Doublethink in governmental accounting: development of an RPA to identify inconsistencies in financial reporting

Luiz Otavio Schmall dos Santos. Fundacao Getulio Vargas and Petroleo Brasileiro S/A (Petrobras), Brazil, schmall@petrobras.com.br

Ricardo Lopes Cardoso. Fundacao Getulio Vargas, Brazil, ricardo.cardoso@fgv.br

Felipe Buchbinder. Fundacao Getulio Vargas, Brazil, felipe.buchbinder@fgv.br

Miklos Vasarhelyi. Rutgers Business School, USA, miklosv@business.rutgers.edu

Abstract. Aiming to assess the reliability of governmental accounting under an armchair-audit approach, we develop a framework to compare the financial reports submitted by municipalities to two different agencies: the Ministry of Finance and the respective Court of Accounts. We developed a framework using concepts of RPA in conjunction with OCR to download, extract, organize, and finally compare the reliability of the financial reports submitted by the municipalities. The results indicate that a framework of RPA is helpful to automate many tasks necessary to armchair-audit municipalities' financial reports. The results also indicate that many Brazilian municipalities submit inconsistent data to the monitoring agencies, i.e., the Ministry of Finance and respective Court of Accounts. Additionally, our findings suggest that more computerized entities are less prone to present inconsistencies in their accounting data.

Keywords: Accounting; robotic process automation (RPA); optical character recognition (OCR); armchair-audit; inconsistency.

1. INTRODUCTION

In Brazil, municipalities render their accounts to multiple agencies. For example, they must submit their complete financial reports both to the Ministry of Finance (for consolidating the whole of government accounts) and to the Court of Accounts (that audits the mayor's accounts). Municipalities must also present specific information to other agencies, such as healthcare expenditures to the Ministry of Health, and education expenditures to the Ministry of Education.

Under an open government approach, financial data from municipalities are readily available to citizens at different websites, such as the municipality's web portal, the Ministry of Finance's repository of public sector financial reports, and at the websites maintained by the Court of Accounts, by the Ministry of Health and by the Ministry of Education. Although it may seem that it guarantees a high level of transparency, this is not the case.

Governmental transparency (both active and passive transparency), either through open government websites or freedom of information, is crucial for citizen engagement and democracy (Koningisor, 2020; Pernagallo & Torrisi, 2020). Hence, a myriad of portals where citizens can access financial information about a single municipality would benefit transparency if the information were consistent. However, discrepancies among "the same" information disclosed at different portals impair its reliability and transparency. Depending on the source where citizens access the municipality's financial reports, they retrieve different (hence, unreliable) information. Considering that the financial reports represent past events and are prepared under a single set of accounting standards despite being disclosed at the Ministry of Finance repository or the Court of Accounts website, we should expect that the financial information from a single municipality made available at different portals do not differ.

This research aims to develop a workflow using robotic process automation (RPA) in conjunction with optical character recognition (OCR) techniques to investigate whether municipalities' financial reports made available at the Ministry of Finance repository and the Court of Accounts website are consistent, i.e., identical. The algorithm developed can be helpful to evaluate the reliability of the data submitted by municipalities to different control entities. The research also aims at investigating the correlation between computerization level and the propensity of presenting unreliable information.

The research applies the concepts of "design science research" (Walls et al., 1992) with the objective of solving a problem of transparency created by the divergence of information made available by the analyzed municipalities and by the difficulty to access the data submitted to the control bodies.

To better insert the research developed here within a theoretical basis of "design science" as proposed by Thuan et al. (2019), we answer the following question: "How should we implement a computational solution capable of indicating possible inconsistencies in financial reports of municipalities?"

To achieve this objective, we developed several source codes in Python, using public libraries, i.e., exploring a crowdsourcing environment; automatically downloaded the financial reports from 721 municipalities (from two states in Brazil, Rio Grande do Sul and Piaui) which are available at Ministry of Finance's repository and their respective Court of Accounts' websites; extracted the appropriate data from the reports, which in many cases are scanned images in PDF files (requiring OCR); organized the data; and finally, compared their consistency. This research indicates that an RPA framework can be helpful to automate many tasks necessary to armchair-audit municipalities' financial reports. The results also indicate that many Brazilian municipalities submit inconsistent data to the monitoring agencies, i.e., the Ministry of Finance and the respective Court of Accounts. Therefore, the research contributes on the design science literature by developing a whole framework to automate all the tasks involved in analyzing the reports submitted by municipalities from two different Brazilian states, which can be used by researchers and auditors interested in financial data from Brazilian municipalities or adapted by those interested in analyzing similar data from other governmental entities, including from other jurisdictions. Also, the research improves the understanding about the consistency of the financial reports published and available to the general public.

Citizens should be aware that governments disclose conflicting accounting information among multiple channels. Hence, depending on the open data portal where citizens collect the financial reports, they might perceive a different reality about the government's budgetary and financial performance and position.

The remining sections of this manuscript are organized as follows: in the next section we present the literature review and propositions; section three describes the analyzed context; section four highlights the research problem; section five

describes the methodological procedures; section six presents and discusses the results; section seven presents the conclusions and provides suggestions for future studies.

2. BACKGROUND

The lack of reliability in governmental accounting impairs democracy since the incumbent 'rewrites' the past, such as in the Oceanian province of Airstrip One, from the dystopian "1984" written by George Orwell, where doublethink is a "vast system of mental cheating […] those who have the best knowledge of what is happening are also those who are furthest from seeing the world as it is. In general, the greater the understanding, the greater the delusion; the more intelligent, the less sane. […] The contradictions are not accidental, nor do they result from ordinary hypocrisy; they are deliberate exercises in doublethink. For it is only by reconciling contradictions that power can be retained indefinitely" (Orwell, 2009: 152).

There are also examples from the private sector. Entities that do not have audited financial reports may appear in different datasets with varying values attributed to the same account. Even audited companies might present conflicting information to the tax authority and capital providers. Indeed, these situations remind the anecdote presented by Abraham Briloff (1972) in the "Unaccountable Accounting: Games Accountants Play", where the management of a company about to launch an IPO decided which audit firm to hire based on the answer to the question "How much is 2 plus 2?" The hired audit partner had answered: "What number did you have in mind?" This anecdote suggests that unreliable or even conflicting financial reporting is not a new phenomenon, nor is it constrained to the public or private sector, or to any specific jurisdiction; quite on the contrary, it is a pervasive phenomenon.

Despite such pervasiveness, the evolution of digitalization, accountability, transparency practices, and the consecutive implementation of austerity policies by governments trying to cope with reiterated crises have exhorted citizens to armchair-audit governments (Dai & Li, 2016; O'Leary, 2015; Cesario, 2020). The problem is that depending on the source where citizens collect data; they can perceive a different reality about the analyzed public sector entity and searching for inconsistency in governmental reporting is very time-consuming for humans.

Advances in hardware and computing technologies have exponentially increased since 1947, the year of the transistor invention (Schaller, 1997). This technological revolution influenced all areas of expertise to adopt different forms of automation. Vasarhelyi and Halper (1991) proposed Continuous Process Audit Methodology (CPAM) to increase the efficiency and accuracy of an audit performed at AT&T. This was achieved by designing a system that kept track of all transactions performed, different from the traditional audit procedure, that checks the entire financial reports and respective evidence only once a year.

A vital enabler to make Continuous Auditing possible was the development of Real-Time Accounting (RTA) systems (Rezaee et al., 2001). These systems substituted traditional business transactions documentation with electronic evidence, thereby enabling auditors to keep continuous tracking of all business transactions performed.

 IT resources applied by auditors vary with company size (Janvrin et al., 2008). 181 auditors from national, regional, and local audit firms were interviewed to analyze audit applications, IT productivity tools, and working paper review methodology. Their results indicate that auditors use several audit applications, but large-sized firms use more sophisticated IT tools in their work.

More recently, many studies have been conducted to apply artificial intelligence (AI), data mining, and RPA. Iguma and Riccio (2018) performed a bibliometric analysis and found that data mining and big data were the most cited keywords in studies related to technology application in auditing. Even though the companies are investing in digital transformation, Gotthardt et al. (2020) found that only 15% of the companies consider themselves mature in the use of RPA. The numbers are even lower when considering AI; only 5% of the companies consider themselves mature in the use of AI. Gotthardt et al. (2020) conclude that the successful implementation of RPA and AI in auditing and accounting requires cross-department support.

Lacurezeanu et al. (2020) performed a literature survey about the use of RPA in accounting and auditing. They found that although the theme is new, it is gaining popularity, and there are many examples of applications in the literature recently. Some examples of these applications can be found at Rozario and Vasarhelyi (2018a; 2018b), Huang and Vasarhelyi (2019) and Moffitt et al. (2018) who analyzed the application and the potential impact of RPA in auditing. Huang et al.

(2020) applied OCR with RPA to automate financial bill recognition and classification.

Following the methodology proposed by Geerts (2011), Hevner et al. (2004) and Pfeffers et al. (2006), we identified a problem of transparency in the information provided by the Brazilian municipalities of Piauí and Rio Grande do Sul, since the information has divergences when comparing the data submitted to the Courts of Accounts and the Ministry of Finance. In addition, the data available to the public are scanned images with low quality, thus making it difficult to use automated algorithms.

In this research, we developed a package of algorithms that are capable of automatically downloading the data provided by the courts of accounts and the Ministry of Finance, reading and structuring this data in a database and, automatically, identifying possible inconsistencies in the data presented. The potential inconsistencies appointed by the proposed solution were never evaluated, since the data submitted independently to different control bodies are not compared.

The developed solution was applied to the data provided by the municipalities of Rio Grande do Sul and Piauí to evaluate its efficiency and applicability. Hence, the first proposition:

P1: It is feasible to develop an RPA application to search for inconsistencies in financial reports disclosed at multiple channels.

Much work is invested in the private sector to automate processes and gain efficiency. On the other hand, even in developed countries, there is still a lack of computerization in public service. Joffe and Reck (2019) found that the degree of computerization and accessibility of accounting information from USA state governments is limited. Most USA states publish their reports in PDF format, even with all government incentives to migrate to a more accessible format. They cite cases from Spain and the Brazilian SICONFI (the Accounting and Fiscal Information System of the Brazilian Public Sector) as examples of investment in applying data in XBRL format (eXtensible Business Reporting Language) and data availability to the public.

Salles (2019) studied the relationship between the degree of computerization of states and municipalities and the delivery of financial reports to the Ministry of Finance (through an IT platform known as SICONFI) within the established

deadline. To measure the degree of computerization of the executive power of states and municipalities, Salles (2019) used data on IT expenditures and correlated these expenditures with the deadline for submitting financial reports through SICONFI. The study indicated that less computerized states and municipalities had more significant problems submitting reports on time. Therefore, the second proposition:

P2: More computerized entities provide less conflicting information.

3. CONTEXT

This research uses the financial reports submitted by municipalities from Piauí, and Rio Grande do Sul to the Ministry of Finance (SICONFI) and their respective State Court of Accounts (TCEs). The two states were chosen by their degree of computerization (Salles, 2019). Therefore, we downloaded the data from SICONFI (2020) and the respective TCEs databases (TCE – PI, 2020; TCE – RS, 2020).

The SICONFI database was developed by SERPRO (Federal Data Processing Service) to receive the financial reports submitted by the 5570 Brazilian municipalities. SICONFI applies XBRL technology to receive and store the data. Beyond the advantage of consolidating the data from all Brazilian municipalities, SICONFI is also responsible for standardizing the financial reports and giving high transparency to the public sector accounts in Brazil.

Municipalities must also submit their financial reports to their respective State Court Accounts (TCEs). Each TCE has its own database and procedures. In the case of Piauí's TCE, the financial reports of the municipalities must be submitted in a scanned PDF version and in an electronic version to a system called SAGRES for the exclusive use of governmental auditors. In the Rio Grande do Sul's TCE, the financial reports are submitted directly to an electronic database made available to the general public.

Governmental auditors audit the financial reports municipalities submit to the TCE. In their circularization procedures, they do not necessarily consider verifying whether the information submitted to the TCE is consistent with information mayors present to other stakeholders, such as the Ministries of Finance, Health, or Education.

Piauí is composed of 224 municipalities, with an average monthly household income per capita of BRL 827, Teresina (the capital) being the most populous municipality, with 864,845 inhabitants and Miguel Leão the least populated

municipality, with 1,246 inhabitants. Piauí's total population is estimated at 3,273,227 inhabitants, and 34% of municipalities have less than 5,000 inhabitants (IBGE, 2020).

Rio Grande do Sul comprises 497 municipalities, with an average monthly household income per capita of BRL 1,843, Porto Alegre (the capital) being the most populous municipality, with 1,483,771, and Engenho Velho the least populated municipality, with 1,034 inhabitants. The state's total population is estimated at 11,377,239 inhabitants and 46% of the municipalities have less than 5,000 inhabitants (IBGE, 2020).

4. PROBLEM IDENTIFICATION

Brazilian municipalities are required by law to annually submit their accounting reports to various public entities (control bodies), e.g., Ministry of Finance, State Courts of Accounts, Ministry of Health, Ministry of Education, among others. Each of these public entities has its own database, its respective system, input format, and schedule for receiving data.

The lack of standardization of formats, in addition to the lack of integration between the existing bases, makes it difficult to compare the reports presented, which can potentially generate divergences in them.

As already discussed, there is a lot of research carried out with the objective of evaluating possible inconsistencies in the accounting reports of public entities, but we did not find any work comparing the same data presented to independent bodies.

In this way, we need to develop a package of tools to download data in different databases, convert images and unstructured data into a structured database and finally compare them to detect possible inconsistencies.

5. METHODOLOGY

This research develops a workflow to compare the consistency of the data submitted by Brazilian municipalities to two different monitoring agencies, the Ministry of Finance, and the Court of Accounts. To achieve this goal, we developed a framework to download the financial reports, extract the data, organize it, and check their consistency. We used RPA in conjunction with OCR to download and extract the data presented in the municipalities' reports. The sample comprises the financial reports prepared by all municipalities from the two Brazilian states during two

consecutive years (i.e., 2017 and 2018). The choice was based on the different levels of computerization presented by Piauí, and the Rio Grande do Sul municipalities (Silva, 2019), where the Court of Accounts from the Rio Grande do Sul is more computerized than Piauí.

Hence, we collected data from the Ministry of Finance website (SICONFI, 2020) and the respective Court of Accounts (TCE - PI, 2020; TCE - RS, 2020). The Ministry of Finance website is prepared for download in big chunks, i.e., it is possible to select all the municipalities and download in one click their financial reports in CSV format. On the other hand, the Courts of Accounts websites do not permit downloading several municipalities' reports at once. Indeed, in order to download a single report of each municipality, several clicks are necessary. To overcome this issue, we developed an RPA application to download all the required data at once, i.e., all municipalities' reports from 2017 and 2018. The source code we developed is available on GitHub (Santos, 2021).

Another challenge was the data extraction from each report, especially in the data from the Piauí's Court of Accounts. There, reports were available in scanned images converted to PDF. Thus data extraction required an OCR application. Several commercial solutions were tested (e.g., PDFelement, Adobe Acrobat Pro DC, and OmniPage Ultimate), including AWS Textract service, which was our final choice. The AWS Textract service has the advantage of managing multiple files in a batch, but the user must develop a code to upload the files at the AWS cloud and process the files asynchronously. There are freeware libraries available at GitHub, that were adapted for these needs (AWS - Textract, 2020).

As depicted in Figure 1 (a sample balance sheet retrieved from the Piauí's Court of Accounts) and Figure 2 (a sample budgetary report retrieved from the same Court of Accounts), there are several barriers for the data extraction, e.g., stamps and signatures, low-quality image, holes for archiving and tables not necessarily in the horizontal direction (tilted).

ANEXO 14 - BALANÇO PATRIMONIAL
Dezembro(31/12/2017)
CONSOLIDADO - Desconsiderando as contas INTRA-OFFS

Exercício de 2017   1 of 3

ATIVO			PASSIVO			
ESPECIFICAÇÃO	Exercício Atual	Exercício Anterior	ESPECIFICAÇÃO		Exercício Atual	Exercício Anterior
ATIVO CIRCULANTE	441.798,08	436.120,89	PASSIVO CIRCULANTE		297.226,05	370.539,18
CAIXA E EQUIVALENTES DE CAIXA	290.189,95	449.939,59	OBRIGAÇÕES TRABALHISTAS, PREVIDENCIÁRIAS E ASSISTENCIAI		-187.767,41	0,00
CAIXA E EQUIVALENTES DE CAIXA EM MOEDA NACIONAL	284.211,24	449.939,59	PESSOAL A PAGAR		-138.767,52	0,00
BANCOS CONTA MOVIMENTO - DEMAIS CONTAS F	1.019,57	218,38	PESSOAL A PAGAR DE EXERCICIOS ANTERIORES	F	-138.767,52	0,00
BANCOS CONTA VINCULADAS-DEMAIS CONTAS F	4.436,42	253,82	ENCARGOS SOCIAIS A PAGAR		-48.999,89	0,00
APLICACOES FINANCEIRAS DE LIQUIDEZ IMEDIATA F	278.755,25	449.467,39	CONTRIBUIÇÕES PREVIDENCIÁRIAS - DÉBITO PARCELADO	P	-34.485,30	0,00
CAIXA E EQUIVALENTES DE CAIXA EM MOEDA ESTRANGEIRA	5.978,71	0,00	CONTRIBUIÇÕES SOCIAIS - DÉBITOS PARCELADOS	P	-3.535,09	0,00
APLICACOES FINANCEIRAS DE LIQUIDEZ IMEDIATA F	5.978,71	0,00	OUTROS ENCARGOS SOCIAIS	P	-10.979,50	0,00
DEMAIS CRÉDITOS E VALORES A CURTO PRAZO	149.090,13	-13.818,70	FORNECEDORES E CONTAS A PAGAR A CURTO PRAZO		-8.658,65	0,00
DEPÓSITOS RESTITUÍVEIS E VALORES VINCULADOS	119.693,03	0,00	FORNECEDORES E CONTAS A PAGAR NACIONAIS A CURTO PRAZ		-8.658,65	0,00
CONTA ESPECIAL - PRECATÓRIOS F	119.693,03	0,00	FORNECEDORES NACIONAIS	F	-128.351,68	0,00
OUTROS CRÉDITOS A RECEBER E VALORES A CURTO PRAZO	29.397,10	-13.818,70	PRECATÓRIOS DE FORNECEDORES NACIONAIS - REGIME ESPECIAL	F	119.693,03	0,00
CRÉDITOS A RECEBER POR REEMBOLSO DE SALÁRIO FAMÍLIA PAGO F	54.111,96	10.896,16	DEMAIS OBRIGAÇÕES A CURTO PRAZO		493.652,11	370.539,18
CRÉDITOS A RECEBER POR REEMBOLSO DE SALÁRIO MATERNIDADE F	-24.714,86	-24.714,86	VALORES RESTITUÍVEIS		493.652,11	370.539,18
ESTOQUES	2.518,00	0,00	CONSIGNACOES	F	489.581,43	371.353,60
ALMOXARIFADO	2.518,00	0,00	DEPOSITOS NAO JUDICIAIS	F	4.070,68	-814,42
MATERIAL DE CONSUMO P	338,00	0,00	TOTAL PASSIVO		297.226,05	370.539,18
GÊNEROS ALIMENTÍCIOS P	2.180,00	0,00	PATRIMÔNIO LÍQUIDO			
ATIVO NÃO CIRCULANTE	7.792.123,45	7.337.543,91	ESPECIFICAÇÃO		Exercício Atual	Exercício Anterior
IMOBILIZADO	7.792.123,45	7.337.543,91	PATRIMÔNIO LÍQUIDO		7.936.695,48	7.403.125,62
BENS MÓVEIS	2.176.124,41	1.751.091,87	PATRIMÔNIO SOCIAL E CAPITAL SOCIAL		7.026.355,95	7.026.355,95
BENS DE INFORMÁTICA P	7.320,00	7.320,00	PATRIMÔNIO SOCIAL		7.026.355,95	7.026.355,95
DEMAIS BENS MÓVEIS P	1.657.878,62	1.657.878,62	PATRIMÔNIO SOCIAL - CONSOLIDAÇÃO	P	7.026.355,95	7.026.355,95
VEÍCULOS P	173.440,00	0,00	RESULTADOS ACUMULADOS		910.339,53	376.769,67
MÓVEIS E UTENSÍLIOS P	52.045,00	0,00	SUPERÁVITS OU DÉFICITS ACUMULADOS		910.339,53	376.769,67
MÁQUINAS, APARELHOS, EQUIPAMENTOS E FERRAMENTAS P	285.440,79	85.893,25	SUPERÁVITS OU DÉFICITS DO EXERCÍCIO	P	533.569,86	376.769,67
BENS IMÓVEIS	5.615.999,04	5.586.452,04	SUPERAVITS OU DEFICITS DE EXERCICIOS ANTERIORES	P	376.769,67	0,00
BENS DOMINICAIS P	3.300,00	3.300,00	TOTAL PATRIMÔNIO LÍQUIDO		7.936.695,48	7.403.125,62
BENS IMÓVEIS EM ANDAMENTO P	61.500,00	61.500,00	TOTAL		8.233.921,53	7.773.664,80
BENS DE USO ESPECIAL P	189.045,83	189.045,83				
BENS DE USO COMUM DO POVO P	189.630,40	189.630,40				
INSTALAÇÕES P	4.472,00	1.172,00				
DEMAIS BENS IMÓVEIS P	5.168.050,81	5.141.803,81				
TOTAL	8.233.921,53	7.773.664,80				

ATIVO FINANCEIRO	439.280,08	436.120,89	PASSIVO FINANCEIRO (346.225,94)+ Restos não Processado(0,00)		346.225,94	370.539,18
ATIVO PERMANENTE	7.794.641,45	7.337.543,91	PASSIVO PERMANENTE		-48.999,89	0,00
			SALDO PATRIMONIAL		7.936.695,48	7.403.125,62

Raimundo Nonato de Alencar
Prefeito Municipal
CPF:178.968.275-49

Edivaldo da Silva Fontes
Contador
CRC-PI 4.497

Figure 1. Example of a balance sheet available to the public at Piauí's Court of Accounts (notice the signature over the data)

With the AWS Textract service, we converted the report files into tables in CSV and JSON formats, but there was still the need for data organization before usage. Santos (2021) provides the source code developed to organize and compare data. It is important to emphasize that this was the most challenging phase of the framework development since OCR's conversion often does not keep the hierarchy between the items. Consequently, data referenced by the same keyword can be confused by the extraction algorithm. Likewise, files scanned with low resolution or with stamps, or signature marks can also reduce the efficiency of the OCR algorithm.

To minimize these errors, tables were extracted before Textract submission. This approach, despite laborious, significantly minimized the subsequent work of verifying the consistency of the data converted.

RECEITAS ORÇAMENTÁRIAS	PREVISÃO INICIAL	PREVISÃO ATUALIZADA (a)	RECEITAS REALIZADAS (b)	SALDO c = (b-a)
RECEITAS CORRENTES	96.748.533,13	96.748.533,13	115.596.780,97	18.848.247,84
RECEITA TRIBUTARIA	5.450.000,00	5.450.000,00	5.187.453,86	-262.546,14
Impostos	4.100.000,00	4.100.000,00	4.745.841,33	645.841,33
Taxas	1.320.000,00	1.320.000,00	441.612,53	-878.387,47
Contribuicao de Melhoria	30.000,00	30.000,00	0,00	-30.000,00
RECEITA DE CONTRIBUICOES	6.253.000,00	6.253.000,00	7.856.003,71	1.603.003,71
Contribuicoes Sociais	5.053.000,00	5.053.000,00	5.209.993,29	156.993,29
Contribuição de Iluminação Pública	1.200.000,00	1.200.000,00	2.646.010,42	1.446.010,42
RECEITA PATRIMONIAL	2.031.500,00	2.031.500,00	1.360.654,90	-670.845,10
Receitas Imobiliárias	812.000,00	812.000,00	387.632,57	-424.367,43
Receitas de Valores Mobiliários	1.199.500,00	1.199.500,00	973.022,33	-226.477,67
Outras Receitas Patrimoniais	20.000,00	20.000,00	0,00	-20.000,00
RECEITA DE SERVIÇOS	8.303.500,00	8.303.500,00	7.658.392,01	-645.107,99
TRANSFERENCIAS CORRENTES	74.416.533,13	74.416.533,13	92.336.831,76	17.920.298,63
Transferências Intergovernamentais	72.294.533,13	72.294.533,13	92.115.470,46	19.820.937,33
Transferências de Convênios	2.122.000,00	2.122.000,00	221.361,30	-1.900.638,70
OUTRAS RECEITAS CORRENTES	294.000,00	294.000,00	1.197.444,73	903.444,73
Multas e Juros de Mora	176.000,00	176.000,00	725.417,64	549.417,64
Indenzações e Restituições	25.000,00	25.000,00	17.303,87	-7.696,13
Receita da Divida Ativa	45.000,00	45.000,00	308.869,86	263.869,86
Receitas Decorrentes Aportes Per. Amort. Déficit Atuarial do RPPS	1.000,00	1.000,00	0,00	-1.000,00
Receitas Correntes Diversas	47.000,00	47.000,00	145.853,36	98.853,36
RECEITAS DE CAPITAL	12.773.628,80	12.773.628,80	1.098.562,12	-11.675.8
OPERAÇÕES DE CRÉDITO	500.000,00	500.000,00	0,00	-500.000,00
Operações de Crédito Internas	280.000,00	280.000,00	0,00	-280.000,00
Operações de Crédito Externas	220.000,00	220.000,00	0,00	-220.000,00
ALIENAÇÃO DE BENS	240.000,00	240.000,00	0,00	-240.000,00
Alienação de Bens Móveis	120.000,00	120.000,00	0,00	-120.000,00
Alienação de Bens Imóveis	120.000,00	120.000,00	0,00	-120.000,00
TRANSFERENCIAS DE CAPITAL	11.780.628,80	11.780.628,80	1.098.562,12	-10.682.066,68
Transferências Intergovernamentais	2.176.000,00	2.176.000,00	250.210,00	-1.925.790,00
Transferências de Convênios	9.604.628,80	9.604.628,80	848.352,12	-8.756.276,68
OUTRAS RECEITAS DE CAPITAL	253.000,00	253.000,00	0,00	-253.000,00
Receitas de Capital Diversas	253.000,00	253.000,00	0,00	-253.000,00
SUBTOTAL DAS RECEITAS (I)	109.522.161,93	109.522.161,93	116.695.343,09	7.173.181,16
REFINANCIAMENTO (II)	0,00	0,00	0,00	0,00
Operações de Crédito Internas	0,00	0,00	0,00	0,00
Mobiliária	0,00	0,00	0,00	0,00
Contratual	0,00	0,00	0,00	0,00
Operações de Crédito Externas	0,00	0,00	0,00	0,00
Mobiliária	0,00	0,00	0,00	0,00
Contratual	0,00	0,00	0,00	0,00
SUBTOTAL COM REFINANCIAMENTO (III) = (I+II)	109.522.161,93	109.522.161,93	116.695.343,09	7.173.181,16
DÉFICIT (IV)			0,00	
TOTAL (V) = (III+IV)	109.522.161,93	109.522.161,93	116.695.343,09	
SALDOS DE EXERCÍCIOS ANTERIORES			0,00	0,00
(UTILIZADOS PARA CRÉDITO ADICIONAIS)			0,00	0,00
Superávit Financeiro			0,00	0,00
Reabertura de Créditos Adicionais			0,00	0,00

Figure 2. Example of a budgetary report available to the public at the Piauí's Court of Accounts (notice the archiving holes)

6. RESULTS

We developed a complete workflow using RPA in conjunction with OCR, capable of downloading, extracting, comparing, and identifying inconsistencies in data submitted by Brazilian municipalities to the Courts of Accounts and the Ministry of Finance. The results show that RPA can be an essential tool for inspecting governmental financial reports, confirming the first proposition. Figure 3 presents the complete flowchart developed in the study, and Santos (2021) presents the source code used to download, organize, and compare the municipalities' financial reports' data.
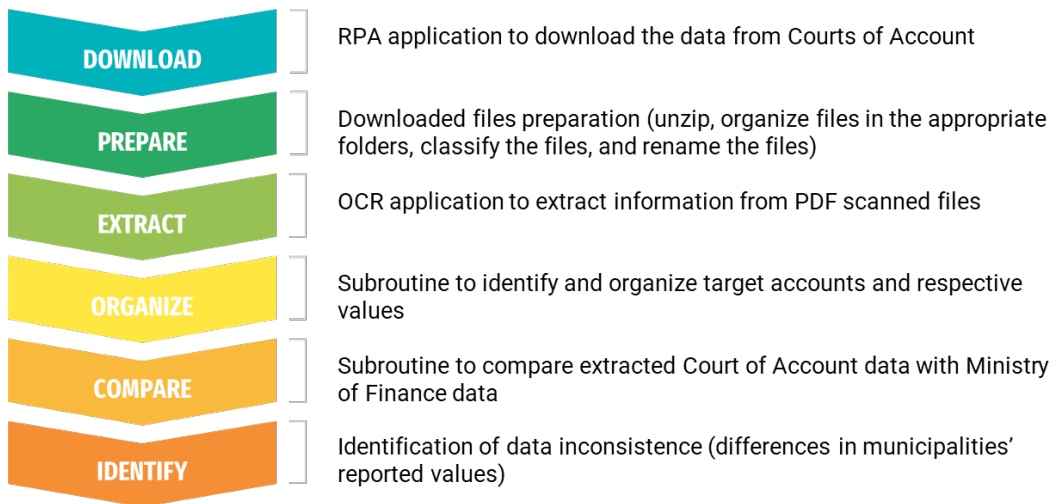
Figure 3. Workflow developed to automatize the consistency verification

Using our workflow, we were able to evaluate the consistency of municipalities' financial reports. Straightforward evidence of such consistency (or lack thereof) is the number of inexistent reports. Although city halls must submit financial reports to both databases (SICONFI, 2020; TCE – PI, 2020; and TCE – RS, 2020), there is a significant discrepancy in the data available when comparing the municipalities . In 2017, of the 224 municipalities in Piauí, only 213 had balance sheet data in both bases, 11 of which were not available at the Court of Accounts' website, and three of which were not available at the SICONFI database (Annual Accounts database). An even worse situation was verified for the budgetary reports of the Piauí municipalities; only 130 had data in both bases, seven municipalities' budgetary reports were not available at the Court of Accounts database, and 91 municipalities' budgetary reports were not available at the SICONFI database. In the case of Rio Grande do Sul, no data were missing from the analyzed databases in 2017.

The absence of data (financial reports that are not available at the analyzed databases) was also observed in 2018 for the municipalities of Piauí but again not for Rio Grande do Sul. Among the 224 municipalities in Piauí, only 199 presented the balance sheet at both bases, 25 of which were not available at the Piaui Court of Accounts database, and 11 were not available at the SICONFI (2020) database. In the budgetary reports, 200 municipalities were included in both databases, being that 19 were not available at the Piauí Court of Accounts database, and six were not available at the SICONFI database. Table 1 presents the total number of missing

financial reports in the Ministry of Finance and the Court of Accounts databases in 2017 and 2018 for municipalities from both states.

| | Missing reports in at least one of the databases | | | |
| | Piauí | | Rio Grande do Sul | |
	Balance sheet	Budgetary report	Balance sheet	Budgetary report
2017	11	94	0	0
2018	25	24	0	0
Total number of municipalities in the state	224		497	

Table 1. Total number of missing reports from the Rio Grande do Sul municipalities and Piauí municipalities at the Ministry of Finance and Courts of Accounts' databases 2017 and 2018.

Besides the absence of data, we could also automatically verify municipalities from Piauí that submitted to the Court of Account reports that were identical to those of other municipalities, suggesting municipalities were just 'copying & pasting' from each other. In 2017, for example, the reports from Santa Cruz dos Milagres were identical to those from Santa Cruz do Piauí, both available for public consultation at the Piauí's Court of Accounts website. We also found that the entire balance sheet from the year ended on 31 December 2018 from São José do Piauí was identical to that from Sussuapara in the Ministry of Finance website. In Rio Grande do Sul, we did not find any municipality with reports fully duplicated; however, some accounts were presented with identical balances. Table 2 presents a list of municipalities from Rio Grande do Sul with equal balance sheet balances, where the 'account code' follows the Ministry of Finance standardized chart of accounts (STN, 2021).

Year	Municipality	Account code	Balance (BRL)
2017	TRAVESSEIRO	2.2.0.0.0.00.00	597916.69
	SÃO VALENTIM	2.2.0.0.0.00.00	597916.69
2018	TUCUNDUVA	2.2.0.0.0.00.00	13490709.51
	VALE VERDE	2.2.0.0.0.00.00	13490709.51
	CORONEL PILAR	2.2.8.0.0.00.00	158205.88

BOA VISTA DO BURICÁ	2.2.8.0.0.00.00	158205.88
POÇO DAS ANTAS	2.2.8.0.0.00.00	158205.88
IVOTI	2.2.8.0.0.00.00	158205.88
BOM RETIRO DO SUL	2.2.8.0.0.00.00	158205.88
SEDE NOVA	2.1.2.0.0.00.00	175000.02
ALTO ALEGRE	2.1.2.0.0.00.00	175000.02
ESPUMOSO	2.1.2.0.0.00.00	249999.96
BOM JESUS	2.1.2.0.0.00.00	249999.96

Table 2. Balance sheet records with the very same balance for the same account from different municipalities from the Rio Grarnde do Sul

This framework was also able to automatically analyze a total of 28 accounts from the balance sheet (i.e., assets, liabilities, and equity accounts) and 18 accounts from the budgetary reports (i.e., budgetary revenues and expenditures). Although the data submitted by municipalities to both monitoring agencies should be equal, we identified a significant portion of municipalities with at least one divergent data in the financial reports submitted to SICONFI and their respective TCE. Table 3 presents the consolidated percentage of municipalities that submitted divergent data in at least one analyzed account in the financial reports, i.e., municipalities whose data in the Ministry of Finance database was divergent from their corresponding data in the Court of Accounts database. In Table 3, we calculated the percentage of the inconsistent reports in three ways: Panel A – excluding the missing data, i.e., the percentage was calculated by the number of municipalities with divergent data (considering just the municipalities which submitted the reports to both agencies) by the total number of municipalities which submitted the reports to both agencies; Panel B – excluding the missing data and excluding misclassification data, i.e., the percentage was calculated by the number of municipalities with divergent data in the consolidated accounts, e.g. current assets, non-current asset, etc. (considering just the municipalities which submitted the reports to both agencies) by the total number of municipalities which submitted the reports to both agencies; and in Panel C – including the missing report, i.e., the percentage was calculated by dividing the

total number of municipalities with divergent data, plus the total number of municipalities which did not submit the report to at least one of the controlling agencies, all divided by the total number of municipalities.

Panel A: Cases of missing reports excluded from the analysis

		% of Municipalities with inconsistent data		
Year	Report	Piauí	Rio Grande do Sul	Total
	Balance Sheet	40%	31%	34%
2017	Budgetary Report, income accounts	27%	22%	24%
	Budgetary Report, expenditure accounts	27%	15%	19%
2018	Balance Sheet	34%	13%	20%
	Budgetary Report, income accounts	25%	23%	24%
	Budgetary Report, expenditure accounts	26%	12%	16%

Panel B: Cases of missing reports and cases of data misclassification excluded from the analysis

		% of Municipalities with inconsistent data		
Year	Report	Piauí	Rio Grande do Sul	Total
	Balance Sheet	26%	17%	20%
2017	Budgetary Report, income accounts	25%	17%	19%
	Budgetary Report, expenditure accounts	25%	14%	17%

Year	Report			
	Balance Sheet	30%	13%	18%
2018	Budgetary Report, income accounts	22%	16%	18%
	Budgetary Report, expenditure accounts	25%	10%	15%

Panel C: Cases of missing reports included in the analysis

		% of Municipalities with inconsistent data		
Year	Report	Piauí	Rio Grande do Sul	Total
	Balance Sheet	43%	31%	35%
2017	Budgetary Report, income accounts	58%	22%	33%
	Budgetary Report, expenditure accounts	58%	15%	28%
	Balance Sheet	41%	13%	22%
2018	Budgetary Report, income accounts	33%	23%	26%
	Budgetary Report, expenditure accounts	34%	12%	19%

Table 3. Percentage of municipalities which presented divergent data in at least one analyzed account in the financial reports.

Adding to the inconsistencies observed from both Rio Grande do Sul and Piauí, 421 reports had some variances in 2017. This number dropped to 349 in 2018, indicating an improvement in data quality. Indeed, 292 municipalities did not present inconsistencies over the two years analyzed. However, 336 municipalities presented at least one inconsistent report in 2017, and 175 municipalities did so in 2018, indicating an improvement in data quality from 2017 to 2018.

The results from Table 3 also indicate that the Rio Grande do Sul municipalities are more prone to submit data with fewer inconsistencies to the monitoring agencies than the municipalities from Piauí. The Court of Accounts from the Rio Grande do Sul is more computerized than Piauí (Silva, 2019) consistent with proposition 2.

Table 4 shows the percentage of municipalities that presented divergent reports in 2017 and 2018, considering the total number of municipalities that submitted at least one inconsistent report. The data indicates that most municipalities that submit inconsistent data in 2017 repeated problems in 2018.

	% of municipalities that presented inconsistent data in both years, by municipalities that presented inconsistency in at least one report	
	Piauí	Rio Grande do Sul
Balance Sheet	53%	40%
Budgetary report	69%	65%

Table 4. Percentage of municipalities which presented divergent reports both years (2017 and 2018), considering the total number of municipalities that submitted at least one divergent report.

7. CONCLUSIONS

This study developed a workflow that automatically assesses the consistency of the financial reports made available by the city halls from Piauí (224 municipalities) and the Rio Grande do Sul (497 municipalities) to two public bodies responsible for their external control (monitoring agencies), i.e., Courts of Accounts and the Ministry of Finance. Following transparency and open government data policy, the Ministry of Finance and the Courts of Accounts make the financial reports prepared by the municipalities publicly available at their internet portals. Thus, any citizen interested in knowing any Brazilian municipality's financial and budgetary performance can get reports from either portal (SICONFI, 2020; TCE - PI, 2020; TCE - RS, 2020). It is expected that the data disclosed at different databases to be the same, regardless of the source. However, it appears that the values of the same

accounts and the same reports presented to different public bodies are often divergent, as shown in Table 3.

Although the data is available online and can be accessed by any citizen interested in inspecting governmental accounts, reports are often made available in scanned images converted to PDF (it is a reality among municipalities from Piauí). There were barriers to downloading multiple reports from the Court of Accounts. Thus, it was necessary to prepare an RPA application to automatize this task. There was no standard format for the presented reports. Therefore, considering these issues the transparency of municipalities' accounts (enforced by the law) is not a full reality.

Examining Proposition P1, we developed a workflow combining RPA and OCR to download, extract, and organize the financial data submitted by municipalities from Rio Grande do Sul and Piauí. We also automatically compared the data presented by municipalities to two controlling agencies (the Ministry of Finance and the respective Court of Accounts) to assess data inconsistencies.

The results indicate that 60% of municipalities, on average, submit inconsistent data. The results also show over time a significant decrease in inconsistencies (there are significantly fewer inconsistencies in 2018 than in 2017), and that the municipalities from the Rio Grande do Sul (which Court of Account has implemented a higher level of computation) presented significantly fewer inconsistencies than from Piauí, as suggested by Proposition 2.

The prevalence of inconsistencies within numerous financial reports highlights a concerning lack of accountability within city administrations. While the law mandates transparency and reliability in these reports, many municipalities appear to disregard their legal obligations without facing any repercussions. When city mayors present inconsistent financial data to different oversight bodies such as the Court of Accounts and the Ministry of Finance, it may seem as though they are meeting the requirements of both authorities. However, upon closer examination of these inconsistencies, it becomes evident that the mayor is not adhering to neither transparency requirements. This discrepancy is distinct from "*de jure* compliance," which refers to the extent to which local officials follow federal law by implementing legal measures to guide the application of national laws at the local level (Michener & Nichter, 2022).

These results can serve armchair auditors interested in using public sector financial reports and to motivate the public sector to invest in computerization and facilitate access to monitoring the application of public resources. Although there are user-friendly RPA packages (such as Uipath, Blue Prism, and Automation Anywhere), we opted to develop the codes in Python because it is opensource and counts with many libraries that are also consumable at no cost. Hence, those intended to armchair audit the consistency of public sector accounts can use the codes we provide on GitHub (Santos, 2021). Indeed, researchers in general that need to consume documents only available on image PDF can also use Santos's (2021) codes. Hence, this research contributes on the design science literature by developing a whole framework to automate all the tasks involved in analyzing the reports submitted by municipalities from two different Brazilian states, which can be used by researchers and auditors interested in financial data from Brazilian municipalities or adapted by those interested in analyzing similar data from other governmental entities, including from other jurisdictions.

This research, focused on the comparison of the data submitted by municipalities to two control bodies (State Court of Accounts and Ministry of Finances), it is suggested to extend the comparison of data to other public bodies that also require the provision of financial reports by municipalities, e.g., Ministry of Health and Ministry of Education. Likewise, the same evaluations presented here can be carried out for municipalities in other states, as regional differences prevent the generalization of the conclusions found for other contexts. In the same way, studying other countries realities could also be interesting to compare how culture and level of bureaucracy can impact the reliability of data made available by governmental entities to the general public.

A significant limitation of this manuscript lies in its primary focus on the identification of inconsistencies rather than delving into their underlying causes. An extension of this research could pivot towards a more in-depth exploration of the factors contributing to these discrepancies, with particular emphasis on investigating whether these inconsistencies may be indicative of fraudulent activities or not. A minor limitation refers to the sample representativeness; we only analyzed financial data from all municipalities of two Brazilian states, future research could investigate the inconsistencies on non-financial data disclosed by other governmental entities.

8. REFERENCES

AWS - Textract. (2020). GITHUB AWS TEXTRACTOR. https://github.com/aws-samples/amazon-textract-textractor. Accessed 13 December 2020

Briloff, A. J. (1972). *Unaccountable Accounting: Games Accountants Play*. HarperCollins. ISBN 13: 9780060104719

Cesario, G. (2020). *The sousveillance of politically connected entities and individuals in the public sector: an armchair audit approach of related party transactions*. Rio de Janeiro: Fundação Getúlio Vargas. https://hdl.handle.net/10438/29659. Accessed 12 April 2021

Dai, J., & Li, Q. (2016). Designing audit apps for armchair auditors to analyze government procurement contracts. *Journal of Emerging Technologies in Accounting*, 13 (2), 71-88. https://doi.org/10.2308/jeta-51598

Geerts, G. L. (2011). A design science research methodology and its application to accounting information systems research. *International Journal of Accounting*, 12 (2), 142-151. https://doi.org/10.1016/j.accinf.2011.02.004

Gotthardt, M., Koivulaakso, D., Paksoy, O., Saramo, C., Martikainen, M., & Lehner, O. (2020). Current state and challenges in the implementation of smart robotic process automation in accounting and auditing. *ACRN Journal of Finance and Risk Perspectives*, 9 (1), 90-102. https://doi.org/10.35944/jofrp.2020.9.1.007

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. *MIS quarterly*, 28 (1), 75-105. https://doi.org/10.2307/25148625

Huang, F., & Vasarhelyi, M. A. (2019). Applying robotic process automation (RPA) in auditing: A framework. *International Journal of Accounting Information Systems*, 35, 100433. https://doi.org/10.1016/j.accinf.2019.100433

Huang, Z., Qian, L., Sun, S., Cai, D., Hang, T., & Feng, Q. (2020). Design of Dynamic Cooperative Robot System for Financial Bill Recognition. *4th International Conference on Automation, Control and Robots (ICACR)*, 44-47. https://doi.org/10.1109/ICACR51161.2020.9265492

IBGE. (2020). Estatística População IBGE 2019. https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html?edicao=25272&t=resultados. Accessed 13 December 2020

Iguma, M. K., & Riccio, E. L. (2018). The use of Information Technology tools in auditing: a Bibliometric Analysis. *15th CONTECSI International Conference on Information Systems and Technology Management*. http://dx.doi.org/10.5748/9788599693148-15CONTECSI/PS-5734

Janvrin, D., Bierstaker, J., & Lowe, D. J. (2008). An examination of audit information technology use and perceived importance. *Accounting Horizons*, 22, 1-21. http://dx.doi.org/10.2308/acch.2008.22.1.1

Joffe, M., & Reck, J. (2019). Applying XBRL to US State and Local Government Audited Financial Reports. *Mercatus Research Paper*, 1-20. http://dx.doi.org/10.2139/ssrn.3311695

Koningisor, C. (2020). Transparency deserts. *Northwestern University Law Review*, 114 (6), 1461–1548. https://heinonline.org/HOL/P?h=hein.journals/illlr114 &i=1497 Accessed 12 April 2021

Lacurezeanu, R., Tiron-Tudor, A., & Bresfelean, V. P. (2020). Robotic Process Automation in Audit and Accounting. *Audit Financiar*, 160, 752-770. http://dx.doi.org/10.20869/AUDITF/2020/160/024

Michener, G., & Nichter, S. (2022). Local compliance with national transparency legislation. *Government Information Quarterly*, 39 (1), 101659. https://doi.org/10.1016/j.giq.2021.101659.

Moffitt, K. C., Rozario, A. M., & Vasarhelyi. (2018). Robotic process automation for auditing. *Journal of Emerging Technologies in Accounting*, 15 (1), 1-10. https://doi.org/10.2308/jeta-10589

O'Leary, D. E. (2015). Armchair auditors: Crowdsourcing analysis of government expenditures. *Journal of Emerging Technologies*, 12 (1), 71-91. https://doi.org/10.2308/jeta-51225

Orwell, G. (2009). *1984*. Companhia das Letras. ISBN-13: 978-8535914849

Pernagallo, G., & Torrisi, B. (2020). A logit model to assess the transparency of Italian public administration websites. *Government Information Quarterly*, 37 (4), 101519. https://doi.org/10.1016/j.giq.2020.101519

Pfeffers, K., Tuunanen, T., Gengler, C. E., Rossi, M., Hui, W., Virtanen, V., & Bragge, J. (2006). The design science research process: A model for producing and presenting information systems research. *Proceedings of the First International*. Claremont (CA) - USA. https://doi.org/10.48550/arXiv.2006.02763

Rezaee, Z., Elam, R., & Sharbatooghlie, A. (2001). Continuous auditing: the audit of the future. *Managerial Auditing Journal*, 16 (3), 150-158. https://doi.org/10.1108/02686900110385605

Rozario, A. M., & Vasarhelyi, M. A. (2018a). Auditing with Smart Contracts. *International Journal of Digital Accounting Research*, 18, 1-27. http://dx.doi.org/10.4192/1577-8517-v18_1

Rozario, A. M., & Vasarhelyi, M. A. (2018b). How robotic process automation is transforming accounting and auditing. *The CPA Journal*, June, 46-49.

https://www.cpajournal.com/2018/07/02/how-robotic-process-automation-is-transforming-accounting-and-auditing/ Accessed 12 April 2021

Salles, E. B. (2019). *O Timeliness dos Relatórios resumidos de Execução Orçamentária e os Gastos com Tecnologia da Informação no Setor Público Brasileiro*. FUCAPE Fundação de Pesquisa e Ensino, Vitória. https://anpcont.org.br/pdf/2020_CSP147.pdf Accessed 12 April 2021

Santos, L. S. (2021). https://github.com/luizschmall/tce_siconfi_inconsistencies (v2.1). Zenodo. https://doi.org/10.5281/zenodo.4557655.

Schaller, R. R. (1997). Moore's law: past, present and future. *IEEE spectrum*, 34 (6), 52-59. https://doi.org/10.1109/6.591665

SICONFI (2020). Secretaria do Tesouro Nacional: https://siconfi.tesouro.gov.br/siconfi/index.jsf;jsessionid=RJ+q0EI48fS5wCXlJE2plNYV.node3 Accessed 13 December 2020

Silva, Z. M. (2019). *Transparência das Informações Contábeis dos Municípios pelos Tribunais de Contas*. Universidade Federal de Uberlândia - UFU, Uberlândia. https://doi.org/10.51891/rease.v8i10.7357

STN, Secretaria do Tesouro Nacional. (2021). *Instruções de Procedimentos Contábeis*.https://siswebTesouro.gov.br/apex/f?p=2501:9:::::9:P9_ID_PUBLICACAO_ANEXO:8726. Accessed 22 June 2021

TCE - PI. (2020). Tribunal de Contas do Estado do Piaui. https://www.tce.pi.gov.br/fiscalizado/pesquisa-de-processos/ Accessed 13 December 2020

TCE - RS. (2020). Tribunal de Contas do Estado do Rio Grande do Sul. http://dados.tce.rs.gov.br/dados/municipal/recebimentos/2018.html Accessed 13 December 2020

Thuan, N. H., Drechsler, A., & Antunes, P. (2019). Construction of Design Science Research Questions. *Comunications of the Association for Information Systems*, 44, 332-363. https://doi.org/10.17705/1CAIS.04420

Vasarhelyi, M. A., & Halper, F. B. (1991). The Continuous Audit of Online Systems. *Auditing: A Journal of Practice and Theory*, 10 (1), 110-125. https://doi.org/10.1108/978-1-78743-413-420181004

Walls, J. G., Widmeyer, G. R., & Sawy, O. A. (1992). Building an information system design theory for vigilant EIS. *Information Systems Research*, 3 (1) 36-59. https://doi.org/10.1287/isre.3.1.36