

|                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                     |                                              |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|----------------------------------------------|
| <b>Institution:</b> University of Cambridge                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                     |                                              |
| <b>Unit of Assessment:</b> UoA 21                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |                                                     |                                              |
| <b>Title of case study:</b> DIGIWHIST: using Big Data methods to transform how we measure and manage corruption risks in government contracting                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |                                                     |                                              |
| <b>Period when the underpinning research was undertaken:</b> 2014-2018                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                     |                                              |
| <b>Details of staff conducting the underpinning research from the submitting unit:</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                     |                                              |
| <b>Name(s):</b>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | <b>Role(s) (e.g. job title):</b>                    | <b>Period(s) employed by submitting HEI:</b> |
| Mihály Fazekas                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | Research Associate                                  | 01.03.2014-30.04.2018                        |
| Lawrence Peter King                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | Professor of Sociology and Political Economy (2012) | 01.10.2006-31.08.2017                        |
| <b>Period when the claimed impact occurred:</b> 2014-present (ongoing)                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |                                                     |                                              |
| <b>Is this case study continued from a case study submitted in 2014?</b> No                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |                                                     |                                              |
| <p><b>1. Summary of the impact</b> (indicative maximum 100 words)</p> <p>DIGIWHIST (The Digital Whistleblower) is used by policy-makers, investigators and civil society organisations in the UK and Europe to tackle corruption in government contracting by providing access to novel datasets, measurement tools and policy-relevant evidence. The project has:</p> <ol style="list-style-type: none"> <li>1) had transnational policy impact across major European and global institutions;</li> <li>2) improved anti-corruption practices, including risk assessment and auditing; and</li> <li>3) enabled civil society organisations to monitor public procurement activity and to organise anticorruption action.</li> </ol> <p>DIGIWHIST has significantly shifted international policy debates on anticorruption and public contracting towards a data-driven approach. DIGIWHIST data and indicators now form an integral part of risk-assessment methods employed by investigators and auditors and have been adopted by major organisations such as the Organisation for Economic Co-operation and Development (OECD) to assess country administrative quality. Civil society watchdog web portals running on DIGIWHIST data and indicators have been established across Europe.</p>                                                                                                                                                            |                                                     |                                              |
| <p><b>2. Underpinning research</b> (indicative maximum 500 words)</p> <p>Corruption has featured among the top five concerns of citizens across Europe throughout the 2010s. It carries a variety of significant social, political and economic costs including loss of trust in public institutions. Public procurement amounts to about 15% of GDP annually in the EU. According to the OECD, public procurement is the highest corruption risk in developed economies, even higher than areas traditionally associated with bribery, such as paying taxes or accessing healthcare. Estimates vary, but a recent report put the annual cost of corruption risk in public procurement at around five billion euros [E6].</p> <p>DIGIWHIST promotes government accountability in procurement activities. Our rich, reliable and actionable databases and risk indicators provide usable, low-cost transparency for policy makers, bidding companies and civil society, enabling them to mobilise against corruption and build integrity.</p> <p>The DIGIWHIST research project (2014-18), led by researchers at Cambridge's Department of Sociology (MihályFazekas and Lawrence King) used Big Data methods to mine large amounts of online administrative documents to identify corruption risk patterns in government purchasing processes across 35 European countries and to enhance anticorruption measures using new, objective and reliable data.</p> |                                                     |                                              |

Almost all governments publish procurement data (tenders, contracts etc.) online in a format that only allows reading of a single document at a time. This hinders systematic analysis of large numbers of documents. In 2018, DIGIWHIST developed an automated method for collecting public procurement tender and contract data and organised the information into a flexible database structure freely analysable and downloadable at [www.opentender.eu](http://www.opentender.eu). The DIGIWHIST datasets contain over 19 million contracts worth 5-10% of annual GDP in 32 European countries (including all EU institutions and member states).

Using these large-scale datasets, DIGIWHIST has created new quantitative tools for risk measurement based on an established understanding of high-risk scenarios, such as convoluted tender specifications that seem to be tailored to particular firms and may thus limit the number of bids in an otherwise competitive market [R4]. By estimating the prevalence of such high-risk situations, DIGIWHIST produced corruption risk indicators that can be used to identify risks, both in individual contracts, and (when aggregated) in countries, regions, markets, or government agencies [R2, R3]. The DIGIWHIST risk indicators have been widely adopted by practitioners and academics as tools for the measurement and management of corruption risk.

The datasets and risk indicators created by DIGIWHIST were used by the research team to generate policy-relevant research findings evaluating the effectiveness of anticorruption policies and targeted monitoring interventions. This research activity includes work on the role of checks and balances in preventing state capture in Hungary [R5], and the effectiveness of bureaucratic controls in safeguarding EU Structural Funds from fraud and corruption [R1].

For the work on EU structural funds [R1], DIGIWHIST data from over 100,000 public procurement contracts was analysed to further elaborate DIGIWHIST's framework of corruption risk indicators (see above) and link them to agency-level data in the public sector. Propensity scores matching estimations suggested that EU funds increase corruption risk by up to 34 percent, largely due to overly formalistic compliance and EU Funds overriding domestic accountability mechanisms in public organizations entirely dependent on external funds.

The work on state capture in Hungary [R5] also used DIGIWHIST data to create a novel analytic framework; this time linking DIGIWHIST risk indicators to social network analysis to measure capture as clustered networks of organisations bearing high levels of corruption risk. The power of the new analytical framework was demonstrated by exploring how the radical centralization of the governing elite following the 2010 elections in Hungary affected centralization of state capture.

### 3. References to the research (indicative maximum of six references)

- R1: Fazekas, M., and King, L.P. (2018). Perils of development funding? The tale of EU Funds and grand corruption in Central and Eastern Europe. *Regulation & Governance*, 13(3), 405-430. [\[DOI\]](#)
- R2: Charron, N., Dahlström, C., Fazekas, M., and Lapuente, V. (2017). Careers, connections and corruption risks: Investigating the impact of bureaucratic meritocracy on public procurement processes. *Journal of Politics*, 79(1), 89-104. [\[DOI\]](#)
- R3: Fazekas, M., and Kocsis, G. (2017). Uncovering high-level corruption: Cross-national objective corruption risk indicators using public procurement data. *British Journal of Political Science*, 50(1), 155-164. [\[DOI\]](#)
- R4: Fazekas, M., Tóth, I. J., and King, L. P. (2016). An objective corruption risk index using public procurement data. *European Journal on Criminal Policy and Research*, 22, 369-397. [\[DOI\]](#)

R5: Fazekas, M. and Tóth, I. J. (2016). From corruption to state capture: A new analytical framework with empirical applications from Hungary. *Political Research Quarterly*, 69(2), 320-334. [DOI](#)

This research has been published in peer-reviewed and reputable publications and therefore meets the 2\* minimum REF requirement. It was supported by the EU Horizon2020 grant, *The Digital Whistleblower: Fiscal Transparency, Risk Assessment and Impact of Good Governance Policies Assessed*, led by King (01.03.2015 – 28.02.2018) GBP1.26M

#### 4. Details of the impact (indicative maximum 750 words)

Over 1,000 weekly users [E1] visit the DIGIWHIST data portal which gives free public access to over 19 million contracts across European countries and EU institutions. Users include government agencies, watchdog groups, journalists, researchers and NGOs. For the first time, European public procurement data are directly accessible in a structured format, increasing the transparency and accountability of public officials and public organisations across Europe. DIGIWHIST enables organisations to hold public officials to account, and it allows public organisations to fulfil their regulatory functions by enabling measurement and management of corruption risk. DIGIWHIST watchdog portals reach several thousand users each week. The overall reach of the portal is likely to be around 100,000 users per annum (unique users based on IP address as reported by Google analytics).

DIGIWHIST has had international impact on public procurement and anticorruption policy, oversight institutions carrying out investigations and audits, and civil society organisations active in the anticorruption movement.

##### 1) Transnational policy impact: European and global institutions

DIGIWHIST has had impact on transnational institutions at both European and global levels. Hitherto, corruption measurement has relied on investigated or audited cases, a method that underestimates corruption. DIGIWHIST has transformed practice, using a big data approach and quantitative risk indicator tools to enable real-time monitoring and targeted audits.

- The European Court of Auditors has recommended that the European Commission adopt the DIGIWHIST approach in order to enhance the measurement and management of corruption and fraud risks across the EU [E2, p.21, p.77].
- The OECD's work on data-driven risk assessment methodologies [E3, p.44] has used DIGIWHIST tools and research findings to measure the scale of corruption and corporate policy capture [E4] and is promoting this approach across the world [E5, p.10-11]. This represents a significant shift for the OECD - from policy analysis relying on opinion surveys and expert interviews towards a Big Data approach. OECD [E3] has highlighted the application of DIGIWHIST risk-assessment tools by the Mexico City International Airport, citing it as an exemplar of how the approach can be used in large-scale public projects.
- The European Parliament's flagship study on the cost of corruption across the EU devoted an entire chapter to DIGIWHIST, using our data and methodology to produce a novel estimate for the cost of corruption [E6, section 2.3]. This supported the European Parliament's push for stronger corruption controls in the EU. The European Commission's Directorate-General for Regional and Urban Policy (DG REGIO) flagship

report on EU territorial cohesion has used DIGIWHIST research to gain an insight into the quality of regional governance [E7, p158]. These two cases exemplify the impact in supporting the marked shift in EU policies on public procurement and anticorruption towards one based on risk indicators derived from Big Data methods.

## 2) Improving anti-corruption practices

DIGIWHIST's work with national and international anticorruption and public procurement agencies has led to tangible improvements in measuring the progress of anti-corruption policies.

- The OECD/Sigma program and the European Commission use DIGIWHIST indicators and benchmarks for measuring accession and neighbourhood countries' progress towards EU norms and the fulfilment of accession criteria [E8, p225].
- The World Bank promotes the DIGIWHIST approach in its lending and grants projects across the world,  
 "Even beyond Europe, the project's public procurement data analytics have had a significant impact as they have inspired and contributed to new ways of approaching the assessment of corruption risk and public integrity in a variety of governance contexts. For the World Bank, the analytical methods and corruption risk indicators developed by DIGIWHIST have informed out thinking about our own risk indicators and helped us identify what is feasible with Big Data analytics in this domain. Furthermore, we have used DIGIWHIST's analytical approaches in providing technical assistance to national governments, for example in Mexico" [E9, Senior Governance Specialist, Governance Global Practice, The World Bank].

## 3) Impact on major organisations' risk assessment and auditing

The European Investment Bank (EIB) finances projects worth over GBP45 billion annually. DIGIWHIST has worked with EIB on multiple events and projects, and its Fraud Investigations Division states that DIGIWHIST has had a broad impact on their approach to risk measurement and management. Since 2018, the organisation has shifted from reliance on whistle-blower reports, towards a more data-driven approach to risk assessment based on adoption of DIGIWHIST's corruption risk indicators [E3, p.44].  
 "Thanks to DIGIWHIST's work, public procurement data across Europe have been systematically collected, structured, analysed, and disseminated, increasing transparency and accountability of public officials across the continent. The project's public procurement data analytics have had a broad impact on the thinking of the Inspectorate General on measuring and managing risks of EIB operations. Through a number of co-organised events and projects, the European Investment Bank was supported by DIGIWHIST data and indicators. Among others, we have incorporated DIGIWHIST's corruption risk indicators for quantitative risk-scoring in order to select high-risk counterpart organizations (borrowers) for our Proactive Integrity Reviews with the aim of mitigating risks and avoiding large financial losses," [E10, Head of Fraud Investigations Division, and Head of PIR Unit, Inspectorate General, European Investment Bank].

- DIGIWHIST data has supported audits and criminal investigations by various investigative bodies, including the EU's European Anti-fraud Office and the Romanian anticorruption agency. While the exact impact is not public due to privacy considerations,

DIGIWHIST data made audits and investigations more efficient by reducing data collection costs.

#### 4) Impact on civil society

There are a great number of NGOs across Europe which monitor public procurement corruption and organise anticorruption action using DIGIWHIST. Most prominently, a growing number of online watchdog portals have been created across Europe, especially in Eastern Europe, focusing on government contracting. Many of these portals, such as those in the Netherlands and Poland [E11], use DIGIWHIST data and indicators. This increase in civil society monitoring of corruption has been enabled by Digiwhist's work providing online access to procurement data in a format that allows for systematic large-scale data analysis, as the statement from the Founder of Barometr Ryzyka Nadużyć makes clear, "Thanks to the DIGIWHIST team and Dr Mihaly Fazekas we have created the Corruption Risk Barometer [E11], an interactive tool for the citizens, journalists and researchers measuring the risk of corruption on the Polish public procurement market," [E12 Founder of Barometr Ryzyka Nadużyć, Head of the Public Integrity Program at the Stefan Batory Foundation]. This democratisation of procurement monitoring demonstrates the scale and reach of Digiwhist's impact.

#### 5. Sources to corroborate the impact (indicative maximum of 10 references)

E1: Websites: Making Public Tenders More Transparent.

E2: European Court of Auditors. (2019) *Fighting fraud in EU spending: action needed*. Special report No 01/2019.

E3: OECD. (2020). *Analytics for integrity: Data-driven approaches for enhancing corruption and fraud risk assessment*.

E4: OECD forum on financing democracy and averting policy capture. Paris, 3-4 December 2014.

E5: OECD. (2017). *Preventing policy capture: Integrity in public decision making*. OECD Public Governance Reviews. OECD Publishing. [\[DOI\]](#)

E6: Hafner, M., Taylor, J., Disley, E., Thebes, S., Barberi, M., Stepanek, M., and Levi, M. (2016). *The cost of non-Europe in the area of organised crime and corruption: Annex II – corruption*. European Parliamentary Research Service. [\[DOI\]](#)

E7: EU Commission. (2017). *My region, my Europe, our future: 7th report on economic, social and territorial cohesion*. Publications Office of the European Union. [\[DOI\]](#)

E8: OECD/Sigma. (2017). *Methodological framework for the principles of public administration*. Sigma.

E9: Testimonial: Senior Governance Specialist, Governance Global Practice, The World Bank.

E10: Testimonial: Head of Fraud Investigations Division, and Head of PIR Unit, Inspectorate General, European Investment Bank.

E11: Websites: (i) Elvis [\[Link\]](#) (ii) Corruption Risk Barometer (*Barometr Ryzyka Nadużyć*) [\[Link\]](#)

E12: Founder of Barometr Ryzyka Nadużyć, Head of the Public Integrity Program at the Stefan Batory Foundation (Warsaw, Poland) and assistant professor at the Collegium Civitas (Warsaw, Poland).