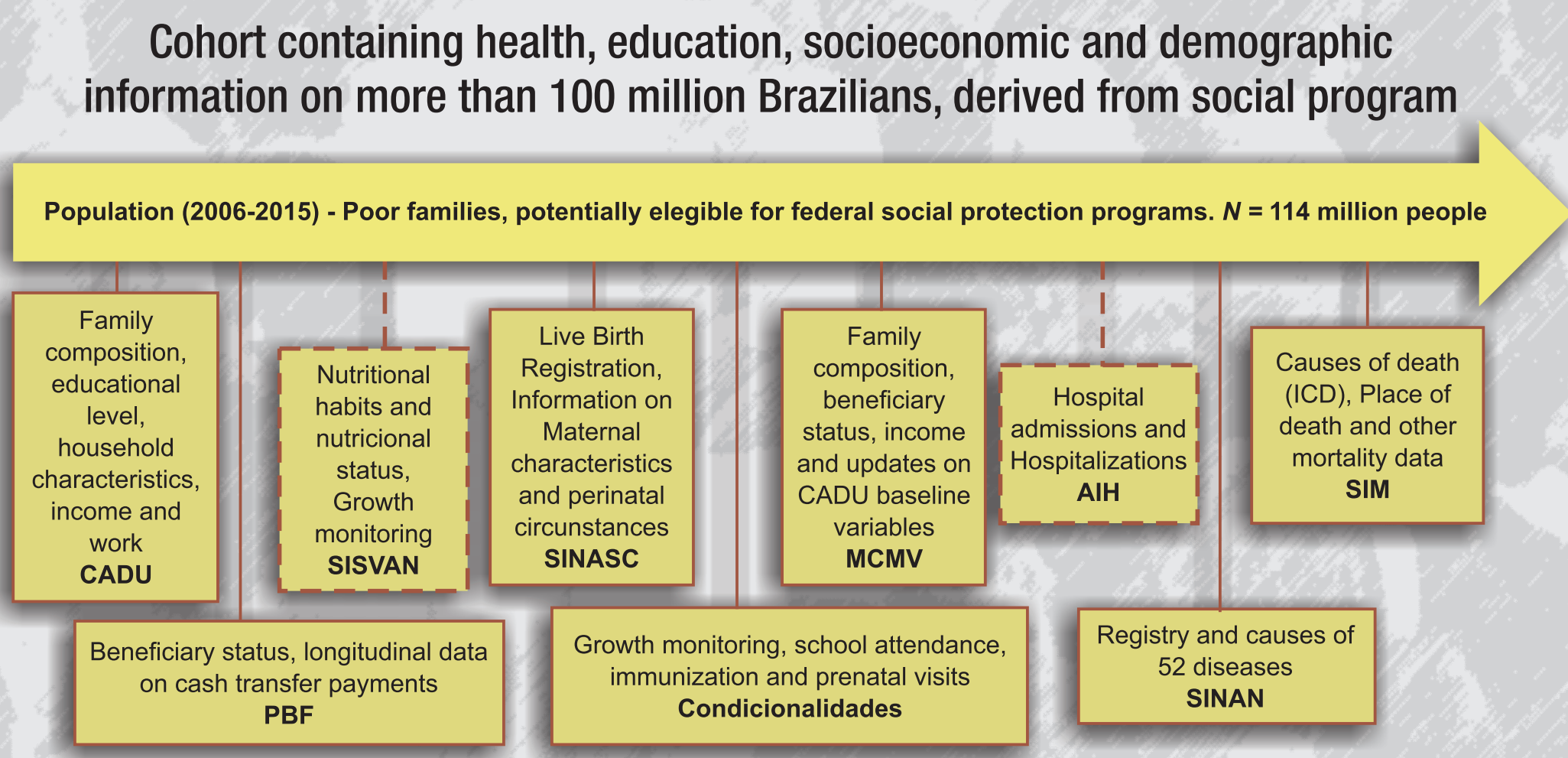A participant observation case.

## Results
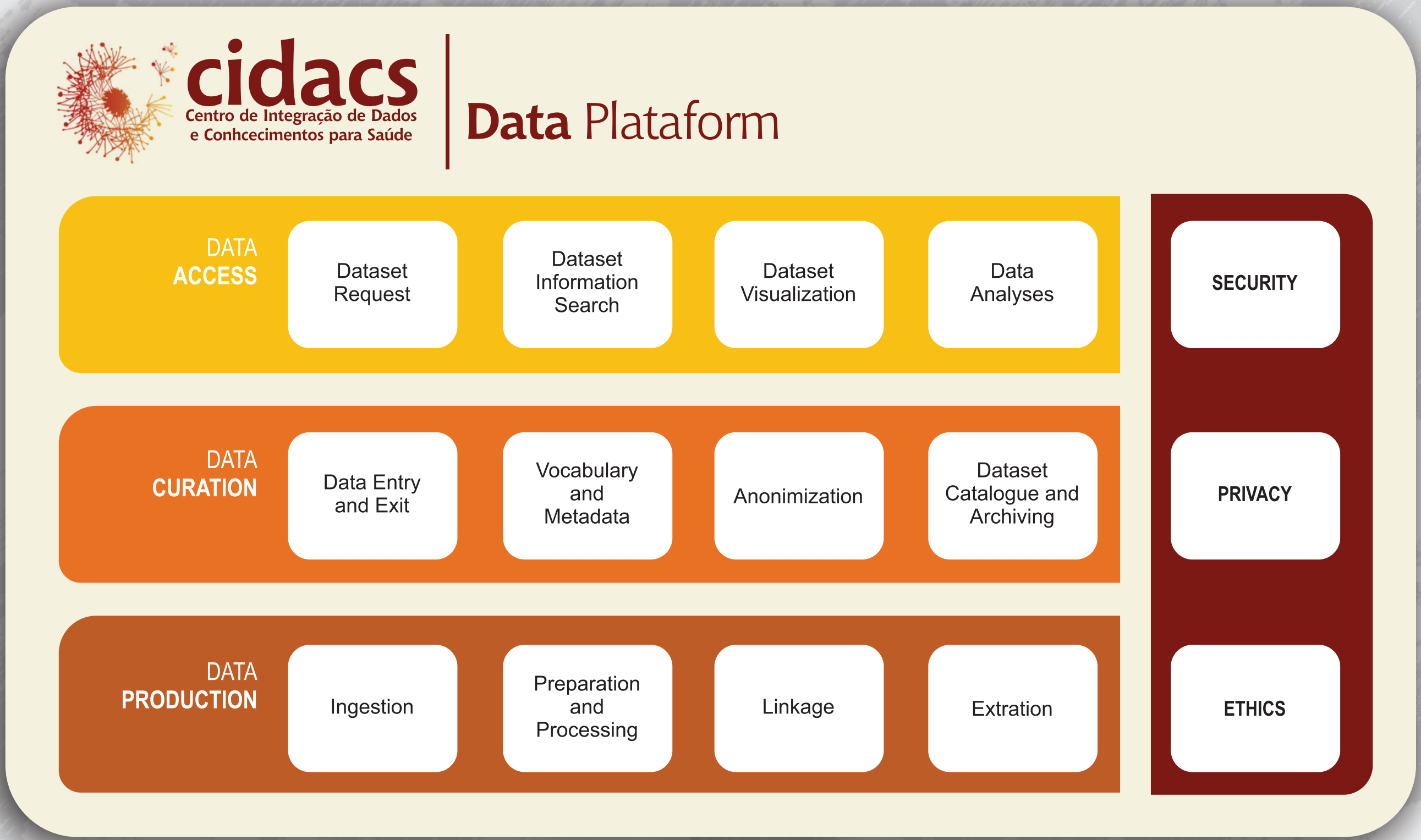


Source: CIDACS, 2018.

By addressing the issues surrounding the treatment of administrative data containing personal information, as well as the use and reuse of linked anonymized datasets to support public health research and evidence-based decision-making to inform policy, CIDACS also intends to explore other efforts to adhere to open data practices in the mold of FAIR Principles, including distinguishing between data and metadata to support a wide range of special circumstances.



Source: Australian National Data Service - FAIR Data Training, 2018.

Cohort containing health, education, socioeconomic and demographic information on more than 100 million Brazilians, derived from social program



At this time, approximately 20 studies are being conducted using this cohort, including the effects of social determinants, as well as of conditional cash transfers or housing programs for low-income populationson health outcomes.

## Conclusion

We believe that big-data platforms, such as that of Cidacs, hold the potential to create data ecosystems capable of generating knowledge and evidence to support scientific investigation and decisions made by policymakers, administrators and others throughout society in an effort to fact the problems that impact the health of populations.

## Acknowledgements

## References

Australian National Data Service - FAIR Data Training. (2018). Retrieved from https://www.ands.org.au/working-with-data/fairdata/training?utm_source=home-page-carousel&utm_medium=Web-banner&utm_campaign=fair-data-training

Connelly, R., Playford, C.J., Gayle, V. & Dibben, C. (2016). The role of administrative data in the big data revolution in social science research. Social Science Research, (59),1–12. DOI: 10.1016/j.ssresearch.2016.04.015

European Commission. (2016). H2020 Programme: guidelines on FAIR Data Management in Horizon 2020, v.3. Retrieved from http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf.

Go Fair. (2018). Global Open, Findable, Accessible, Interoperable, Reusable. Retrieved from https://www.go-fair.org/.

Harron, K., Dibben, D., Boyd, J., Hjern, A., Azimaee, M., Barreto, M.& Goldstein, H. (2017). Challenges in administrative data linkage for research. Big Data & Society. July–December, 1–12. DOI: 10.1177/2053951717745678

Leonelli, S. The Epistemology of Data Use. (2017). Lecture in The Center of Philosophy of Science - University of Pittsburgh. Retrieved from https://www.youtube.com/watch?v=qc1aQep4DE8.

Santos, P., Almeida, B. & Henning, P. (2018). Livro Verde - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz. Retrieved from https://www.arca.fiocruz.br/handle/icict/24117.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W.,da Silva Santos, L.B., Bourne. P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K, Zhao, J.& Mons B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data3:160018 DOI: 10.1038/sdata.2016.18.

Wright, M. (2014). Using Administrative Data In Research: Learning from the Public Dialogues. Administrative Data Research Network. Retrieved from http://www.esrc.ac.uk/files/research/administrative-data-taskforce-adt/using-administrative-data-in-research-learning-from-the-public-dialogues-melanie-wright/