

RESUMO

A Importância da Replicabilidade na Ciência Política: O Caso do SIGOBR

O principal objetivo desse trabalho é difundir o padrão de replicabilidade nas Ciências Sociais no Brasil em geral e na Ciência Política em particular. Para tanto, reproduzimos alguns argumentos da literatura em favor da replicabilidade como procedimento padrão da atividade de pesquisa. Defendemos que a replicabilidade tem três principais dimensões: (1) substantiva, na medida em que contribui para o aprimoramento e acúmulo do conhecimento científico; (2) pedagógica, já que facilita a compreensão de noções básicas de análise de dados e (3) transparência, na medida em que protege a comunidade acadêmica não só contra erros honestos, como também de fraudes intencionais. Além disso, apresentamos as principais características do projeto Sistema de Informações de Governo (SIGOBR), conduzido pelo Departamento de Ciência Política (DCP) da Universidade Federal de Pernambuco (UFPE), que tem o padrão de replicabilidade como um dos seus componentes essenciais.

Palavras-chave: *replicabilidade; ciência política; SIGOBR.*

ABSTRACT

The Importance of Replication in Political Science: the case of SIGOBR

The principal aim of this paper is to diffuse the replication standard in Brazilian Social Sciences in general and in Political Science in particular. To do so, we reproduce some arguments that favor replication as a default standard in research activity. We argue that replication has three main dimensions: (1) substantive, since it contributes to both increase quality and cumulation of scientific knowledge; (2) pedagogical, since it facilitates the understanding of data analysis basic concepts and (3) transparency, when it protects scientific community not only from honest mistakes but also from intentional frauds. In addition, we present the main characteristics of the Government Information System project (SIGOBR), conducted by the Political Science Department (DCP) of the Federal University of Pernambuco (UFPE), which has replication standard as one of its key features.

Keywords: *replication; political science; SIGOBR.*

A Importância da Replicabilidade na Ciência Política: O Caso do SIGOBR¹²

Enivaldo Carvalho da Rocha³

Ranulfo Paranhos⁴

Dalson Britto Figueiredo Filho⁵

Erinaldo Ferreira do Carmo⁶

Few events in academic life are more frustrating than investing enormous amounts of time, effort, and pride in an article or book, only to have it ignored by the profession, not followed up by other researchers, not used to build upon for succeeding research, or not explored in other contexts.

Gary King

INTRODUÇÃO⁷

Em 1995, o periódico *Political Science and Politics* (PS) lançou um dossiê sobre a replicabilidade nas Ciências Sociais. Entre as diferentes contribuições, Gary King alertou a comunidade acadêmica a respeito da importância da replicabilidade dos resultados de pesquisa como componente central do conhecimento científico. Além dos benefícios substantivos associados ao desenvolvimento e à acumulação do conhecimento, a replicabilidade é uma ferramenta pedagógica extremamente eficiente (BOYER, 2003). Ela permite que estudantes de graduação e pós-graduação possam seguir os mesmos passos percorridos por pesquisadores mais experientes. Ainda, a replicabilidade protege a comunidade científica não só de

1 Esse produto contou com suporte financeiro do CNPq. Eventuais imprecisões nem delegamos nem partilhamos: são todas nossas.

2 Esse trabalho se beneficiou do treinamento recebido durante o Berkeley Initiative for Transparency in the Social Sciences (BITSS) realizado na Universidade de Berkeley, Califórnia, Estados Unidos entre 2 e 6 de junho de 2014.

3 Professor do Departamento de Ciência Política da Universidade Federal de Pernambuco (DCP/UFPE), Pós-doutorando do Departamento de Ciência Política da Universidade Federal de Minas Gerais (DCP/UFMG) e Doutor em Estatística pela Universidade Federal do Rio de Janeiro (UFRJ).

4 Professor do Instituto de Ciências Sociais da Universidade Federal de Alagoas (ICS/UFAL), Mestre e Doutor em Ciência Política pelo Departamento de Ciência Política da Universidade Federal de Pernambuco (DCP/UFPE)

5 Pós-doutor, Doutor e Mestre em Ciência Política pelo Departamento de Ciência Política da Universidade Federal de Pernambuco (DCP/UFPE)

6 Professor do Colégio de Aplicação do Centro de Educação da Universidade Federal de Pernambuco (UFPE). Doutor e Mestre em Ciência Política e Graduado em Ciências Sociais (Bacharelado e Licenciatura Plena) pela mesma instituição.

7 Para o leitor interessado em saber mais sobre o padrão de replicabilidade sugerimos acompanhar o dossiê publicado pela *Political Science and Politics* (PS) em 1995. Ver também Sieber (1991) e Boyer (2003).

erros honestos, mas também de fraudes deliberadas. Ao se adotar o padrão de compartilhamento de dados e replicabilidade, qualquer tentativa de manipulação indevida pode ser facilmente detectada por outro pesquisador.

Apesar de ser amplamente utilizada na Ciência Política norte-americana, a replicabilidade é um conceito ainda pouco difundido no Brasil. Por exemplo, salvo melhor catalogação, analisamos o conteúdo de todos os artigos publicados em quatro prestigiosos periódicos nacionais – Revista DADOS, Revista Brasileira de Ciências Sociais (RBCS), Revista de Sociologia e Política (RSP) e Opinião Pública – e não encontramos uma única referência a esse termo. O objetivo central desse trabalho é difundir o padrão de replicabilidade nas Ciências Sociais no Brasil em geral e na Ciência Política em particular. Para tanto, reproduzimos os principais argumentos da literatura em favor da replicabilidade como padrão desejável na atividade de pesquisa. Além disso, apresentamos as principais características do projeto Sistema de Informações de Governo (SIGOBr), conduzido pelo Departamento de Ciência Política (DCP) da Universidade Federal de Pernambuco (UFPE) que tem a replicabilidade como um dos seus eixos estruturadores.

O restante do trabalho está organizado em três partes. A próxima seção apresenta o conceito e as vantagens do padrão de replicabilidade. Depois disso, descrevemos as principais características do projeto Sistema de Informações Governamentais (SIGOBR). A última seção sumariza nossas principais recomendações.

2. REPLICAR É PRECISO⁸

O padrão de replicabilidade tem três principais dimensões. Primeiramente, ele contribui para o aprimoramento da qualidade dos resultados de pesquisa e acúmulo do conhecimento científico (substantiva). Em segundo lugar, a replicabilidade facilita que alunos de graduação e pós-graduação compreendam noções básicas de análise de dados (pedagógica). Por fim, o padrão de replicabilidade protege a comunidade científica não só contra erros honestos, como também de fraudes deliberadas⁹ (transparência). As próximas seções desenvolvem esses argumentos.

2.1 O que é replicabilidade?¹⁰

8 Parte dessa seção foi extraída do artigo *Levando Gary King a sério: desenhos de pesquisa em Ciência Política*, veiculado pela *Revista Eletrônica de Ciência Política*: <<http://ojs.c3sl.ufpr.br/ojs2/index.php/politica/article/view/29614/21569>>

9 Para Fowler (1995), “among the likely benefits of replication, several stand out: the creation of cumulative body of work founded on similar methods and data; the improvement of graduate training through enhanced student access to major studies conducted by prominent political scientists; and the enhancement of methodological standards in the discipline” (FOWLER, 1995: 478).

10 Evidentemente, existem argumentos contra a replicabilidade. Por exemplo, Gibson (1995) acredita que a adoção estrita do padrão de replicabilidade reduz os incentivos à produção de novas bases de dados. Isso porque os pesquisadores tenderiam mais a utilizar bases já existentes do que a gastar tempo e recursos na elaboração de bases de dados originais. Herrnson (1995) afirma que “this data relinquishment or verification policy is mistakenly referred to as a replication policy, and would harm researchers, journals, the discipline, and the acquisition of knowledge about politics” (HERRNISON, 1995: 452). Maisel (1995) também discorda das sugestões de King (1995). Fowler (1995) adverte a respeito das diferentes consequências negativas não-intencionais associadas à adoção do padrão de replicabilidade.

Para King (2006), o padrão de replicabilidade requer a disponibilização de informação suficiente para compreender, avaliar e replicar os resultados de um determinado trabalho sem informação adicional do autor do estudo¹¹. Além disso, King (1995) afirma que “the replication standard does not actually require to anyone replicate the results of an article or book. It only requires sufficient information to be provided (...) so that the results could in principle be replicated” (KING, 1995: 444)¹². O componente básico do padrão de replicabilidade é que o pesquisador deixe claro o passo a passo de como os dados foram coletados e analisados. O leitor não deve ser tratado como detetive ou clarividente. Pelo contrário, o pesquisador deve deixar claro todos os procedimentos metodológicos adotados na elaboração de sua pesquisa. Uma pesquisa que não descreve exatamente como os dados foram coletados torna-se irreplicável e, conseqüentemente, não falseável¹³.

2.2 Dimensão substantiva

De acordo com Gleditsch, Metelits e Strand (2003), autores que disponibilizam seus dados são duas vezes mais citados do que aqueles que não o fazem. Tem-se aqui o primeiro incentivo para compartilhar bases de dados¹⁴. Uma segunda vantagem do compartilhamento de bases de dados é a replicabilidade (replication), ou seja, o processo pelo qual novas análises podem ser realizadas a partir de um banco de dados já existente com o objetivo de aprimorar os resultados de pesquisa. A replicação permite que outros pesquisadores melhorem nossas análises, demonstrando inclusive que nossos resultados estavam errados. Por exemplo, Dewald, Thursby e Anderson (1986) replicaram os resultados de diferentes artigos e constataram que erros sistemáticos são frequentes. Resultados de pesquisa de especialistas que compartilham bases de dados são mais facilmente replicáveis e, portanto, mais facilmente falseáveis¹⁵. Para o pesquisador comprometido com o desenvolvimento do conhecimento científico, quanto mais falseáveis forem seus resultados, tanto melhor. Quão frutífero seria para o desenvolvimento da Ciência Política se ao acessar um artigo qualquer, o estudante/pesquisador pudesse ter acesso aos dados utilizados? Meier (1995) identifica

11 No original, “*replication standard holds that sufficient information exists which to understand, evaluate and build upon a prior work if a third party could replicate the results without any additional information from the author*” (KING, 1995: 444).

12 Para Herrson (1995), *replication* e *reanalysis*, *verification* ou *secondary analysis* são conceitos diferentes. A reanálise investiga o mesmo problema examinado por outro estudo. Ela pode ou não utilizar a base de dados original. Se o pesquisador coleta os dados de forma independente (nova coleta de dados) tem-se a replicabilidade. Se o banco de dados original é utilizado tem-se a verificação. A análise secundária consiste em utilizar um banco de dados já existente para investigar outro problema de pesquisa. Dentro dessa definição, a replicabilidade apenas ocorre quando há um novo esforço de coleta de dados.

13 Para King, Keohane e Verba (1994), “*scholars should always record the exact methods, rules, and procedures used to gather information and draw inferences so that another researcher can do the same thing and draw the same conclusion*” (KING, KEOHANE e VERBA, 1994: 27).

14 Para Fienberg (1994), *data sharing*: (1) *reinforces open scientific inquiry*; (2) *Encourages diversity of analysis and opinions*; (3) *promotes new research and allows for testing new or alternative methods*; (4) *reduces costs by avoiding duplicate data collection efforts*; e (5) *provides an important resource for training in research*.

15 Para Popper (1968) não importa se nossas hipóteses de pesquisa foram extraídas da teoria, da observação empírica, ou se simplesmente tropeçamos nelas: em termos metodológicos, o que realmente importa é se elas são falseáveis. O falsificacionismo consiste em tentar repetidas vezes demonstrar, a partir da realidade, que uma determinada teoria é falsa. Quanto mais a teoria resistir à tentativa de ser falseada, tanto melhor será a teoria.

dois principais benefícios associados ao padrão de replicabilidade: “the author is more sure that the results are valid, and the study has greater credibility simply because the author is willing to release the data and documentation to others” (MEIER, 2006: 456).

Nesse sentido, a primeira dimensão do padrão de replicabilidade diz respeito ao aspecto substantivo da atividade de pesquisa. Quanto mais replicáveis forem os resultados de pesquisa, tanto mais rápido é o falseamento de teorias. Além disso, a disponibilização de bases de dados facilita que outros pesquisadores possam investigar a mesma questão ou até mesmo uma questão totalmente nova a partir de um banco de dados já existente.

2.3 Dimensão pedagógica

Pedagogicamente, o padrão de replicabilidade facilita que estudantes deem seus primeiros passos no mundo da pesquisa empírica. Para King (1995), “having students replicate the results of existing articles has proven to be an effective teaching tool” (KING, 1995: 447). Por exemplo, aprender que a correlação de Pearson mede o grau de associação entre peso e altura é importante. No entanto, a aprendizagem é mais eficiente quando os estudantes podem estimar o grau de associação entre variáveis utilizadas em desenhos de pesquisa em Ciência Política.

Na Universidade Federal de Pernambuco (UFPE), o professor Enivaldo Carvalho da Rocha incorporou o padrão de replicabilidade como um dos componentes da disciplina Estudos Avançados de Metodologia de Pesquisa (EAMP), ofertada aos alunos do doutorado em Ciência Política. Primeiramente, os alunos são orientados no sentido de ler alguns textos básicos sobre o padrão de replicabilidade. Depois disso, o monitor da disciplina apresenta diferentes bases de dados publicamente disponíveis que podem ser utilizadas. Por fim, os estudantes são orientados a elaborar um desenho de pesquisa e/ou incorporar ao seu projeto de doutorado uma ou mais bases de dados, formular e testar uma hipótese teoricamente orientada. Cada estudante elabora um produto diferente que é utilizado como um dos critérios de avaliação da disciplina. Como fruto dos esforços pedagógicos do professor Rocha, registra-se a publicação de nove artigos sobre métodos e técnicas de pesquisa. Em 2011: (1) O que fazer e o que não fazer com a regressão: pressupostos e aplicações do modelo linear de mínimos quadrados ordinários¹⁶, publicado pela Revista Política Hoje; (2) What is R2 all about?¹⁷, publicado pela revista Leviathan. Em 2012: (3) Levando Gary King a sério: desenhos de pesquisa em Ciência Política¹⁸, publicado pela Revista Eletrônica de Ciência Política; (4) Classificando regimes políticos utilizando análise de conglomerados¹⁹, veiculado pela revista Opinião Pública. E em 2013 (5) When is statistical significance not significant?²⁰ publicado pela Brazilian Political Science Review; (6) Causalidade e mecanismos em Ciência Política²¹, veiculado pela Revista Mediações; (7) Inferências causais falsificáveis em Ciência Política²², publicado pela Revista Eletrônica de Ciência Política;

16 <http://www.revista.ufpe.br/politica/hoje/index.php/politica/article/view/117/88>

17 http://www.flch.usp.br/dcp/leviathan/index.php/leviathan/article/viewFile/85/pdf_12

18 <http://ojs.c3sl.ufpr.br/ojs2/index.php/politica/article/view/29614/21569>

19 <http://www.scielo.br/pdf/op/v18n1/v18n1a06.pdf>

20 <http://www.bpsr.org.br/index.php/bpsr/article/view/154/144>

21 <http://www.uel.br/revistas/uel/index.php/mediacoes/article/view/17664/14109>

22 <http://ojs.c3sl.ufpr.br/ojs/index.php/politica/article/view/29613/21587>

(8) Análise de componentes principais para construção de indicadores sociais²³, publicado pela Revista Brasileira de Biometria; e (9) Corra que o survey vem aí: noções básicas para cientistas sociais²⁴, publicado pela Revista Latinoamericana de la Investigación Social. Atualmente, parte da equipe do projeto está trabalhando em quatro diferentes artigos sobre metodologia: (a) Regressão logística na ciência política: uma introdução; (b) Tudo que você sempre quis saber sobre análise de dados, mas tinha medo de perguntar; (c) Cluster analysis for dummies e (d) Tobit or not tobit: making the most with censored data. Para 2014, o artigo Desvendando os mistérios da correlação de Pearson: o retorno, já se encontra no prelo da Revista Leviathan.

No presente momento, parte da equipe do SIGOBr está trabalhando no sentido de incorporar o padrão de replicabilidade nas ementas das disciplinas de métodos quantitativos ofertadas na graduação. Ao que tudo indica, devemos iniciar os trabalhos com palestras e aulas expositivas sobre a importância da replicabilidade. Depois disso, o próximo passo é ensinar aspectos básicos de como replicar um trabalho científico. Esperamos incorporar esse padrão na Universidade Federal de Alagoas (UFAL).

2.4 Dimensão transparência²⁵

A terceira dimensão importante do padrão de replicabilidade é a transparência. O matemático inglês Charles Babbage (1830), em *Reflections on the Decline of Science in England*, identificou três diferentes modalidades de práticas fraudulentas: (1) Aparar; (2) Peneirar e (3) Falsificar. O aparamento consiste no ajustamento dos dados no sentido de deixá-los mais precisos. Aqui o pesquisador leva às últimas consequências a máxima de que se “devidamente torturados os dados sempre confessam”. O peneiramento consiste em reportar seletivamente os resultados observados. O truque aqui é escolher, deliberadamente, os resultados de pesquisa que corroboram a hipótese de trabalho. A falsificação, por sua vez, consiste na invenção de dados com o objetivo de descrever um determinado padrão²⁶.

Um trabalho publicado no *Proceedings of the National Academy of Sciences* analisou 2.047 artigos veiculados e, posteriormente, desacreditados na área de Biomedicina desde 1973. Em mais de 40% dos casos foi detectado fraude. Em 23% o problema foi plágio²⁷.

Em 2009, a Corte de Apelação (Court of Appeals) norte-americana confirmou a decisão do Corte Distrital, indeferindo um pedido de ação de difamação de Lott contra Levitt. O nome de Lott foi citado no livro *Freaknomics* como exemplo de trabalho não replicável. Lott argumentou que no contexto citado ele estaria sendo acusado de desonestidade acadêmica. Em 2011, veio a público o caso do psicólogo social Diederik Stapel. Considerado um dos mais prestigiados pesquisadores da Holanda, tendo publicado artigo na revista *Science*. De acordo com as investigações, Stapel inventava a maior parte dos seus dados. Curiosamente, seus estudantes de doutorado reportaram que Stapel nunca lhes permitia participar da

23 http://jaguar.fcav.unesp.br/RME/fasciculos/v31/v31_n1/A5_Dalson_Ranulfo.pdf

24 <http://relmis.com.ar/ojs/index.php/relmis/article/view/74/61>

25 De acordo com a reportagem veiculada pela Folha de São Paulo (<http://www1.folha.uol.com.br/ciencia/896418-quimico-da-unicamp-e-acusado-de-fraudar-11-estudos-cientificos.shtml>), em 16 anos as fraudes científicas cresceram 161% nos Estados Unidos.

26 Para maiores informações ver <http://www.fraudes.org/showpage1.asp?pg=328>

27 Ver <http://hypescience.com/biomedicina-fraudes-aumentam-em-estudos-cientificos/>

coleta dos dados²⁸.

Em 2004, o sul-coreano Woo-Suk Hwang forjou dados de um artigo publicado na prestigiada Science. Dois anos depois a mentira foi descoberta e os resultados desacreditados pela comunidade acadêmica. Em 2002, o norte-americano George Ricaurte da Universidade John Hopkins (EUA) também teve seus achados de pesquisas desacreditados. Não por fraude deliberada, mas sim por erro na manipulação do tratamento em seu experimento.

Em 2011, agora no Brasil, uma investigação apontou fraude em 11 artigos científicos publicados por um professor titular de Química da Universidade Estadual de Campinas (UNICAMP). De acordo com a reportagem veiculada pela Folha da São Paulo, a Elsevier afirmou que os sinais de manipulação são conclusivos. De acordo com a investigação, o pesquisador teria falsificado imagens de ressonância magnética. Em síntese, a replicabilidade protege a comunidade científica não só de fraudes deliberadas, mas também de erros honestos e das limitações técnicas dos pesquisadores.

King (1995) afirma que algumas medidas podem incentivar o padrão de replicabilidade. A primeira delas é incluir a disponibilização de banco de dados como um indicador de produtividade acadêmica como a publicação de um artigo ou livro. Outra medida é exigir que alunos de mestrado e doutorado disponibilizem os bancos de dados utilizados em suas pesquisas. O terceiro envolve a política de publicação dos periódicos especializados. Alguns periódicos em Ciência Política já exigem a disponibilização do banco de dados em algum repositório de acesso público como condição para publicação do artigo. Boyer (2003) afirma que depois da publicação do dossiê sobre replicabilidade, muitos periódicos adotaram padrões de compartilhamento de dados. Em particular, dos 28 periódicos analisados em 2001, 18% adotavam uma política explícita de replicabilidade. Ao se comparar por área, os periódicos de Ciência Política (13%) apresentaram menor tendência à replicabilidade quando comparados com os periódicos de Relações Internacionais (23%). Passados mais de 10 anos desde esse diagnóstico, muitos outros periódicos internacionais adotaram alguma política de replicabilidade como padrão de submissão de artigos.

No plano internacional tem-se o Inter University Consortium for Political and Social Research (ICPRS) da Universidade de Michigan. No Brasil, o Consórcio de Informações Sociais (CIS) conta com um acervo significativo de bases de dados que podem ser publicamente acessadas, mediante cadastro no sistema²⁹. No que diz respeito a periódicos, destacamos a iniciativa da Revista Leviathan, que além de exigir o banco de dados como critério de publicação, aceita artigos em formato LaTeX. Comparativamente, o Dataverse Network, organizado pelo The Institute for Quantitative Social Science da Universidade de Harvard, disponibiliza o maior repositório de dados em Ciências Sociais do mundo. O sistema é

28 Ver <http://veja.abril.com.br/noticia/brasil/holandeses-se-surpreendem-com-fraude-de-prestigiado-psicologo>

29 No Brasil, o Consórcio de Informações Sociais (CIS) é um sistema de intercâmbio de informações científicas sobre a sociedade brasileira. Tem como objetivo oferecer gratuitamente dados qualitativos e quantitativos resultantes de pesquisas sobre vários aspectos da vida social. O pressuposto central do CIS é que as pesquisas produzem dados que são passíveis de reutilização para outros propósitos, diferentes dos originais. Para permitir esta reutilização, o CIS oferece aos seus usuários não só bancos de dados, mas também a literatura científica já produzida com seu uso. Com isso, amplia-se a infraestrutura de informações disponíveis para a pesquisa social. O Consórcio de Informações Sociais é mantido pelo Departamento de Sociologia da Universidade de São Paulo e pela Associação Nacional de Pós-graduação e Pesquisa em Ciências Sociais (ANPOCS) e conta com suporte material e financeiro da USP e CNPq.

público e bastante amigável.

No entanto, apesar dessas iniciativas, não existe uma política sistemática de incentivo ao compartilhamento de dados por parte da comunidade acadêmica. King (1995) questiona a importância das inferências produzidas por um desenho de pesquisa não replicável, afirmando que “at a minimum, some protection should be afforded to keep researchers from wasting their time reading these works” (KING 1995, p. 445). Concordamos com o professor King.

O CASO DO SIGOBr

O projeto Sistema de Informações de Governo (SIGOBr) é conduzido pelo Departamento de Ciência Política (DCP) da Universidade Federal de Pernambuco (UFPE). Em particular, o projeto é coordenado pelo professor Enivaldo Carvalho da Rocha e conta com a participação de pesquisadores de diferentes instituições (ver anexo). Substantivamente, o SIGO é composto por cinco atividades centrais: (1) compilação de bases de dados; (2) replicabilidade; (3) repositório de dados; (4) cursos de capacitação e (5) divulgação. O objetivo central é integrar diferentes práticas desejáveis em um único produto.

Para tanto, a compilação de bases consiste na junção e agregação de diferentes bancos de dados sobre diferentes temas/assuntos em dois principais níveis de análise: estadual e municipal. Em termos substantivos, espera-se munir acadêmicos e formuladores de políticas públicas de informações não só facilmente acessíveis, mas principalmente altamente confiáveis.

A segunda atividade é a replicabilidade. A meta é incentivar a replicação de trabalhos já publicados com o objetivo de refinar os resultados de pesquisa, treinar alunos de graduação e pós-graduação e verificar os resultados de pesquisa contra eventuais erros e/ou fraudes. Esperamos consolidar a metodologia de replication como procedimento padrão não só nas Ciências Sociais em geral, mas também em outros ramos do conhecimento científico e na Ciência Política em particular.

De forma complementar, o SIGOBr funcionará como repositório permanente de dados para fins de pesquisa, consolidando a terceira atividade. A meta aqui é emular o que vem sendo feito pelo Dataverse no plano internacional e pelo CIS no âmbito nacional.

A quarta atividade prevista é a oferta de cursos de capacitação em métodos quantitativos. O objetivo é fornecer treinamento técnico específico para indivíduos e/ou instituições interessados em aprender técnicas de análise de dados. Os cursos poderão ser agendados através do sítio eletrônico do projeto.

Por fim, o projeto prevê atividades de divulgação dos resultados (artigos finalizados, artigos em desenvolvimento, palestras e seminários) com a finalidade de publicizar esses dados. A seguir apresentamos a proposta do projeto do SIGOBr.

1. Identificação do Projeto

Título do Projeto: **SIGOBr – Sistema de Informações sobre Governo no Brasil**

Palavras-Chaves: **1.Banco de Dados; 2.Estados; 3.Municípios; 4.Dados Agregados**

Enivaldo Rocha, Ranulfo Paranhos e Dalson Figueiredo Filho

- Sede:** UFPE - Universidade Federal de Pernambuco
DCP - Departamento do Ciência Política
Recife-PE
- Financiamento:** CNPq – Centro Nacional de Desenvolvimento Científico e Tecnológico
- Parcerias:** IPESPE – Instituto de Pesquisas Sociais, Políticas e Econômicas
Praça Dr. Fernando Figueira, nº 30, 10º andar - Ilha do Leite
CEP: 50070-520 – Recife- Pernambuco - Brasil
- UFAL - Universidade Federal de Alagoas**
Instituto de Ciências Sociais
Campus A. C. Simões - Av. Lourival Melo Mota, Cid. Universitária - Maceió - AL
CEP:57072-900
- UFAL - Universidade Federal de Alagoas**
Instituto de Ciências Sociais
Campus A. C. Simões - Av. Lourival Melo Mota, Cid. Universitária - Maceió - AL
CEP:57072-900
- UFC – Universidade Federal do Ceará**
Av. da Universidade, 2853 - Benfica - Fortaleza – CE, CEP 60020-181
- UFMG - Universidade Federal de Minas Gerais**
Depto. de Ciência Política
Av. Antonio Carlos, 6.627 - Prédio da FAFICH - 4º andar - Campus Pampulha
Belo Horizonte, Minas Gerais, Brasil. CEP: 31270-901
- IFMT - Instituto Federal de Educação, Ciência e Tecnologia do Mato Grosso**
R. Comandante Costa, 1144, Ed. Tarcom - Sala 12, Centro
Cuiabá –MT CEP 78020-400
- UNICAP – Universidade Católica de Pernambuco**
Rua do Príncipe, 526 - Boa Vista - Recife, PE - Cep: 50050-900
- Faculdade Carlos Drummond de Andrade**
R. Professor Pedreira de Freitas, 401/405 - Tatuapé - São Paulo / SP

Membros da equipe

Função	Nome	Currículo Lattes	Instituição
01 Coordenador	Enivaldo Carvalho da Rocha (enivaldocrocha@gmail.com)	http://lattes.cnpq.br/3629594575120398	Univ. Fed. de Pernambuco (UFPE)
02 Pesquisador	Ernani Rodrigues de Carvalho Neto (ernani_carvalho@hotmail.com)	http://lattes.cnpq.br/7185025416007213	Univ. Fed. de Pernambuco (UFPE)
03 Pesquisador	Dalson Britto Figueiredo Filho (dalsonfigueiredofo@ufpe.br)	http://lattes.cnpq.br/6683806605359913	Univ. Fed. de Pernambuco (UFPE)
04 Pesquisador	Mariana Batista da Silva (mariana.bsilva@gmail.com)	http://lattes.cnpq.br/5144800413272494	Univ. Fed. de Pernambuco (UFPE)
05 Pesquisador	José Mário Wanderley Gomes Neto (jmariow@hotmail.com)	http://lattes.cnpq.br/8519132753277329	Univ. Católica de Pernambuco (UnicaP) / Univ. Fed. de Pernambuco (UFPE)
06 Pesquisador	Diego H. Leonel de Oliveira Costa (diegocosta21@yahoo.com.br)	http://lattes.cnpq.br/3502805847381596	Univ. Fed. de Pernambuco (UFPE)
07 Pesquisador	Ranulfo Paranhos dos Santos Filho (ranulfoparanhos@me.com)	http://lattes.cnpq.br/5670542372454107	Univ. Fed. de Alagoas (UFAL) / Univ. Fed. de Pernambuco (UFPE)
08 Pesquisador	José Alexandre da Silva Júnior (jasjunior2007@yahoo.com.br)	http://lattes.cnpq.br/5100245942386773	Univ. Fed. de Alagoas (UFG)

A Importância da Replicabilidade na Ciência Política

09	Pesquisador	Manoel Leonardo W. D. Santos (manoelsantos@fafich.ufmg.br)	http://lattes.cnpq.br/2653461047567152	Univ. Fed. de Minas Gerais (UFMG)
10	Pesquisador	Jakson Alves Aquino (jaa@ufc.br)	http://lattes.cnpq.br/8056757049654019	Univ. Fed. do Ceará (UFC)

2. Resumo:

O principal objetivo desse projeto é elaborar um banco de dados para analisar diferentes temas/assuntos referentes às unidades da federação (municípios/estados). Em termos substantivos, espera-se munir acadêmicos e formuladores de políticas públicas de informações não só facilmente acessíveis, mas principalmente altamente confiáveis. Metodologicamente, utilizamos análise documental com o objetivo de replicar as principais qualidades de bancos de dados já existentes, enfatizando as características do Quality of Government compilado pela universidade de Gothenburg. Além disso, utilizamos diferentes técnicas de estatística multivariada para analisar os dados compilados pelo SIGOBr. Com esse projeto, esperamos consolidar a metodologia de replication como procedimento padrão não só nas Ciências Sociais, mas também em outros ramos do conhecimento científico. De forma complementar, o SIGOBr funcionará como repositório permanente de dados para fins de pesquisa. Por fim, o projeto prevê atividades de divulgação dos resultados (palestras e seminários) com a meta de publicizar esses dados além de oferecer treinamento específico na área de metodologia quantitativa.

3. Etapas/Atividades

Os dados serão coletados em dois principais níveis de análise: municipal e estadual. No entanto, para aumentar a acessibilidade, o SIGOBr disponibilizará informações em outros níveis (regional e nacional).

ATIVIDADE	DESCRIÇÃO
Compilar os dados temáticos por estados/municípios	O SIGOBr vai contemplar diferentes temas (educação, segurança pública, saúde, meio ambiente, etc.), tomando como parâmetro comparativo os valores máximos e mínimos de cada unidade da federação. O objetivo é oferecer um diagnóstico dos principais avanços e problemas por municípios e Estados, munindo não só o gestor público, mas também a comunidade acadêmica de informações precisas e confiáveis.
Replication	De acordo com Gleditsch, Metelits e Strand (2003), autores que disponibilizam seus dados de pesquisa são duas vezes mais citados do que aqueles que não o fazem. Tem-se aqui o primeiro incentivo para compartilhar bases de dados. Uma segunda vantagem do compartilhamento de bases de dados é a replicabilidade (<i>replication</i>), ou seja, o processo pelo qual novas análises podem ser realizadas a partir de um banco de dados já existente com o objetivo de aprimorar os resultados de pesquisa. A replicação permite que outros pesquisadores possam melhorar nossas análises, demonstrando inclusive que nossos resultados estavam errados. Dessa forma, um dos objetivos do SIGOBr é difundir a metodologia de <i>replication</i> como padrão a ser seguido nas ciências sociais de forma geral e na ciência política brasileira, em particular.

Repositório de Dados	Um componente indissociável da metodologia de <i>replication</i> é o compartilhamento de dados. King (1995) afirma que algumas medidas podem incentivar o padrão de replicabilidade. A primeira delas é incluir a disponibilização de banco de dados como um indicador de produtividade acadêmica tal como a publicação de um artigo ou livro. Outra medida é exigir que alunos de mestrado e doutorado disponibilizem os bancos os dados utilizados em suas pesquisas no endereço eletrônico do programa de pós-graduação a que estão vinculados.
Cursos de capacitação	Com o objetivo de facilitar o uso do <i>SIGOB</i> , pretendemos ofertar a cursos de capacitação que auxiliem o uso da plataforma a professores e pesquisadores de universidades e centros pesquisa do Brasil.
Divulgação	Como estratégia de divulgação, prevemos a apresentação de seminários acerca da plataforma <i>SIGOB</i> para professores, pesquisadores e alunos das mais variadas áreas do conhecimento, em universidades e centros de pesquisa do Brasil. Outra estratégia eficaz é a elaboração de um tutorial em forma de vídeo que será disponibilizado no próprio endereço eletrônico do <i>SIGOB</i> .

4. Definição do Problema

Estimar em que medida as políticas públicas produzem os seus efeitos esperados é um dos principais desafios enfrentados pelos estudiosos do assunto. Com efeito, esse argumento ganha ainda mais força ao se considerar a realidade de países que carecem de instituições responsáveis pela coleta, processamento e divulgação sistemática de informações. Fundamentalmente, esse impedimento gera uma série de efeitos perversos. Primeiro, a falta de informações limita as possibilidades objetivas de investigação da realidade. Segundo, restringe a realização de estudos em perspectiva comparada, limitando a difusão de práticas institucionais eficientes. Terceiro, dificulta o controle social sobre a gestão governamental ao reduzir a transparência das ações públicas, violando o princípio da publicidade. Em conjunto, esses obstáculos comprometem não só o aperfeiçoamento das políticas, mas principalmente a sua efetividade. Dessa forma, a compilação de dados tem duas diferentes dimensões: a primeira, mais teórica/acadêmica, se justifica na medida em que a oferta de dados confiáveis é um elemento central na construção do conhecimento científico. A segunda dimensão diz respeito à aplicação prática desse conhecimento seja em pesquisa aplicada seja na formulação de políticas públicas mais eficientes e efetivas.

5. Objetivos

Geral:

Elaborar um banco de dados para analisar diferentes temas/assuntos referentes às unidades da federação (municípios/estados).

Específicos:

1. Facilitar o acesso a diferentes tipos de bancos de dados;
2. Fomentar a produção e a publicação de trabalhos acadêmico/científicos sobre as temáticas presentes no *SIGOB*;
3. Sistematizar informações que auxiliem o gestor público na elaboração e implementação de políticas públicas em diferentes áreas;
4. Disponibilizar o Banco de Dados (*SIGOB*) para a comunidade acadêmica e

em geral através do Consórcio de Informações Sociais da Universidade de São Paulo (CIS/USP³⁰), sítios virtuais do Departamento de Ciência Política da Universidade Federal de Pernambuco (DCP/UFPE³¹), Instituto de Ciências Sociais da Universidade Federal de Alagoas (ICS/UFAL³²) e *Institute for Quantitative Social Science - Dataverse Network* (IQSS/Universidade de Harvard³³).

6. Metodologia

Poucos especialistas discordariam de que a fertilização cruzada de técnicas de pesquisa é benéfica ao conhecimento científico. Inicialmente, esse projeto utiliza análise documental de outros bancos de dados, em particular, do *Quality of Government* compilado pela universidade de Gothenburg, Suécia, com o objetivo de elaborar um modelo ótimo de banco de dados³⁴. Em um segundo momento, adota-se uma perspectiva quantitativa para analisar os dados, utilizando diferentes técnicas multivariadas (regressão linear de mínimos quadrados ordinários, análise fatorial, análise de conglomerados, modelos lineares generalizados, testes de comparação de médias, análise espacial, análise de correspondência, além de testes específicos de associação para variáveis categóricas, etc.).

Para Swanson (1971), “pensar sem comparação é impensável. E, na ausência da comparação, toda a investigação e pesquisa científica” (SWANSON, 1971). Mas por que realizar um projeto em perspectiva comparada? Lijphart (1971) define o método comparativo como um dos métodos básicos - sendo os demais o experimental, o estatístico (N-grande) e o estudo de caso - para estabelecer proposições empíricas gerais (LIJPHART, 1971: 682). Similarmente, Ragin (1987) argumenta que a comparação fornece ao pesquisador a possibilidade de realizar afirmações sobre regularidades empíricas e avaliar e interpretar casos a partir de critérios teóricos e substantivos.

E quais são as vantagens de um desenho de pesquisa em perspectiva comparada? A primeira é observar como diferentes conceitos podem “viajar” para analisar outras realidades. A segunda vantagem da comparação é verificar em que medida os resultados de pesquisa observados em um determinado contexto podem ser encontrados em outros desenhos institucionais. Para os nossos propósitos, trabalharemos com dois principais níveis de análise: municipal e estadual, sempre em perspectiva comparada.

7. Resultados Esperados

1. Contribuir para o fortalecimento do campo de estudos empíricos e para o fortalecimento da teoria positiva e política comparada na área de Ciência Política sobre as unidades da federação;
2. Elaborar um banco de dados para analisar os estados brasileiros em perspectiva comparada;
3. Os resultados da pesquisa serão divulgados na forma de artigos científicos;

30 <http://www.nadd.prp.usp.br/cis/>

31 <http://www.politica.ufpe.br/>

32 <http://www.ufal.edu.br/unidadeacademica/ics>

33 <http://dvn.iq.harvard.edu/dvn/>

34 <http://www.qog.pol.gu.se/>

4. Divulgação dos resultados da pesquisa em congressos científicos da área (IPSA, APSA, ABCP, ANPOCS entre outros);
5. Fortalecimento da Cooperação acadêmica no âmbito nacional entre instituições interessadas em estudar os mais diferentes temas sobre o Brasil.

8. Relevância Científica

Esse projeto de pesquisa visa assegurar a criação de um banco de dados sobre diferentes eixos temáticos (saúde, educação, segurança, agropecuária, política, etc.), atendendo à demanda acadêmica em fornecer informações de forma mensurável/quantificável. Fundamentalmente espera-se facilitar o acesso a uma variedade de informações válidas e confiáveis que possam ser utilizadas não só pelo cidadão comum, mas também pela comunidade acadêmica e gestores públicos na formulação de políticas públicas mais eficientes. Além disso, a criação de um banco de dados com esse teor atende ao princípio da transversalidade das atividades de ensino, pesquisa e extensão que se apresentam como fundamental para a formação acadêmica. A relação estabelecida entre essas dimensões exerce um efeito positivo sobre a qualidade do processo pedagógico. Ao mesmo tempo em que a extensão possibilita a democratização do conhecimento acadêmico e, por meio dela, este conhecimento retorna à universidade, testado e reelaborado, sendo capaz de atender às demandas sociais, estabelecendo um elo entre teoria e prática.

REFERÊNCIAS DO PROJETO SIGOBr

- GLEDITSCH, N. P., Metelits, C. & Strand, H. Posting your data: Will you be scooped or will you be famous? *International Studies Perspectives*, 4, 2003.
- KING, G. "Publication, Publication." *PS: Political Science and Politics*, 39, p. 119-125, 2006.
- KING, G. "The Future of Replication." *International Studies Perspectives*, 4, p. 443-499, 2003.
- KING, G. "Replication, Replication." *PS: Political Science and Politics*, 28, p. 443-499, 1995.
- KING, G., KEOHANE, R., e VERBA, S. *Designing social inquiry: scientific inference in qualitative research*. Princeton: Princeton University Press, 1994.
- LIJPHART, A. "Comparative Politics and the Comparative Method" in *American Political Science Review*, 65:3, pp. 682-693, 1971.
- RAGIN, C. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley and Los Angeles: University of California Press, 1987.
- SWANSON, G. *Framework for Comparative Research: Structural Anthropology and Theory of Action: Essay in Trends and Applications*. Berkeley: University of California, 1971.

CONCLUSÃO

Esse artigo apresentou uma introdução à replicabilidade. Nosso principal objetivo é difundir essa prática como padrão a ser seguido nas Ciências Sociais no Brasil em geral e na Ciência Política em particular. Defendemos aqui que o padrão de replicabilidade apre-

sentia, entre outras características desejáveis, três importantes dimensões: (1) substantiva, na medida em que contribui para o aprimoramento e acúmulo do conhecimento científico; (2) pedagógica, já que facilita a compreensão de noções básicas de análise de dados e (3) transparência, na medida em que protege a comunidade acadêmica não só contra erros honestos, como também de fraudes intencionais. Para Meier (1995), “as a political scientist, I strong support adopting Gary King’s replication policy. I consider it the single most significant contribution in turning political science into a rigorous discipline in my professional lifetime” (MEIER, 1995: 456). Endossamos essa afirmativa e acreditamos que a Ciência Social nacional pode avançar significativamente se optar pelo padrão de replicabilidade. Em 1995, Kenneth Meier previu que em 2010 todos os periódicos em Ciência Política adotariam alguma política de replicabilidade. Ao que parece, no entanto, isso ainda não ocorreu. No Brasil, em particular, a replicabilidade ainda é uma prática pouco difundida entre os cientistas sociais. Esperamos com esse artigo difundir a replicabilidade como procedimento padrão na atividade de pesquisa, ajudando a realização da profecia de Meier.

REFERÊNCIAS BIBLIOGRÁFICAS

- BABBAGE, C. *Reflections on the Decline of Science in England*. London, 1830.
- BOYER, M. A. Symposium on replication in International studies research. *International Studies Perspectives* 4, 72-107, 2003.
- DEWALD, W. G.; J. G. THURSBY, and R. G. ANDERSON. “Replication in empirical economics: The Journal of Money, Credit, and Banking project” *The American Economic Review* 76:587-603, 1986.
- FIENBERG, S. *et al.* (1994). *Sharing Research Data*. Washington, DC: National Academy Press.
- FIGUEIREDO FILHO, D. B.; PARANHOS, R.; ROCHA, E. C. da; SILVA, M. B.; SILVA JUNIOR, J. A.; SANTOS, M. L.; MARINO, J. G. “When is statistical significance is not significant?”. *Brazilian Political Science Review*, v. 7, p. 31-55, 2013.
- FIGUEIREDO FILHO, D. B.; PARANHOS, R.; ROCHA, E. C. da; SILVA JUNIOR, J. A.; MAIA, R. “Análise de componentes principais para construção de indicadores sociais”. *Revista Brasileira de Biometria*, v. 31, p. 61-78, 2013.
- FIGUEIREDO FILHO, D. B.; ROCHA, E. C. da; SILVA JUNIOR, J. A.; PARANHOS, R. Causalidade e Mecanismos em Ciência Política. *Revista Mediações* (UEL), v. 18, p. 10-27, 2013.
- FIGUEIREDO FILHO, D. B., SILVA JÚNIOR, J. A., ROCHA, E. C. “What is R2 all about?” *Leviathan - Cadernos de Pesquisa Política*, v. 3, p. 60-68, 2011.
- _____. “O que fazer e o que não fazer com a regressão: pressupostos e aplicações do modelo linear de Mínimos Quadrados Ordinários (MQO)” *Revista Política Hoje*, v. 20, 1, 44-99, 2011.
- _____. “Classificando regimes políticos utilizando análise de conglomerados.” *Opinião Pública* (UNICAMP. Impresso), 2012.

- FOWLER, L. L. "Replications as Regulation." *PS. Political Science and Politics* 28 (3): 478-481, 1995.
- GIBSON, J. L. "Cautions reflections on a Data-Archiving Policy for Political Science" *PS. Political Science and Politics* 28 (3): 473-476, 1995.
- GLEDITSCH, N. P, METELITS, C. and STRAND, H. "Posting Your Data: Will You be Scooped or Will You be Famous?" *International Studies Perspectives* 4: 89-97, 2003.
- HERRNANSON, P. S. "Replications, Verifications, Secondary Analysis, and Data Collections in Political Science." *PS. Political Science and Politics* 28 (3): 452-455, 1995.
- JUPP, V. *The Sage Dictionary of Social Research Methods*. London: Sage, 2006.
- KING, G. "Replication, Replication." *PS: Political Science and Politics* 28: 443-499, 1995.
- _____. "The Future of Replication." *International Studies Perspectives* 4: 443-499, 2003.
- _____. "Publication, Publication." *PS: Political Science and Politics* 39: 119-125, 2006.
- KING, G., KEOHANE, R. e VERBA, S. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton. N.J.: Princeton University Press, 1994.
- LEVITT, S. *Freakonomics - o lado oculto e inesperado de tudo que nos afeta*. Rio de Janeiro: Elsevier, 2005.
- MAISEL, L. S. (1995). "On the Inadequacy and Inappropriateness of the Replication Standard". *Political Science and Politics* 28(3): 467-70.
- MEIER, K. J. (2006). "Replication: A View from the Streets." *PS. Political Science and Politics* 28 (3): 456-459.
- PARANHOS, R.; FIGUEIREDO FILHO, D. B.; ROCHA, E. C. da; SILVA JUNIOR, J. A. "Corra que o survey vem aí: Noções básicas para cientistas sociais". *Revista Latinoamericana de Metodología de la Investigación Social*, v. 6, p. 7-24, 2013.
- _____. "INFERÊNCIAS CAUSAIS FALSIFICÁVEIS EM CIÊNCIA POLÍTICA". *Revista Eletrônica de Ciência Política - RECP*, v. 4, p. 264-283, 2013.
- POPPER, K. *The Logic of Scientific Discovery*. Londres: Hutchinson, 1968.
- SIEBER, J. E. Introduction to Sharing Social Science Data: Advantages and Challenges, edited by J. E. SIEBER, pp. 1-18 . London: Sage, 1991.

OUTRAS FONTES DE REFERÊNCIAS

- Disponível em <<http://www1.folha.uol.com.br/ciencia/896418-quimico-da-uni-camp-e-acusado-de-fraudar-11-estudos-cientificos.shtml>> Acesso em: 03 fev, 2013.
- Disponível em <<http://www.fraudes.org/showpage1.asp?pg=328>> Acesso em: 03 fev, 2013.
- Disponível em <<http://hypescience.com/biomedicina-fraudes-aumentam-em-estudos-cientificos/>> Acesso em: 03 fev, 2013.
- Disponível em <<http://veja.abril.com.br/noticia/brasil/holandeses-se-surpreen->

A Importância da Replicabilidade na Ciência Política

dem-com-fraude-de-prestigiado-psicologo>Acesso em: 03 fev, 2013.

Disponível em <<http://www.nadd.prp.usp.br/cis/>>Acesso em: 10 set, 2013.

Disponível em <<http://www.politica.ufpe.br/>>Acesso em: 10 set, 2013.

Disponível em <<http://www.ufal.edu.br/unidadeacademica/ics>>Acesso em: 10 set, 2013.

Disponível em <<http://dvn.iq.harvard.edu/dvn/>>Acesso em: 10 set, 2013.

Disponível em <<http://www.qog.pol.gu.se/>>Acesso em: 10 set, 2013.

Disponível em <<http://ojs.c3sl.ufpr.br/ojs2/index.php/politica/article/view/29614/21569>>Acesso em: 10 set, 2013.

Disponível em <<http://ojs.c3sl.ufpr.br/ojs2/index.php/politica/article/view/29614/21569>>Acesso em: 10 set, 2013.

Disponível em <<http://www.scielo.br/pdf/op/v18n1/v18n1a06.pdf>>Acesso em: 29 set, 2013.

Disponível em <<http://www.bpsr.org.br/index.php/bpsr/article/view/154/144>>Acesso em: 29 set, 2013.

Disponível em <<http://www.uel.br/revistas/uel/index.php/mediacoes/article/view/17664/14109>>Acesso em: 07 out, 2013.

Disponível em <<http://ojs.c3sl.ufpr.br/ojs/index.php/politica/article/view/29613/21587>>Acesso em: 07 out, 2013.

Disponível em <http://jaguar.fcav.unesp.br/RME/fasciculos/v31/v31_n1/A5_Dalson_Ranulfo.pdf>Acesso em: 07 out, 2013.

Disponível em <<http://relmis.com.ar/ojs/index.php/relmis/article/view/74/61>>Acesso em: 07 out, 2013.

Disponível em <<http://hypescience.com/biomedicina-fraudes-aumentam-em-estudos-cientificos/>>Acesso em: 07 out, 2013.



UNIVERSIDADE
FEDERAL
DE PERNAMBUCO



Departamento de
Ciência Política
Programa de Pós-Graduação
em Ciência Política