A Review of How Whistleblowing is Studied in Software Engineering, and the Implications for Research and Practice

Lucy Hunt
L.Hunt1@lancaster.ac.uk
Lancaster University
Lancaster, England, UK

Maria Angela Ferrario
M.Ferrario@qub.ac.uk
Queen's University Belfast
Belfast, Northern Ireland, UK

ABSTRACT

Harmful software has resulted in loss of life, societal and environmental damage alongside economic losses from fines and sales embargoes. When someone perceives their team or organisation is creating or operating harmful software (e.g., defective, vulnerable, malicious or illegal), one way to attempt to change the situation is to "blow the whistle" and disclose the situation internally or externally. Studying harmful situations and the effectiveness of interventions, up to and including whistleblowing, can help identify technical and human successes and failings in software engineering (SE).

The aim of this paper is to explore the extent to which whistleblowing is studied in SE with the objective of identifying themes, research approaches, gaps and concerns, and the implications for future SE research and practice.

We find that whistleblowing is an under-explored area of SE research, and where research exists, it often takes the view that reporting harm is a matter of individual moral responsibility; we argue this poorly reflects SE collaborative practice where professional responsibilities are distributed across the software development lifecycle. We conclude by 1) recommending approaches that can help a more timely identification and mitigation of harm in SE; 2) suggesting mechanisms for improving the effectiveness and the personal safety of harm-reporting in SE, and 3) reflecting on the role that professional bodies can have in supporting harm reporting, up to and including whistleblowing.

CCS CONCEPTS

• **Software and its engineering** → *Software creation and management*; • **Social and professional topics** → **Codes of ethics**;

KEYWORDS

Software Engineering, Whistleblowing

ACM Reference Format:
Lucy Hunt and Maria Angela Ferrario. 2022. A Review of How Whistleblowing is Studied in Software Engineering, and the Implications for Research and Practice. In *Software Engineering in Society (ICSE-SEIS'22), May 21–29, 2022, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3510458.3513013

ICSE-SEIS'22, May 21–29, 2022, Pittsburgh, PA, USA

ACM ISBN 978-1-4503-9227-3/22/05.
https://doi.org/10.1145/3510458.3513013

1 INTRODUCTION

Whistleblowing has a long history spanning at least three centuries [118]. However, the act of "blowing the whistle" seems to be a relatively rare occurrence in the science and technology community, particularly when compared to the number of reported whistleblowing cases amongst health, education, government and finance professionals [110]. Many definitions of whistleblowing exist with varying levels of detail regarding roles, responsibilities, obligations and evidence. Jubb, writing in a business ethics journal, [43] describes it as "a deliberate non-obligatory act of disclosure, which gets onto public record and is made by a person who has or had privileged access to data or information of an organisation, [...] to an external entity having potential to rectify the wrongdoing." A broader view, by the United Nations, describes disclosures through internal mechanisms, external oversight mechanisms or in some cases to the media [46]. We construct a definition specifically for SE practice anchored around the software development lifecycle and SE literature: "the responsible evidencing and disclosure of actions and artefacts perceived to be contrary to accepted ethical and professional standards [1] or software life-cycle standards (e.g. [27]), carried out to mitigate harm to self, others or wider society."

We acknowledge that whistleblowing is not the only way to address harmful situations, and recognise the extensive work in both the research and practitioner community to detect, reduce and prevent harm (e.g. ethical SE [37] and responsible AI [6]). Yet despite the best efforts of individuals, teams and organisations there are circumstances where blowing the whistle, on some level, may be the final option to raise concerns and change outcomes. Studying harmful SE situations and effective interventions, up to and including whistleblowing, is evidence of the success or failure of harm mitigation approaches in SE research and practice.

Goal and aim- The ultimate goal of our research is to understand how SE practitioners and organisations make disclosures about perceived or actual harmful software or SE practices, with a particular focus on those who find it "necessary to blow the whistle" [1]. We are motivated to explore challenging situations as opportunities for reflection and improvement of SE practices, and to support the SE community to uphold professional values and standards. Working towards this goal, the specific aim of this paper is to bring to light how whistleblowing has been studied in the SE research community. With no literature reviews on the phenomenon within SE identified, we explore it through the following research questions:

- *RQ1: What are the prevalent whistleblowing themes in SE?*
- *RQ2: How is whistleblowing researched in SE?*
- *RQ3: What are the gaps and concerns for SE research and practice?*

Contribution- The main contribution of this paper is a critical analysis of how the phenomenon of whistleblowing is studied in SE literature. Not only do we find that the effectiveness of whistleblowing is an under-explored area of SE research, but where SE research exists, it takes the view that reporting wrongdoing is a matter of *individual* morals and responsibility. We find that this view is not only prevalent in research studies, but is also a concern in SE education and practice. We argue that this view may be unhelpful and particularly problematic for SE, since software system design and development is rarely an individual effort, but a collaborative practice, distributed across different teams and complex lines of responsibility.

Paper structure- The rest of the paper is organised as follows: in Section 2 we introduce some key perspectives on whistleblowing from organisational studies, decision sciences, and from a broad legal and policy overview. In Sections 3 we summarise our method, results and findings from the literature review. Section 4 discusses how our research questions are answered followed by the implications for research and practice. We conclude with validity and future work reflections in Sections 5 and 6.

2 BACKGROUND

The early 1970s saw whistleblowing at the centre of heated debates in the engineering community, triggered by Nader's call to scientists and engineers to "hold responsibility" [55]. In the 1980s high profile cases such as the Challenger disaster re-ignited the discussion in the community. Zelby's 1989 feature article written for the *IEEE Technology and Society Magazine* [118] concedes that "the issue of blowing the whistle is complicated", but also calls for something to be done beyond codes of ethics, standards, and legislation; the article concludes by arguing for the establishment of an ombudsman, seen as a possible mechanism to "protect both the whistleblower and those on whom the whistle was blown". More recently, high profile whistleblower stories involving computing professionals have attracted public attention; these include Christopher Wylie's revelations on Cambridge Analytica's unlawful use of personal data [31], and Snowden's leaked documents on NSA mass-surveillance [106]. By contrast, at Volkswagen no one publicly raised concerns about the software designed to mislead emission tests, until external emission test reports identified the issues in 2015 [65].

The ACM Code of Ethics (Principle 1.2 Avoid Harm) [1] states that "a computing professional has the obligation to report any signs of systems risks that may result in harm" and that "if leaders do not act [...] it may be necessary to blow the whistle". It then advises to "avoid misguided reporting" and to "carefully assess relevant aspects of the situation". A perceived lack of support, compounded with the risk of serious personal consequences from whistleblowing, may explain why computing professionals chose to leave a workplace rather than or as a result of reporting up issues.

In 2019, an IPSOS Mori online survey [67] of over 1000 UK computing professionals found 28% of respondents had witnessed decisions about a technology that could have negative consequences for people or society. Their concerns included the development of products causing harms such as addictiveness, job losses or isolation, alongside failures in safety, security and testing practices before product releases. As detailed in Table 1, the survey showed 90% of participants took some action(s) about the issue, ranging from talking to a colleague up to reporting to an external body. However, 18% reported leaving their company and 10% took no action. Statista, using data from the UK's Office for National Statistics, reports that there were over 408,000 "programmers and software development professionals" in the UK in 2020 [100]. The potential scale of people leaving a job is worrying, more so if the number of computing professionals worldwide is also considered. The impact of staff turnover and the resources to replace and train new staff should not be under estimated by individuals or their organisations [75], nor indeed the wider impact that unreported harmful software can have on society. Harmful software has resulted in loss of life, societal and environmental damage alongside economic losses from fines and sales embargoes (Volkswagen [65], Boeing [42] and Facebook [31]).

Table 1: Actions taken by tech workers after witnessing decisions about technology with potentially negative consequences (N=287, from survey of N=1010). An IPSOS Mori survey commissioned by Miller and Coldicott [67].

Action(s) taken by tech workers	%
Took no action	10
Left the company	18
Considered leaving the company	28
Reported concerns to external body	29
Raised concerns with manager or HR	47
Raised concerns with a colleague	51

2.1 Research Context

Whistleblowing research primarily sits at the intersection of decision sciences [107], organisational studies [68], and business ethics [66]. With studies spanning over 30 years, Dozier, Miceli and Near have developed widely cited works on whistleblowing effectiveness and decision making models [38, 68, 69]. Alongside this, moral psychology and business ethics scholars such as Alford [4] have focused on the lives and experiences of whistleblowers.

In 2019, a social psychology paper by Anvari et al. [5] contended that much whistleblowing research has focused on "the impact of individual and organisational factors" and proposes models to widen the scope of studies to include wider social identities and group memberships. Similarly to Anvari et al., we also argue that viewing whistleblowing as an individual act of responsibility confronting wrongdoing is problematic, particularly within the SE context.

In Figure 1 we show Anvari et al.'s whistleblowing escalation boundary hierarchy, adapted for the SE domain. SE practice is based on complex interactions between individuals, teams, organisations, and wider society with responsibilities distributed across the software development lifecycle [16]. In proposing a model for the behavioural and social aspects of a software engineer's work, Lenberg et al. [57] describe a "unit of analysis" that includes individuals, groups and organisations.

Taking on board Anvari et al. and Lenberg et al.'s work, we argue that in SE, reducing any studies to individual ethics, morals and actions will fail to fully understand how and why whistleblowing

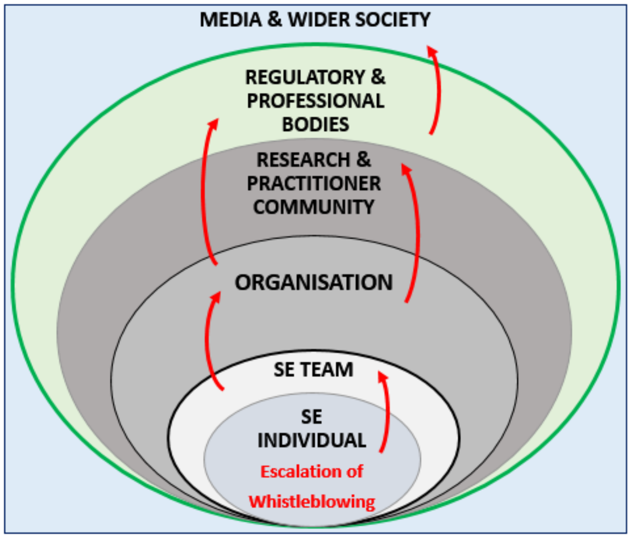
MEDIA & WIDER SOCIETY
REGULATORY & PROFESSIONAL BODIES
RESEARCH & PRACTITIONER COMMUNITY
ORGANISATION
SE TEAM
SE INDIVIDUAL
Escalation of Whistleblowing

Figure 1: Proposed SE whistleblowing escalation boundary hierarchy based on Anvari et al. [5]

happens (or indeed does not happen). Nor will it factor in advice and support needed by computing professionals and organisations when choosing to take or respond to whistleblowing actions.

2.2 Legal Status and Protection

Whistleblower protection varies greatly across countries, employment, and contract types. A government employee may be bound by state secret acts, a freelancer by non-disclosure agreements, and in some sectors such as healthcare there are disclosure channels provided [14]. In the UK, Vandekerckhove's 2012 survey of 2000 British adults found 81% of people supportive of whistleblowers [109]. Despite this, and with laws supporting whistleblowers, case studies overwhelming report that those reporting harm or wrongdoing are victimised and ostracised in the workplace by both colleagues and management [4], even when their disclosures are vindicated after lengthy investigations or court cases [25].

In Europe, prior to 2019, few EU Member States fully protect whistleblowers by law. To address this fragmented legislative landscape, new EU whistleblower laws were adopted in 2019, requiring protection by both companies and governments, including safe channels for reporting internally and externally [23, 24]. In the UK, protection for whistleblowers is included in the Public Interest Disclosure Act 1998 [15]. If a worker makes a protected disclosure, compensation can be claimed for any victimisation following such disclosures. The 1989 Whistleblower Protection Act is the USA's equivalent, with the False Claims Act offering financial incentives for disclosing fraud committed against the U.S. government.

The SE industry is a diverse community including consultants, contractors, third party, and outsourced organisations that may fall outside whistleblower protection policies or laws in a particular organisation or country. While whistleblowing may be assumed beneficial for society, encouraging it may be inappropriate without understanding factors that increase the likelihood of its effectiveness to change a situation [69]. When serious software issues are discovered by SE teams there needs to be a clear understanding of channels and actions available, which in extreme cases could include whistleblowing.

Research and advocacy practices [109, 110] stress the importance for organisations to understand their obligations upon the discovery and disclosure of harmful practices, as they have a responsibility to explain and uphold their policies and processes for reporting and managing the issues. They also highlight the need for individuals to become aware of the support and protection available to them should they choose to take action. Kenny et al.'s recent organisational research [51] finds that, even after reprisals, whistleblowers continue to be passionate about their professional integrity, and stress the importance of sustained practical and material support for those who blow the whistle. We reflect on this and ask how should technology organisations, trade unions and professional bodies such as the IEEE, IET and ACM support the SE community in such situations?

3 METHOD

Our method is guided by Kitchenham's systematic literature review procedures, an approach designed for established research fields [52]. While whistleblowing is not new, it is not a mature SE research field. Some of Kitchenham's SLR specifics could not be applied in full. For instance, we need to prioritise coverage (the review of any articles or studies about whistleblowing in SE peer reviewed literature) over quality assessment. Similarly, we do not use publication dates or citation counts as cut off points for inclusion.

Since we are looking for items *specifically* about whistleblowing in SE, our primary search was based on the title, abstract and key words of items, following approaches used in similar SE mapping studies by Glass et al. [30] and Shaw [95]. The items' selection was organised in three broad cycles: 1) find any whistleblowing items in the chosen information sources; 2) remove false positives where whistleblowing is used in its literal meaning in sport, football, sailing, acoustics or video editing for example; 3) apply SE specific selection criteria to the remaining items for their relevance to SE practitioners, projects or practices. The specific steps used to prepare and execute our primary search, and carry out items' selection are described below:

(1) Identify the need for a literature review
(2) Formulate research questions
(3) Define search strategy and string (whistleblowing)
(4) Search relevant information sources
(5) Remove duplicates (SCOPUS contains ACM and IEEE items)
(6) Remove false positives
(7) Apply specific inclusion criteria (software engineering)
(8) Apply general inclusion criteria (date, publication year, type)

Selected items were then inductively coded [18] (grounded in data, not existing concepts) to support the analysis and interpretation of the results:

(1) Extract meta data from selected items (title, abstract, author, publication, citations, source, country(s), keywords, document type)

(2) Authors independently review and inductively code themes for subset of items
(3) Authors compare and discuss themes
(4) Authors agree and define set of primary themes
(5) Authors independently allocate primary theme to all items
(6) Use Cohen's Kappa (K) [17] to calculate primary theme inter-rater reliability score
(7) Authors discuss and resolve discrepancies
(8) Authors run second round of primary theme allocation based on defined themes
(9) Finalise results for analysis and interpretation

3.1 Search and Selection Strategy

3.1.1 Information Sources. We performed automated searches on IEEE Xplore Digital Library, ACM Digital Library and the SCOPUS (computer science) database. The IEEE is the largest professional organisation for technology and the ACM covers computing and information technology; any items in here could be related to software or software professionals. SCOPUS is a broader database covering many disciplines, so the search was limited to computer science (COMP-SCI) items only.

3.1.2 Whistleblowing Search. The word whistleblow defines precisely our area of interest, however it can be phrased in multiple ways. The following 12 whistleblowing variations were used to search the information sources:

"whistle blow" OR "whistleblow" OR "whistle-blow" OR "whistle blowing" OR "whistleblowing" OR "whistle-blowing" OR "whistle blower" OR "whistleblower" OR "whistle-blower" OR "blowing the whistle" OR "blown the whistle" OR "whistle blown".

3.1.3 Removing False Positives. The titles and abstracts from the whistleblowing search were manually reviewed to remove false positives, for instance where whistleblowing is used in its literal sense (e.g. sport, football, sailing, acoustics or video editing).

3.1.4 Removing Duplicates. Duplicate entries, where SCOPUS returned matching items as the IEEE or ACM, were identified. SCOPUS and IEEE functionality provided the richest export data, therefore ACM duplicates were the most frequently removed.

3.1.5 Selecting Software Engineering Items. The abstract and body of whistleblowing items were manually reviewed for terms or narratives relevant to SE:

- Items about SE projects (creating products or services)
- Items about studies or reports on SE practitioners

3.1.6 Inclusion Criteria. Finally, the following inclusion and exclusion criteria were applied to the items:

Accepted

- Any publication year
- Peer reviewed book chapters and magazine items
- Peer reviewed conference and journal papers (any length)

Rejected

- Papers not in English
- Item behind paywall and not locatable in library
- Entire books, book reviews, or volumes of proceedings

Volumes of proceedings, such as [108], were rejected as relevant items were returned as part of the main search. In this example [87] was returned, although subsequently rejected as not specifically a software engineering item.

3.2 Results

A full text and metadata search for whistleblowing (or its variation) was run against each source database. When the metadata for these was analysed only 311 of the 1311 items contained whistleblowing in the title, abstract or author keyword fields, as detailed in Table 2. This review selected items *specifically* about whistleblowing such that it is *explicitly* mentioned in the title, abstract or keyword fields. Selecting and classifying papers based on their title, abstract and keywords is an approach used in similar SE studies [30, 95]. The full text data result set will be revisited in future research, to examine the frequency and context of the use of whistleblowing terms [84].

Table 2: Preliminary Whistleblowing Search Results

Whistleblowing Search	IEEE	ACM	SCOPUS	Total
Full Text and Metadata	422	264	625	1311
Title, Abstract, Keyword	70	94	147	**311**

3.2.1 First round selection. Metadata from the 311 preliminary items was downloaded into a spreadsheet. An initial search identified 124 false positive or duplicate items, as shown in Table 3, giving 187 first round selection items.

Table 3: First Round Whistleblowing Selection Results

Whistleblowing Search	IEEE	ACM	SCOPUS	Total
Title, Abstract, Keyword	70	94	147	311
False Positive, Duplicate	13	72	35	124
First round selection	57	18	112	**187**

3.2.2 Second round SE relevancy. The title and abstracts of the 187 items from the first round selection were then analysed for the words "software" or "engineer", 34 unique items were found. A search for "project" (things that software engineers work on) in the 187 items also run and returned 20 items and gave us a further 11 unique items to add to the existing 34 items. Finally, a manual review of the 187 items abstracts and bodies was made to identify SE project and product related items not detected by the previous systematic key term search as described in Section 3.1.5. Examples of words found included "IT", "computing industry", "professional", "safety-critical system" and "technology". This gave us a further 29 unique items for inclusion. All remaining search items were, where available, downloaded and checked for relevancy based on the inclusion criteria in Section 3.1.6, 14 were excluded. This gave us a final list of 60 items for analysis, as shown in Table 4.

A thematic analysis followed to inductively code [18] items:

(1) 20 items independently reviewed and coded for themes
(2) Primary and sub-themes compared and discussed
(3) Key theme list agreed and defined (6 primary)
(4) All items independently allocated primary theme

Table 4: Second Round SE Relevancy and Inclusion Results

Second Round Selection	**Count**
Items with "software" or "engineer"	+34
Unique items with "project"	+11
Manual check for relevance to SE	+ 29
Exclusion criteria applied	-14
Final Accepted Items	**60**

(5) Primary inter-rater reliability score calculated (K = 0.82)
(6) Authors discussed and resolved 8 discrepancies
(7) Second round of primary theme allocation run
(8) Disagreed (weakly) on only 2 items
(9) Result set finalised for interpretation and discussions

For a subset of 20 items, each author independently tagged themes and sub-themes (unlimited number of themes to each item). The resulting list of themes included: whistleblowing practices, individual attitude and behaviour, professional ethics, technology for whistleblowing, not whistleblowing (not speaking up), management, governance, organisations, case study, stories, human decision making, industry domains (education, government, engineering, health), SE aspects, security and privacy. The authors compared and discussed their results for each of the 20 items to understand each other's applied theme. Of note was the separating out of "human factors" into "personal and social factors" and "organisational and professional issues" - thus identifying differences between individual and organisational themes. From this we jointly came up with six primary coding themes, with a list of sub-themes for that primary theme. These are shown in Table 5.

All 60 items were then independently allocated one of the primary themes by each author. After the calibration cycle and 2 primary theme iterations both authors agreed strongly (Kappa = 0.96) on a primary theme for each item. As an example of a disagreement, in [77] (Title: "Organisational Factors and Bad News Reporting on Troubled IT Projects") both agreed on the paper having themes of organisation and whistleblowing process, but disagreed on which was the primary one. The final 6 primary themes, sub-themes and papers are detailed in Table 5 and are discussed in Section 4.

3.3 Meta Data Results Summary

Based on item meta data, this section gives a summary of citations, temporal and geographical distribution of reviewed items, alongside an overview of the types and publication sources of items. These results feed into the discussions and findings in Section 4 of the paper.

3.3.1 Temporal View and Types. Figure 2 presents a publication timeline of the search results. It shows that generally there has been an increase in publications *mentioning* whistleblowing, peaking at 18 items in both 2017 and 2018. However, the number of items *specifically* about whistleblowing in software engineering (from the SE relevancy selection process) shows only a low numbers (peak of 5 per year in 2010 and 2019) with only small changes, both up and down, over time. Overall, the items break down into 3 distinct groups - conference and journal papers (49), book chapters (3 chapters) and magazine articles (8). The first 6 items chronologically are whistleblowing and ethics focused magazine articles. Education related papers, with whistleblowing case studies as part of ethics education, start to emerge in the late 1990s. Journal and conference papers begin appearing in the early 2000s with empirical research studies and technology solutions (supporting or detecting whistleblowing).

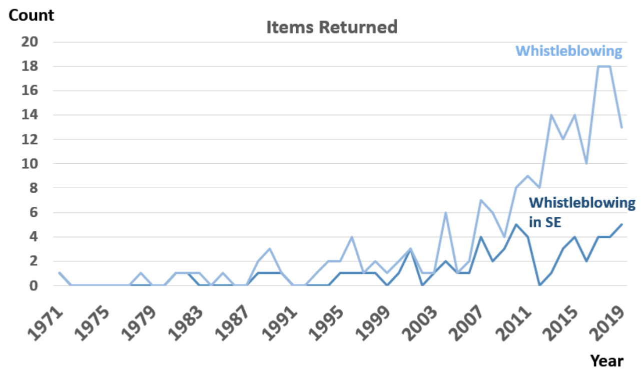

Figure 2: Temporal view (48 years) of search results

3.3.2 Authors and Citations. There are 48 primary authors with a total of 96 authors across all items. There is a total of 655 citations across the 57 non-book chapter items. The collective body of works from Keil, Park, Wang, Smith et al. (13 papers) related to bad news reporting on IT projects are the highest cited (114, 96, 85, 64, 49, 43 dropping to under 20 for the more recent papers) and account for over three quarters of all citations. Over 50% of the items have less than 2 citations, with 13 items having not been previously cited. It seems the field is dominated by a few key authors, indicating SE research is not giving it the attention it warrants and leading to a lack of diversity of voices and approaches to this topic.

3.3.3 Participant and Geographical Distribution. Over half the items are from work carried out in the USA, with at least 19 other countries represented. Empirical study participants are reasonably evenly split between students and professionals, and there is a similarly even split between Eastern and Western cultures, including 2 specific cross-cultural studies with the USA and South Korea. India, South America and Africa are not well represented as authors or participants. Whistleblowing stories are predominantly about USA-based organisations. Given the international spread of the SE development community alongside the complexity and global reach of IT systems, this area deserves looking at further to understand how whistleblowing happens internationally.

3.3.4 Publication Sources. An analysis of the publications for the 49 conference and journal papers reveals there are no items from top-tier SE venues such as ICSE, FSE, TSE and TOSEM. The majority of papers found were from Information Systems (14 items), Education (7 items), Computing in Society (4 items) or Ethics (3 items) publications. The magazine articles were found in a broad range of publications including IEEE Technology and Society, IEEE Spectrum, IEEE Computer and ACM Inroads. Given the impact of harmful software and SE practices to the SE community and wider society, it is encouraging to find articles discussing whistleblowing.

Table 5: Prevalence of Primary Themes and Papers ([19, 77] appear in two categories due to inter-rating differences)

Primary Theme	Sub-themes Include	Count	Individual Papers with Primary Theme
Personal and Social Factors	Personal values, motivation, social identity, decision making, mum effect (keeping quiet about issues), social responsibility	20	[55] [83] [26] [21] [54] [99] [98] [50] [102] [47] [70] [76] [13] [94] [114] [61] [3] [113] [85] [39]
Whistleblowing Stories	Case studies, engineering ethics, government, education, news and social media, history	15	[90] [19] [25] [53] [91] [10] [11] [44] [62] [2] [86] [35] [45] [32] [29]
Technology	Protection or detection. Platforms, protocols, vulnerabilities of technology, privacy enhancing technologies	12	[105] [74] [7] [92] [9] [36] [97] [117] [96] [72] [41] [58]
Whistleblowing Process	Theories, models, motivation, agency theory, decision making, behavioural reasoning theory	7	[118] [77] [78] [12] [48] [73] [106]
Organisational and Professional Issues	Roles, responsibilities, culture, professional code of ethics, deaf effect (not listening to issues), governance, policies, professional bodies, regulation	7	[77] [19] [79] [88] [49] [112] [71]
SE Aspects	Whistleblowing and software development lifecycle, coding, testing, design, open source, software piracy, code reuse, misuse of IT, security	1	[93]

4 DISCUSSION OF FINDINGS

With sixty items identified in nearly fifty years of publications, the main finding of our investigation is that there is a limited amount of research on whistleblowing in SE and none presented in top SE venues such as FSE or ICSE. We find this a concern on three main fronts: first, there are historic and recent case studies demonstrably showing both the potential and actual far-reaching harm from poor software or SE practices [65]; second, in order to identify malpractices in SE decision making processes, SE specialist knowledge is required; finally, the challenges faced by practitioners who witness harm or wrongdoing can be too great resulting in non-reporting and even leaving the workplace [67].

The main objective of this section is to expand and bring to light these issues by mapping the literature review findings onto our three research questions. We do so by first, reporting and discussing the themes identified in the literature through thematic analysis (RQ1). We then report on the approaches used to study whistleblowing in SE and identify some of their strengths and weaknesses (RQ2). Finally, we examine the gaps and concerns in whistleblowing research (RQ3).

4.1 RQ1- Whistleblowing Themes

- *RQ1: What are the prevalent whistleblowing themes in SE?*

We identify *Personal and Social Factors* as the top primary theme (N=20) , which includes sub-themes such as individual values, personal motivations, and social identity (i.e. belonging to a team) as Table 5 shows. The fact that this is the most recurrent theme is not surprising since whistleblowing is widely portrayed - both by research and media - as being about individuals making decisions according to their conscience, sense of responsibility, and "obligations" [1]. *Whistleblowing Stories* also features frequently as a theme (N=15), particularly in the context of higher education where whistleblowers' stories are used in computer and engineering ethics courses to illustrate to students real-life situations in which their morals and responsibilities may be challenged at work [12]. The *Technology* theme follows with 12 identified papers and including 10 items primarily concerned with the development of safe and secure channels for safeguarding whistleblowers' communication, alongside 2 looking at detection of whistleblowing activities. The *Organisational and Professional Issues* theme includes papers investigating organisational culture and structures affecting whistleblowing [114]. Finally, we found only one paper specifically covering harmful SE practices and their disclosure [93], discussing "dirty code" where developers take open source code snippets into closed source systems and so infringe on licenses and intellectual property rights. While *SE Aspects* are not prevalent as a primary theme, the authors still found it an important theme, with 14 items using SE aspects as part of back story in a study (software development lifecycle, coding, testing, design, open source, software piracy,code reuse, misuse of IT, security issues). The next section expands the discussion by presenting a selection of sub-themes identified as most relevant to SE research and practice.

4.1.1 Personal and Social Factors: the Mum Effect. Several items [47–50, 76–78, 85, 90, 98, 99, 112, 114] report on the *mum effect*, that is 'keeping quiet' about known issues or harmful situations. The mum effect is found to be a common behaviour for both practitioners, such as developers and IT auditors, who are found to be generally adverse to "reporting up bad news", and managers, who, in reverse, tend not to listen to risks or harms (the *deaf effect*). Most of the papers reporting on this issue [29, 49, 62, 90] focus on the relationship of this behaviour and the heightened risks of costly project failures. The punitive economic costs of project failures are a known issue in SE literature and practice. However, the impact of keeping quiet can have much wider implications than economic ones: from disrupting democratic processes, to costing people lives.

Keeping quiet is a matter of concern for the SE community. Given the technical complexity of software systems, understanding the harm that a poorly designed or ill-conceived software product may cause, often relies on specialised technical knowledge. The SE community is uniquely positioned to recognise, support and speak up about potentially harmful SE decisions and practices. If we are not looking into this, who can?

4.1.2 Whistleblowing Stories: Role in SE Education. Whistleblowing is often depicted as a personal act of moral responsibility, and as an heroic one, given the harrowing consequences that whistleblowers may face [25]. Whistleblowing stories have hence captured the media attention, and their educational power of telling stories [89] has been understandably harnessed in computing and engineering ethics courses [10–13, 44, 53, 61, 62, 88]. However, this teaching practice risks perpetuating the view that whistleblowing is a personal matter and of individual responsibility, whereas research and SE practice indicates that the responsibility of technology systems design, deployment and operations is much wider and distributed [16].

4.1.3 Technology: as detection or for protection. With whistleblowing sometimes being described as an "insider threat" to an organisation, there are 2 papers presenting technical solutions for the detection of whistleblowing activities. There are also 8 papers specifically looking at technology supporting the protection of a whistleblower's identity or making anonymous the source of whistleblowing submissions. Although, as [55] reflects, the confidentiality of disclosures may not prevent discovery of whistleblower by other means. Two items [41, 117] critique available whistleblowing platforms ahead of developing or selecting a whistleblowing platform solution. Bodo's paper [9] reports on aspects of privacy enhancing technologies, such as Wikileaks, and how it enables insiders to anonymously share private organisational information. On the detection side, Okolica et al.'s work on the ENRON case [74] describes using email analysis to identify signals for potential whistleblowers - i.e. having a hidden interest in sensitive topics. Finally, Maitre et al. [58] present a solution for discovering weak signals in web pages and twitter content, and extracting information "possibly sent by whistleblowers". To avoid such network detection, an innovative solution *AdLeaks*, is provided by Roth et al. [92], using online advertising as a cover for submitting and recovering whistleblowing reports.

This theme raises the wider issue of how online activity is monitored internally and externally to organisations, by whom and for what purposes that may closely relate to whistleblowing (i.e. insider threats, business rivals, regulators or the media looking for evidence and stories). We should therefore reflect on the possibility that technologies to help evidence and report SE harms, could also be used against practitioners who wish to speak up about such issues.

4.2 RQ2- How is Whistleblowing Researched?

- *RQ2: How is whistleblowing researched in SE?*

In this section we look at the approaches (or 'research agendas' [101]) used to report on or study whistleblowing in SE. We do so in two steps: first, we apply Stol and Fitzgerald's *ABC framework* [101]. We then apply Easterbrook et al.'s research categories [22] to guide the discussion around the empirical research papers found. Stol and Fitzgerald's ABC framework proposes an SE research classification that takes in consideration the actors (A) being studied, the measurements of their behaviour (B), and how realistic a context (C) the study is set in. We find this framework particularly relevant to our research, because whistleblowing ultimately focuses on actors (e.g. individuals, teams, organisations), whose behaviours (e.g. speaking-up) are situated in specific contexts. Following this framework, we classify our papers according to four research settings: Natural, Contrived, Neutral and Non-empirical.

Natural setting items feature most highly (39 items) in the review. The majority of items in this category were position, opinion and technology solutions papers based on in-field reports or observations (27 items). There were 6 case studies alongside 3 technology platform reviews, 2 field studies (one of which was based on secondary data) and 1 field experiment. Contrived (laboratory-based) setting research makes up nearly 25% of papers with scenario based experiments used to measure a participant's "intention to whistleblow". Neutral setting items included 3 sample studies and 1 survey. Finally there are 3 non-empirical papers proposing whistleblowing (2 items) and threat modelling (1 item) frameworks. With empirical items accounting for the largest portion of the total set, the discussion that follows primarily focuses on this set of papers, and specifically on research carried out in Natural (Field) and Contrived (Laboratory) settings.

4.2.1 Natural Settings (Field Studies). While there are a large number of opinion pieces and position papers on whistleblowing, there is a relatively low number (6 items) of field study papers in the review. All 6 items are centred around individual (personal) whistleblowers' stories. Their findings are not easily generalisable because of the specifics of each situation, and the varying degree of detail presented. The cases covered include safety critical systems [11, 44, 113], nuclear power [25, 53], health [21], transport [62], academia [19] and defence [29, 44] situations. There were 4 cases where issues were eventually reported to the media (Edward Snowden [106], GEC Nuclear [25], USA missile defense program [10, 29]), some tried disclosing issues internally before going to the media or government agencies. In the 13 individual stories reported, all except two whistleblowers lost their jobs (resigned or fired). Notably three whistleblowers received awards from the IEEE for outstanding service to public interest (Clinch River [12], Air Shield Incubator [54], GEC Nuclear [25]). The ACM debated awarding, but did not give, Snowden a “Making a Difference Award”[2].

Whistleblowing in SE is a complex social-technical phenomenon and its understanding would benefit from in-depth ethnographic studies. However such studies are not present in our review. This is likely for a number of reasons. First, whistleblowing is of a sensitive nature, making the design of research studies challenging, particularly when much of the process leading to whistleblowing happens under the radar. In addition, it may take course over a long time, thus requiring research resources that may not be easily available. Finally, the seriousness of the consequences of whistleblowing may also mean participants adjust behaviours during any observation period to avoid or reduce the risk of detection or retaliation.

4.2.2 Contrived Settings (Lab Experiments). In this review, laboratory experiments are found to be frequently used in the context of project failure scenarios and individual reluctance to report bad news. There were few examples of the misuse of successfully delivered IT systems, as found in the Volkswagen and Cambridge

Analytica stories. Aspects of SE such as software bugs, misunderstood requirements, system limitations, poor testing, outsourced work or 3rd party code are used to form the back story to the scenarios. Situational variables are carefully controlled in role play scenarios, with the likelihood (intention) to whistleblow then captured. Extreme scenarios are used to maximise variance [76, 112]. Situational factors such as professionalism [94], time urgency [76], harm of consequences [47, 77, 78, 99, 102] and proximity to victims [78] are shown to increase the likelihood of reporting bad news. However, the studies focus on an individual's response to a scenario, with only limited reflection on wider group discussions and dynamics (inside and outside an organisations) that might occur and have an impact on knowledge of choice of actions and so possible outcomes.

4.2.3 Combining Settings. The most cited paper in this review is Keil and Robey's 2001 field study [49]. Ten IT auditors were interviewed about their real experiences of past project failure situations and the eventual outcomes for the projects. Conditions were identified that affect the reluctance to blow the whistle and find that many internal auditors remain quiet (the mum effect) instead of asserting their responsibility to report bad news. More studies like this, taken down to the SE practitioner level could offer much insight for the SE community. Alongside this is Smith and Keil's 2003 theory development paper [98], bringing together factors on the reluctance to report bad news into the SE project domain. These papers lead to a group of controlled experiment type studies [47, 48, 50, 73, 76–78, 99, 112–114] spanning more than 10 years that contribute to the understanding of factors affecting responsibility and willingness to report bad news on IT projects.

4.3 RQ3- Research Gaps and Concerns

- *RQ3: What are the gaps and concerns for SE research and practice?*

Our gap analysis is based on the prevalence of themes found in RQ1, study types found in RQ2, and how they relate to human and technical aspects of SE. The combination of case studies, field studies and laboratory controlled experiments has contributed to a better understanding of factors affecting the likelihood of whistleblowing in SE practice. However, the studies found lacked diversity in research perspectives, and the number of authors involved was small.

4.3.1 SE Aspects. Of particular concern is the absence of recent in-practice studies exploring the harmful situations that SE teams see, and how they resolved or escalated serious issues with actions up to and including whistleblowing. When software practitioners perceive their team or organisation is creating or operating harmful software (e.g. is defective, insecure, vulnerable, malicious or illegal) what actions might they take? Miller and Coldicott's IPSOS Mori survey [67] finds 90% of people do take action, but what exactly were those actions, how effective was it and what was the final outcome? We found very limited published work examining these aspects from an SE perspective. Within the software development lifecycle, research is needed to improve our understanding of the escalation and effectiveness of whistleblowing actions to change harmful SE situations. These studies need to include actual harms seen that were attributed to software, SE practices or SE decisions, that trigger actions up to and including whistleblowing.

4.3.2 Missing perspectives. Whistleblowing stories are often described from the whistleblower's perspective, with little focus on the involvement and responses from other SE practitioners, disclosure recipients (internal or external), organisations, and groups in wider society made aware of these SE situations through the news and social media. Reports of high-profile software scandals in the media appear to be increasing and so too the scrutiny of software organisations to show consideration of the social and human impact of the systems they design, build and operate [81]. In 2021, there have been a succession of whistleblowing stories in the media [8], with former SE practitioners from companies such as Google [20] and Facebook [82] speaking out, with evidence, about the harms caused by features of their products (e.g. algorithmic bias and putting profit before public good). Researching new and emerging SE stories, presented as comparable cases studies, would help us map out and understand the growing involvement of the wider SE community and campaign groups, and how that impacts the effectiveness of whistleblowers attempting to mitigate harm. Stories would also provide relevant and modern case studies for SE education and practice and help address the perception that reporting wrongdoing is a matter of individual morals and responsibility.

4.4 Implications for Research and Practice

We conclude with three recommendations for research and practice. First, we recommend more timely identification and mitigation of harm using values-led initiatives in SE practices; second we suggest some possible mechanisms for improving the effectiveness and the personal safety of harm reporting in SE, and finally we reflect on the role that professional bodies can have in supporting harm reporting, up to and including whistleblowing.

4.4.1 Supporting early identification and mitigation of harm. In 2020 it was reported that UK councils and government bodies were withdrawing use of automated decision making software in services such as welfare, visa applications and exam grading [59]. Cancelling reasons ranged from problems in the way the systems work to concerns about negative effects and bias with regards to end users. The SE community must continue working to make software, decisions and artefacts transparent and auditable, this will help identify harmful situations in a timely fashion to mitigate harmful outcomes. For example, this could be achieved by enhancing existing SE practices with values-led tools as described in [115] or through new automated SE values detection techniques [28]. Understanding the effectiveness of SE practices and being transparent about points of success and failure [103] may help practitioners "avoid misguided reporting" and support them to "carefully assess relevant aspects of the situation" as advised by the ACM Code of Ethics [1].

4.4.2 Improving effectiveness and safety of harm reporting. In the USA, a Volkswagen engineer has been jailed for the role he played in a team of engineers responsible for emissions tests cheating software [40]; in the UK there are criminal investigations against the Post Office and Fujitsu IT staff [111] for cover ups linked to IT systems that left hundreds of subpostmasters suspended, sacked or criminally prosecuted for theft, fraud and false accounting. External whistleblowing from campaign groups and regulatory bodies, after many years, finally brought these stories to the public attention. It is not known if internal SE teams tried to evidence and speak out about the issues, and if they did, how effective it was. The research shows that SE practitioners find it difficult to report issues, and even more so to speak up about the actions and behaviours of other SE practitioners, teams, and organisations [47–50, 76–78, 85, 98, 99, 112, 114]. We argue that a deep understanding of factors that contribute to whistleblowing effectiveness [69] is an important stepping stone towards understanding how harm (potential or actual) can be safely, responsibly and effectively disclosed.

As an example, in a recent legal paper on recommendations for the probity of computer evidence, linked to the Post Office trials [111], Marshall et al. report that the kind of documents that are likely to exist, and ought to be disclosed are not generally well-understood by people without professional computing or SE knowledge [60]. This highlights a procedural dilemma: when a request for a disclosure of a software issue is made, the lack of specific SE knowledge to locate and examine the required evidence may result in the dismissal of the request. Their paper proposes a two-step software-issue disclosure process emerging from cases involving the cover-up of harmful software practices and products. Solutions and professional standards to support such legal challenges would benefit all stakeholders in the software development lifecycle, and in particular help reduce the burden on individuals and organisations seeking to produce evidence in support or defence of whistleblowing actions.

4.4.3 The role of professional bodies and organisations. Kline [53] and Hersh [34] call for a study on the actions (and in-actions) of professional bodies in respect of whistleblowing cases, whilst research and advocacy group stress the importance for organisations to understand their obligations, and for individuals to be aware of the support and protection available to them should they choose to take whistleblowing action [109, 110]. An IEEE statement elaborates why they (the IEEE) are not always in a position to support whistleblower cases: "diverse opinions may flow from the same set of circumstances [...] and there are always at least one other side to a story [...] the IEEE does not have the responsibility to make a case for the member submitting a complaint or request for support" [83]. Their statement is 30 years old, does it still ring true? Zelby [118] raised the suggestion of a mediation-like authority to protect all parties involved in whistleblowing cases. Other alternatives are the support of trade unions, who look to protect worker rights and public interests, as recognised by the emergence of new technology specific trade unions [64, 80].

5 THREATS TO VALIDITY

We identify potential threats to the validity pertaining to 1) specific search terms used 2) mis-classification and 3) publication sources. Whilst we cannot guarantee that some SE relevant items have not been excluded or missed, we argue that the review is representative of how whistleblowing in SE is studied, and does not rely on single items for its findings.

5.1 Construct Validity - Search Terms

Whistleblowing is a very specific search term - defining precisely the phenomenon we are looking at. We used 12 variations of it to find matching items. Other key words such as "dissent", "bad news reporting", and "mis-reporting" were explored but we found searching on whistleblowing was specific and sufficient for this paper's research objectives.

5.2 Internal Validity - Mis-classification

Only two researchers were involved in the thematic analysis. To mitigate this, the process was rigorous, systematic and included two rounds of independent coding and calibration [33].

5.3 External Validity - Publication Sources

The search was limited to the IEEE, ACM and Scopus databases with no snowballing of references from or to each item. This does limit the validity of our findings. In mitigation, the 3 databases used are common sources for SE research. Whistleblowing studies outside of the SE domain are being reviewed and may uncover studies that include SE aspects or cases.

6 CONCLUSIONS AND FUTURE WORK

Whistleblowing has received growing media attention in the last few years, and our study has been part-motivated by following stories and reviewing a considerable amount of grey literature, including [8, 56, 65, 104, 116]. There we find whistleblowing often portrayed as a tragic, heroic or responsible act of individuals, with limited value for understanding the specific SE aspects and human factors behind a story. On reviewing literature from SE top-venues to seek actionable whistleblowing guidelines for software engineers, we did not find the research breadth and depth we hoped for. However, we argue that, just because whistleblowing in SE is not common (and is difficult to study), does not make it unimportant. Our future research seeks to create a body of evidence (case studies) to help the SE community identify effective mitigating actions (technical, human and policies) that can help resolve, not escalate, harmful situations relating to software or SE practices. Research suggests that seeing connections between the consequences of software decisions through examples of similar stories may influence and inform future ethical decision making [63]. Developing and publishing SE whistleblowing cases studies, to the SE community will help support this awareness.

Our findings indicate that while whistleblowing is increasingly mentioned in the literature, it is an under-explored area of SE research, with no whistleblowing related papers found in top-tier SE venues such as ICSE and FSE. The SE community has a responsibility to society to demonstrate that it assesses and understands the potential risks and consequences associated with the software we design and build and has mechanisms that allow practitioners to speak out if professional values and standards are not being met or followed. It requires us to have processes and procedures in place for risks and harms to be identified and evidenced whilst creating a safe environment for their reporting. In particular, the burden of responsibility to disclose harmful situations should not be left to individuals, given existing studies find both practitioners and IT auditors are often deterred from speaking up (the mum effect) and management are reluctant to listen to their concerns (the deaf effect).

Moving forward, the SE research community is uniquely positioned to explore and advance the understanding of harmful situations in SE, with whistleblowing being an extreme form of harmful situation reporting. Future work should aim to explore how and why whistleblowing is, or indeed is not, happening in modern SE practice. Carefully designed field based approaches are required, sympathetic to the complexities and risks associated with whistleblowing studies, to evidence and reflect on the types of harmful code, software or SE practices that lead to potential and actual whistleblowing situations.

ACKNOWLEDGMENTS

We thank Dr Emily Winter, Dr Chris Bull and Prof Austen Rainer for their comments on an early version of this paper. We also thank the anonymous reviewers for their detailed feedback. This work received support from the Engineering and Physical Sciences Research Council (EPSRC) UK "Values in Computing" project, grant number: EP/R009600/1.

REFERENCES

[1] ACM. 2018. ACM Code of Ethics and Professional Conduct. https://www.acm.org/code-of-ethics

[2] A. A. Adams. 2014. Report of a Debate on Snowden's Actions by ACM Members. *SIGCAS Comput. Soc.* 44, 3 (Oct. 2014), 5–7. https://doi.org/10.1145/2684097.2684099

[3] Greg Adamson. 2015. Ethical challenges for future technologists: The growing role of technology and the growing ethical responsibility of the technologist. In *2015 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, IEEE, 1–6.

[4] CF Alford. 2007. Whistle-blower narratives: The experience of choiceless choice. *Social Research* 74, 1 (2007), 223–248.

[5] Farid Anvari, Michael Wenzel, Lydia Woodyatt, and S Alexander Haslam. 2019. The social psychology of whistleblowing: An integrated model. *Organizational Psychology Review* 9, 1 (2019), 41–67.

[6] Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, Javier Del Ser, Adrien Bennetot, Siham Tabik, Alberto Barbado, Salvador Garcia, Sergio Gil-Lopez, Daniel Molina, Richard Benjamins, Raja Chatila, and Francisco Herrera. 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58 (2020), 82–115. https://doi.org/10.1016/j.inffus.2019.12.012

[7] Graeme Baxter Bell. 2011. Digital whistleblowing in restricted environments. *Journal of Digital Information* 12, 3 (2011), 1–14.

[8] Johana Bhuiyan. 2021. 'Welcome to the party': five past tech whistleblowers on the pitfalls of speaking out | Technology | The Guardian. https://www.theguardian.com/technology/2021/oct/08/tech-whistleblowers-facebook-frances-haugen-amazon-google-pinterest

[9] Balázs Bodó. 2014. Hacktivism 1-2-3: how privacy enhancing technologies change the face of anonymous hacktivism. *Internet Policy Review* 3, 4 (2014), 1–13.

[10] Kevin Bowyer. 1997. Case study resources for an ethics and computing course. In *Proceedings Frontiers in Education 1997 27th Annual Conference. Teaching and Learning in an Era of Change*, Vol. 1. IEEE, 469–473.

[11] Kevin W Bowyer. 2000. Goodearl and Aldred versus Hughes Aircraft: A whistleblowing case study. In *30th Annual Frontiers in Education Conference. Building on A Century of Progress in Engineering Education. Conference Proceedings (IEEE Cat. No. 00CH37135)*, Vol. 2. IEEE, S2F–2.

[12] Kevin W Bowyer. 2001. *Whistle Blowing*. Wiley-IEEE Press.

[13] Bo Brinkman. 2009. The heart of a whistle-blower: a corporate decision-making game for computer ethics classes. In *Proceedings of the 40th ACM technical symposium on Computer science education*. 316–320.

[14] Care Quality Commission. 2017. Quick guide to raising a concern about your workplace | Care Quality Commission. https://www.cqc.org.uk/news/stories/quick-guide-raising-concern-about-your-workplace

[15] Charity Commission for England and Wales. 2013. The Public Interest Disclosure Act - GOV.UK. https://www.gov.uk/government/publications/the-public-interest-disclosure-act/the-public-interest-disclosure-act

[16] Mark Coeckelbergh. 2012. Moral responsibility, technology, and experiences of the tragic: From Kierkegaard to offshore engineering. *Science and engineering ethics* 18, 1 (2012), 35–48.

[17] Jacob Cohen. 1960. A Coefficient of Agreement for Nominal Scales. *Educational and Psychological Measurement* 20, 1 (1960), 37–46. https://doi.org/10.1177/001316446002000104 arXiv:https://doi.org/10.1177/001316446002000104

[18] Daniela S Cruzes and Tore Dyba. 2011. Recommended steps for thematic synthesis in software engineering. In *2011 international symposium on empirical software engineering and measurement*. IEEE, 275–284.

[19] Jaime A Teixeira da Silva and Judit Dobránszki. 2019. A new dimension in publishing ethics: social media-based ethics-related accusations. *Journal of Information, Communication and Ethics in Society* 17, 3 (2019), 354–370. https://doi.org/10.1108/JICES-05-2018-0051

[20] Tom Dimonite. 2021. What Really Happened When Google Ousted Timnit Gebru | WIRED. https://www.wired.com/story/google-timnit-gebru-ai-what-really-happened/

[21] Joseph F Dyro. 1988. Meditation on ethics in clinical engineering practice. *IEEE Engineering in Medicine and Biology Magazine* 7, 2 (1988), 77–80.

[22] Steve Easterbrook, Janice Singer, Margaret-Anne Storey, and Daniela Damian. 2008. Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering*. Springer, 285–311.

[23] European Commission. 2018. Robust protection for whistleblowers across EU: Commission proposes new rules - European Commission. https://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=620400

[24] European Council. 2019. Better protection of whistle-blowers: new EU-wide rules to kick in in 2021. https://www.consilium.europa.eu/en/press/press-releases/2019/10/07/better-protection-of-whistle-blowers-new-eu-wide-rules-to-kick-in-in-2021/

[25] Karen Fitzgerald. 1990. Whistle-blowing: not always a losing game (engineers). *IEEE Spectrum* 27, 12 (1990), 49–52.

[26] Samuel C Florman. 1982. Careers: A skeptic views ethics in engineering: Competing with recession and unemployment for people's concern, ethics still manages to arouse lively debates. *IEEE Spectrum* 19, 8 (1982), 56–57.

[27] International Organization for Standardization. 2008. *Systems and Software Engineering: Software Life Cycle Processes*. ISO.

[28] Sainyam Galhotra, Yuriy Brun, and Alexandra Meliou. 2017. Fairness testing: testing software for discrimination. In *Proceedings of the 2017 11th Joint Meeting on Foundations of Software Engineering*. 498–510.

[29] Subrata Ghoshroy. 2019. The Price for Blowing the Whistle when Facing Ethical Dilemmas. In *2019 IEEE International Symposium on Technology and Society (ISTAS)*. IEEE, 1–6.

[30] Robert L. Glass, Iris Vessey, and Venkataraman Ramesh. 2002. Research in software engineering: an analysis of the literature. *Information and Software technology* 44, 8 (2002), 491–506.

[31] Emma Graham-Harrison and Carole Cadwalladr. 2018. *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*. The Guardian. https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election

[32] Gehan Gunasekara, Andrew A Adams, and Kiyoshi Murata. 2017. Ripples down under: New Zealand youngsters' attitudes and conduct following Snowden. *Journal of Information, Communication and Ethics in Society* 15, 3 (2017), 297–310. https://doi.org/10.1108/JICES-10-2016-0042

[33] Kilem L Gwet. 2014. *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters*. Advanced Analytics, LLC.

[34] Marion A Hersh. 2002. Whistleblowers—heroes or traitors?: Individual and collective responsibility for ethical behaviour. *Annual reviews in Control* 26, 2 (2002), 243–262.

[35] Arne Hintz and Lina Dencik. 2016. The politics of surveillance policy: UK regulatory dynamics after Snowden. *Internet Policy Review* 5, 3 (2016), 16 pages. https://doi.org/10.14763/2016.3.424

[36] Susan Hohenberger, Steven Myers, Rafael Pass, et al. 2015. An overview of ANONIZE: A large-scale anonymous survey system. *IEEE Security & Privacy* 13, 2 (2015), 22–29.

[37] IEEE. 2018. In *2018 IEEE/ACM International Workshop on Software Fairness (FairWare)*, Vol. 1. IEEE Computer Society, Los Alamitos, CA, USA, 5–6. https://doi.org/10.23919/FAIRWARE.2018.8452907

[38] J. P. Near and M. P. Miceli. 2016. After the wrongdoing: What managers should know about whistleblowing. *Business Horizons* 59, 1 (2016), 105–114.

[39] Lennart Jaeger and Andreas Eckhardt. 2018. When Colleagues Fail: Examining the Role of Information Security Awareness on Extra-Role Security Behaviors. In *26th European Conference on Information Systems: Beyond Digitization - Facets of Socio-Technical Change, ECIS 2018*. https://aisel.aisnet.org/ecis2018_rp/124

[40] Kasperkevic. Jana. 2016. Volkswagen engineer pleads guilty to conspiracy in emissions scandal. https://www.theguardian.com/business/2016/sep/09/volkswagen-engineer-pleads-guilty-conspiracy-emissions-scandal-

[41] Hariharan Jayakrishnan and Ritwik Murali. 2019. A Simple and Robust End-to-End Encryption Architecture for Anonymous and Secure Whistleblowing. In *2019 Twelfth International Conference on Contemporary Computing (IC3)*. IEEE,

1–6.

[42] Jasper Jolly and Gwyn Topham. 2019. Boeing's suspension of 737 Max production rattles airline industry. https://www.theguardian.com/business/2019/dec/17/shares-in-boeing-uk-suppliers-fall-after-737-max-production-halted

[43] Peter B Jubb. 1999. Whistleblowing: A restrictive definition and interpretation. *Journal of Business Ethics* 21, 1 (1999), 77–94.

[44] K. W. Bowyer. 2001. "Star Wars" revisited-a continuing case study in ethics and safety-critical software. In *Proceedings International Symposium on Technology and Society.* IEEE, 51–60.

[45] Iordanis Kavathatzopoulos, Ryoko Asai, Andrew A Adams, and Kiyoshi Murata. 2017. Snowden's revelations and the attitudes of students at Swedish universities. *Journal of Information, Communication and Ethics in Society* 15, 3 (2017), 247–264. https://doi.org/10.1108/JICES-02-2017-0009

[46] David Kaye. 2015. *Promotion and protection of the right to freedom of opinion and expression.* Technical Report. United Nations. 13–15 pages. https://www.undocs.org/A/70/361

[47] Mark Keil, Ghi Paul Im, and Magnus Mähring. 2007. Reporting bad news on software projects: the effects of culturally constituted views of face-saving. *Information Systems Journal* 17, 1 (2007), 59–87.

[48] Mark Keil and ChongWoo Park. 2010. Bad news reporting on troubled IT projects: Reassessing the mediating role of responsibility in the basic whistleblowing model. *Journal of Systems and Software* 83, 11 (2010), 2305–2316.

[49] M Keil and Daniel Robey. 2001. Blowing the whistle on troubled software projects. *Commun. ACM* 44, 4 (2001), 87–93.

[50] Mark Keil, H Jeff Smith, Suzanne Pawlowski, and Leigh Jin. 2004. 'Why didn't somebody tell me?' climate, information asymmetry, and bad news about troubled projects. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems* 35, 2 (2004), 65–84.

[51] Kate Kenny, Marianna Fotaki, and Wim Vandekerckhove. 2020. Whistleblower subjectivities: Organization and passionate attachment. *Organization Studies* 41, 3 (2020), 323–343.

[52] Barbara Kitchenham. 2004. Procedures for performing systematic reviews. *Keele, UK, Keele University* 33, 2004 (2004), 1–26.

[53] Ronald R Kline. 2010. Engineering case studies: Bridging micro and macro ethics. *IEEE Technology and Society Magazine* 29, 4 (2010), 16–19.

[54] Jean Kumagai. 2004. The whistle-blower's dilemma. *IEEE Spectrum* 41, 4 (2004), 53–55.

[55] Harry T Larson. 1971. On whistle blowing. *Computer* 4, 4 (1971), 34–34.

[56] Dave Lee. 2021. US regulator reviews safety concerns at Bezos space group Blue Origin | Financial Times. https://www.ft.com/content/122b28c7-9908-45a7-a96c-ca52399311d0

[57] Per Lenberg, Robert Feldt, and Lars-Göran Wallgren. 2014. Towards a behavioral software engineering. In *Proceedings of the 7th international workshop on cooperative and human aspects of software engineering.* 48–55.

[58] Julien Maitre, Michel Ménard, Guillaume Chiron, Alain Bouju, and Nicolas Sidère. 2019. A meaningful information extraction system for interactive analysis of documents. In *2019 International Conference on Document Analysis and Recognition (ICDAR).* IEEE, 92–99.

[59] Sarah Marsh. 2020. Councils scrapping use of algorithms in benefit and welfare decisions. https://www.theguardian.com/society/2020/aug/24/councils-scrapping-algorithms-benefit-welfare-decisions-concerns-bias

[60] Paul Marshall, James Christie, B Ladkin, Bev Littlewood, Stephen Mason, Martin Newby, Jonathan Rogers, Harold Thimbleby, and M Thomas. 2020. Recommendations for the probity of computer evidence. *Digital Evidence and Electronic Signature Law Review* 18 (2020).

[61] C.D. Martin. 2011. Reasoning with ethics. *ACM Inroads* 2, 1 (2011), 8–9. https://doi.org/10.1145/1929887.1929889 cited By 1.

[62] Donald J McCubbrey and Cynthia V Fukami. 2009. ERP at the Colorado department of transportation: The whistle blower's dilemma. *Communications of the Association for Information Systems* 24, 1 (2009), 7.

[63] Andrew McNamara, Justin Smith, and Emerson Murphy-Hill. 2018. Does ACM's code of ethics change ethical decision making in software development?. In *Proceedings of the 2018 26th ACM joint meeting on european software engineering conference and symposium on the foundations of software engineering.* 729–733.

[64] Emiliano Mellino. 2020. Google and Microsoft staff set to join the UK's first tech trade union | WIRED UK. https://www.wired.co.uk/article/united-tech-and-allied-workers-union

[65] R Merkel. 2015. *Where were the whistleblowers in the Volkswagen emissions scandal.* The Conversation. https://theconversation.com/where-were-the-whistleblowers-in-the-volkswagen-emissions-scandal-48249

[66] Jessica R Mesmer-Magnus and Chockalingam Viswesvaran. 2005. Whistleblowing in organizations: An examination of correlates of whistleblowing intentions, actions, and retaliation. *Journal of business ethics* 62, 3 (2005), 277–297.

[67] C Miller and R Coldicott. 2019. *People, power and technology: The tech workers' view.* DotEveryone (Open Data Institute). https://www.doteveryone.org.uk/report/workersview/

[68] Janet P Near and Marcia P Miceli. 1985. Organizational dissidence: The case of whistle-blowing. *Journal of business ethics* 4, 1 (1985), 1–16.

[69] Janet P Near and Marcia P Miceli. 1995. Effective-whistle blowing. *Academy of management review* 20, 3 (1995), 679–708.

[70] Yuan Niu, Antonis C Stylianou, and Susan J Winter. 2008. Blowing the Whistle on Unethical Information Techmology Practices: The Role of Machiavellianism, Gender and Computer Liteacy. *AMCIS 2008 Proceedings* (2008), 269.

[71] Nor Raihana Asmar Mohd Noor and Noorhayati Mansor. 2019. Exploring the Adaptation of Artificial Intelligence in Whistleblowing Practice of the Internal Auditors in Malaysia. *Procedia Computer Science* 163 (2019), 434–439.

[72] Muhammad Nursalman, Ria Anggraeni, and Muhamad Zulfikar Firdaus. 2018. Application of Layered Architecture in Whistleblowing Information System for Supporting Good University Governance in Indonesia University of Education. In *2018 International Conference on Information Technology Systems and Innovation (ICITSI).* IEEE, 49–54.

[73] Lih-Bin Oh and Hock-Hai Teo. 2010. To blow or not to blow: An experimental study on the intention to whistleblow on software piracy. *Journal of Organizational Computing and Electronic Commerce* 20, 4 (2010), 347–369.

[74] James S Okolica, Gilbert L Peterson, and Robert F Mills. 2007. Using Author Topic to detect insider threats from email traffic. *digital investigation* 4, 3-4 (2007), 158–164.

[75] Oxford Economics. 2014. *The Cost of Brain Drain.* Technical Report. Unum. https://www.oxfordeconomics.com/recent-releases/the-cost-of-brain-drain

[76] ChongWoo Park, Ghiyoung Im, and Mark Keil. 2008. Overcoming the mum effect in IT project reporting: Impacts of fault responsibility and time urgency. *Journal of the Association for Information Systems* 9, 7 (2008), 1.

[77] Chongwoo Park and Mark Keil. 2007. Organizational Factors and Bad News Reporting on Troubled IT Projects. *AMCIS 2007 Proceedings* (2007), 92.

[78] ChongWoo Park, Mark Keil, and Jong Woo Kim. 2008. The effect of IT failure impact and personal morality on IT project reporting behavior. *IEEE Transactions on Engineering Management* 56, 1 (2008), 45–60.

[79] PD Park. 1996. Whistleblowing: an employer's view how vulnerable is your company. *Engineering Management Journal* 6, 4 (1996), 183–186.

[80] Kari Paul. 2021. More than 200 US Google employees form a workers' union. https://www.theguardian.com/technology/2021/jan/04/more-than-200-us-google-employees-form-union

[81] Kari Paul and Dan Milmo. 2021. Congress grills Facebook exec on Instagram's harmful effect on children | Facebook | The Guardian. https://www.theguardian.com/technology/2021/sep/30/facebook-hearing-testimony-instagram-impact

[82] Kari Paul and Dan Milmo. 2021. Facebook putting profit before public good, says whistleblower Frances Haugen | Facebook | The Guardian. https://www.theguardian.com/technology/2021/oct/03/former-facebook-employee-frances-haugen-identifies-herself-as-whistleblower

[83] Tekla S Perry. 1981. Ethics: Knowing how to blow the whistle: Speaking out about an employer's unethical practices may bring public esteem to the whistle blower, but the many possible pitfalls must be considered. *IEEE Spectrum* 18, 9 (1981), 56–61.

[84] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. Systematic mapping studies in software engineering. In *12th International Conference on Evaluation and Assessment in Software Engineering (EASE) 12.* 1–10.

[85] Stacie Petter. 2018. If You Can't Say Something Nice: Factors Contributing to Team Member Silence in Distributed Software Project Teams. In *Proceedings of the 2018 ACM SIGMIS Conference on Computers and People Research.* 43–49.

[86] Charles P Pfleeger. 2016. Looking into Software Transparency. *IEEE Security & Privacy* 14, 1 (2016), 31–36.

[87] SC Plotkin. 1989. Economic Survival And whistle-blowing: One Solution. In *Delicate Balance: Technics, Culture and Consequences.* IEEE, 86–89.

[88] David Preston. 1998. What makes professionals so difficult: an investigation into professional ethics teaching. In *Proceedings of the ethics and social impact component on Shaping policy in the information age.* 58–67.

[89] Bernard R Robin. 2008. Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory into practice* 47, 3 (2008), 220–228.

[90] Johann Rost and Robert L Glass. 2011. *The dark side of software engineering: evil on computing projects: Whistleblowing Chapter 7.* John Wiley & Sons.

[91] Johann Rost and Robert L Glass. 2011. *The dark side of software engineering: evil on computing projects: Whistleblowing Chapter 7 Appendix.* John Wiley & Sons.

[92] Volker Roth, Benjamin Güldenring, Eleanor Rieffel, Sven Dietrich, and Lars Ries. 2013. A secure submission system for online whistleblowing platforms. In *International Conference on Financial Cryptography and Data Security.* Springer, 354–361.

[93] Johan Sarkinen. 2007. An open source (d) controller. In *INTELEC 07-29th International Telecommunications Energy Conference.* IEEE, 761–768.

[94] Richard AM Schilhavy and Ruth C King. 2010. Who Says Professionals Are Ethical? A Cross-sectional Analysis of Ethical Decision Making, Attitudes and Action.. In *AMCIS.* 568.

[95] Mary Shaw. 2003. Writing good software engineering research papers. In *25th International Conference on Software Engineering, 2003. Proceedings.* IEEE, 726–736.

[96] Laurens Sion, Koen Yskout, Dimitri Van Landuyt, and Wouter Joosen. 2018. Risk-based design security analysis. In *Proceedings of the 1st International Workshop*

on Security Awareness from Design to Deployment. 11–18.

[97] Thomas Sloan and Julio Hernandez-Castro. 2015. Forensic analysis of video steganography tools. *PeerJ Computer Science* 1 (2015), 1–e7.

[98] H Jeff Smith and Mark Keil. 2003. The reluctance to report bad news on troubled software projects: a theoretical model. *Information Systems Journal* 13, 1 (2003), 69–95.

[99] H Jeff Smith, Mark Keil, and Gordon Depledge. 2001. Keeping mum as the project goes under: Toward an explanatory model. *Journal of Management Information Systems* 18, 2 (2001), 189–227.

[100] Statista. 2021. Programmers & software developers in the UK 2020 | Statista. https://www.statista.com/statistics/318818/numbers-of-programmers-and-software-development-professionals-in-the-uk/

[101] Klaas-Jan Stol and Brian Fitzgerald. 2018. The ABC of software engineering research. *ACM Transactions on Software Engineering and Methodology (TOSEM)* 27, 3 (2018), 1–51.

[102] Tsjalling Swierstra and Jaap Jelsma. 2006. Responsibility without moralism in technoscientific design practice. *Science, technology, & human values* 31, 3 (2006), 309–332.

[103] Matthew Syed. 2015. *Black box thinking: why most people never learn from Their mistakes–But some Do.* Penguin.

[104] Amelia Tait. 2020. Susan Fowler: 'When the time came to blow the whistle on Uber, I was ready' | Sexual harassment | The Guardian. https://www.theguardian.com/world/2020/mar/01/susan-fowler-uber-whistleblower-interview-travis-kalanick

[105] Noburou Taniguchi, Koji Chida, Osamu Shionoiri, and Atsushi Kanai. 2005. DECIDE: a scheme for decentralized identity escrow. In *Proceedings of the 2005 workshop on Digital identity management.* 37–45.

[106] Herman T. Tavani and Frances S. Grodzinsky. 2014. Trust, Betrayal, and Whistle-Blowing: Reflections on the Edward Snowden Case. *SIGCAS Comput. Soc.* 44, 3 (Oct. 2014), 8–13. https://doi.org/10.1145/2684097.2684100

[107] Steven J Thoma, JR Rest, and Robert Barnett. 1986. Moral judgment, behavior, decision making, and attitudes. *Moral development: Advances in research and theory* 133 (1986), 175.

[108] Chantal Toporow, McCagie Rogers, Nik Warren, and Justin Biddle. 1990. *A Delicate Balance: Technics, Culture, and Consequences: California State University, Los Angeles, October 20-21, 1989.* Institute of Electrical & Electronics Engineers (IEEE).

[109] Wim Vandekerckhove. 2012. *UK public attitudes to whistleblowing.* Technical Report. University of Greenwich, UK.

[110] Wim Vandekerckhove, Cathy James, and Francesca West. 2013. *Whistleblowing: the inside story-a study of the experiences of 1,000 whistleblowers.* Technical Report. Public Concern at Work. https://gala.gre.ac.uk/id/eprint/10296/

[111] Nick Wallis. 2019. *Post Office Trial: Horizon trial judgment is handed down.* PostOfficeTrial.com. https://www.postofficetrial.com/2020/01/horizon-trial-judgment-is-handed-down.html

[112] Jijie Wang, Mark Keil, Lih-bin Oh, and Yide Shen. 2017. Impacts of organizational commitment, interpersonal closeness, and Confucian ethics on willingness to report bad news in software projects. *Journal of Systems and Software* 125 (2017), 220–233.

[113] J. Wang, M. Keil, and L. Wang. 2015. The effect of moral intensity on it employees' bad news reporting. *Journal of Computer Information Systems* 55, 3 (2015), 1–10. https://doi.org/10.1080/08874417.2015.11645766 cited By 2.

[114] Jijie Wang and Lih-Bin Oh. 2011. The Impact Of Relationships And Confucian Ethics On Chinese Employees' Whistle-Blowing Willingness In Software Projects.. In *PACIS*. 208.

[115] Emily Winter, Stephen Forshaw, Lucy Hunt, and Maria Angela Ferrario. 2019. Towards a systematic study of values in SE: tools for industry and education. In *2019 IEEE/ACM 41st International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER).* IEEE, 61–64.

[116] Christopher Wylie. 2019. *Mindf* ck: Inside Cambridge Analytica's plot to break the world.* Profile Books.

[117] Irma Zakia, Asep Hikman Fatahillah, Nana Rachmana Syambas, Asih Setiawati, and Housny Mubarok. 2017. Aspiration and complaint system: From literature survey to implementation. In *2017 11th International Conference on Telecommunication Systems Services and Applications (TSSA).* IEEE, 1–6.

[118] Leon W Zelby. 1989. Whistle-blowing-'Somebody has to take a stand'. *IEEE Technology and Society Magazine* 8, 3 (1989), 4–6.