

How to Navigate Crowd Crushes History? A Compilation of Six Existing Sources

Capucine-Marin Dubroca-Voisin

AREP Flux et Mobilités, Paris, France,

E-mail: capucine-marin.dubroca-voisin@arep.fr

LVMT, École des Ponts, Université Gustave Eiffel, Marne-la-Vallée, France,

E-mail: capucine-marin.dubroca-voisin@enpc.fr

Received: 27 July 2023 / Last revision received: 26 October 2023 / Accepted: 27 November 2023

DOI: [10.17815/CD.2023.145](https://doi.org/10.17815/CD.2023.145)

Abstract We use six existing sources about past crowd crushes and accident to build a merged data base. By doing so, we show that every source has a partial view of the crowd crushes, with coverage rates between 14% and 59% of our merged data base. Each of the sources contains crowd crushes that are cited by none of the other sources.

We then may have a very partial view of past crowd crushes. We examine several biases that can explain under reporting of crowd crushes, notably the less recent ones and the smaller.

This partial view affects any statistical study that we can do on the evolution of crowd crushes. However, our data analysis suggests that the number of crowd crushes par capita is not steadily increasing. Crowd crushes may not be a growing tendency in regard to global population.

Our analysis suggests it is necessary to continue studying crowd crushes, both globally and in-depth, to gain a more global view of their reasons and their tendencies. We propose to use collaborative projects such as Wikidata to do so.

Keywords Crowds · safety · crowd crush · crowd accident · history · stampede

1 Introduction

Crowd crushes, or crowd accidents, have been a matter of concern for event organisers, transport operators and local authorities for decades. They represent the worst scenario for crowd managements, with possible hundreds of deaths and injuries. Many events failures led the organisers to courts, and some of these events, such as the Love Parade disaster in

2010 or the Hillsborough stadium crush in 1989, led to an intense mediatisation and best practice evolution.

It is often heard that these deadly accidents are rapidly growing, with the ongoing urbanisation and the tendency to gather in big groups. In an article published in 2023 [1], several researchers confirm that the number of events and deaths had been on the rise, but they link it to a better availability of information.

We matched their data base with 5 other existing sources on crowd crushes (Wikipedia, Wikidata, Keith Still's list, the World Crowd Disasters Web App and the Working with Crowds website) to see which of these lists were complete. We then ran similar analyses about the number of events and deaths, and also looks at the number of deaths per capita.

Finally, we examined the different biases and their impact on information availability, and we proposed a scenario to collectively gather this data in future work.

2 Methods

2.1 Sources used

We listed every event from the six sources below, not excluding any element before the analysis phase.

2.1.1 English Wikipedia

The anglophone version of the collaborative encyclopedia Wikipedia offers a crowd-sourced list of fatal human crowd crushes [2]. Its criteria for inclusion are clear despite being non explicit: at least one death and being labelled as a crowd crush, a stampede, or a crowd-related structure collapse.

The list started in December 2005 as a list of examples in the Wikipedia article 'Stampede' (created itself in 2004). In December 2013, the list became a separate article called 'List of human stampedes', renamed 'List of human crushes' in October 2015, renamed back in September 2016, renamed 'List of human stampedes and crushes' in November 2019, and finally renamed 'List of fatal crowd crushes' in November 2022. After this last move, 11 non-fatal crushes were removed by a contributor. Before this removal, some events had also been removed, principally because they were non-fatal or because they needed a citation.

Wikipedia contributors use talk pages of the articles to discuss editorial matters, such as the criteria for inclusion and the vocabulary. Most of the discussions, both on the talk page of the list and the talk page of 'Stampede', focus on the vocabulary used, with many users asking for the word 'crush' to be used instead of 'stampede', reflecting a consensus in the academic community. This finally led the list to be named 'List of fatal crowd crushes', correctly reflecting its content, while the more general article to which the list is associated is still named 'Stampede' as of July 2023.

The statistics of both pages, available on xtools ([List](#), [Stampede](#)), show a very progressive construction of list, with events being added one by one (either because they just

occurred or were found in archives). These additions are made by many users throughout the years, making it an especially collaborative piece (in contrast to other Wikipedia pages that have been mainly written by one or a few contributors, despite the collaborative nature of the project).

The Wikipedia list references 191 events, ranging from year 150 to 2022 (we used a version of the page from very early 2023). As all Wikipedia content, it is reusable under the CC-BY-SA 4.0 free license.

We did not use lists from other linguistic versions of Wikipedia, despite the fact they list some events that are not covered by the English one.

2.1.2 Wikidata

Like Wikipedia, Wikidata is a Wikimedia project, maintained thanks to crowd-sourcing. It is also a semantic database, easily usable automatically or semi-automatically, which is notably done on certain Wikipedia articles. Its items are individually managed, with common properties and constraints. Each crowd crush has its own item, described with properties such as 'nature of element', 'date', 'location', 'number of deaths', or 'cause'.

Unlike Wikipedia, there is no curation of a list. Each item is managed individually; in order to make a list of crowd crushes, we have to query all the relevant item in the data base using their properties. We used Wikidata query service to extract all the items having the item 'stampede' (Qid Q2165983) or one of its sub-classes as their nature of element. The query we used is available [on GitHub](#).

Thanks to that query, we can extract a list of 84 items with their associated properties, with dates ranging from 1807 to 2022 [3]. It is the smallest source by the number of listed crowd crushes, but it is the only one with semantic data for these events. It is however the only data base available under the CC0 license, the closest one to the public domain.

2.1.3 Keith Still website

Prof. Dr. Keith Still, who is a renowned specialist of the domain, maintains a list of crowd disasters on his personal website [4]. The list generally does not provide its sources but includes large extracts from press article about the events.

The list is copyrighted but is open to crowd-sourced contributions, with at least one eye witness testifying.

210 events belong to the list, with dates ranging from 1902 to 2019.

2.1.4 World Crowd Disasters Web App V1

This web app made by Dr. Ali Asgary includes a map and its associated database [5]. It looks primarily based on the list by Keith Still but contains many other events, notably in the most recent years and in the global South.

The data table contains structured or semi-structured data, including the number of deaths and injuries, the country and the location, a description of the cause, coordinates,

a link to a reference, and the type of the event (Celebration, Educational, Entertainment, Music Festival, Office, Political, Religious, Shopping, Sport, Transportation and Other).

The license is not stated. The data base contains 255 events, with dates ranging from 1952 to 2022.

2.1.5 Working with Crowds website

This website does not state precisely its publisher (one individual based in the UK) nor its methodology [6]. However it contains an extensive list of 'crowds disasters and incidents', generally accompanied by one or more references. This list contains crowd crushes that are not listed in any of the other sources. The website also proposes a list of near-miss accidents, of which we recognise the value for future research, but that we did not use in the current study.

The license is not stated. The website lists 319 events (the longest list of the six sources) ranging from year 27 to 2022.

2.1.6 Introduction to Crowd Management book

We used the list published in the Introduction to Crowd Management book [7], that has also been published as an open dataset in [8] (CC BY SA 4.0 license) and used for the analysis in an article about trends in crowd accidents [1]. This very complete list has been elaborated thanks to a systematic press analysis, academic literature review, use of Wikipedia lists and the list by Keith Still, and finally some court documents. To be listed, each event should have sufficient information on it in one of the language spoken by the authors (English, Chinese, French, German, Italian, Japanese and Spanish), be crowd-related (excluding direct violence by members of the crowd, deaths mainly due to toxic smoke inhalation, etc.) and having caused at least one fatality or ten injuries.

The data set is recent, provides references for all the crowd crushes, and is based on two lists that we also analyse in the current paper. Therefore, we consider it the reference data set. It provides a very complete list of 277 events, ranging from 1900 to 2019, while the dataset [8] contains 281 events, which could mean the work has been completed by the authors.

2.2 Method of categorisation

While not knowing this work at the time where we did ours, we used very similar criteria to [1] but we kept the elements in the data set, putting them in the category 'Not a crowd' if they were in one of the following cases:

- event that is not crowd-related, deaths and injuries have not been caused by the crowd dynamics (car or plane accident, terror attack, sole person falling, weather-related structure collapses, etc.)
- structure collapse that has been caused by misconception (if the collapse has been caused by overcrowding we consider it a crowd related event)

- fire where smoke and heat were the only responsible of the deaths and injuries, when we could determine that from the sources.

Out of the 545 events, 71 were assigned to this category (13%, only from the sources by Keith Still, the World Crowd Disasters Web App and the Working with Crowds website). The method by the authors of [1] and the crowd-sourcing process on Wikidata and Wikipedia prevented any non-crush event to be included in these sources. The elements that we categorized as 'Not a crush' are excluded of all the analyses and graphics in the rest of this article, except Fig. 1 and Fig. 3.

For the other events, we placed them in 5 categories:

- 'Structure collapse' if the collapse leading to the fatalities or injuries had been caused or may have been caused by overcrowding
- 'Evacuation crush' if a crush occurred during an evacuation (from a fire but also from a riot, or from a false fire alarm)
- 'Interior crowd crush' if a crush occurred in a closed environment (building with a roof, train station...)
- 'Exterior crowd crush' if a crush occurred outside (including stadiums).
- 'Unknown' if it was impossible to know the environment, the nature of the crush, or if it was unsure that a crush occurred (notably in the case of fires, where a crowd crush is likely but not systematic).

Two events were also categorised as 'Collapse/Evacuation crush' as in both cases the collapse may have been triggered by the evacuation itself and not by simple overcrowding.

These categories are not meant to represent assertions on causality of the events, and have a limited analysis value. They are meant to give a first overview of the type of events, and be able to understand the limits of the statistics (typically, we should analyse events from the 'Evacuation crush' category separately as the number of deaths is the total number from the event (including people intoxicated by the smokes for instance) and not the number of deaths directly caused by the crush. It is also useful to distinguish structure collapses as they are deadly events that are not only caused by deficient crowd management, but also potentially by a deficient structure. The distinction between interior and exterior crowd crushes is the less strong, especially for stadiums and other semi-covered spaces, however it can be useful to specifically study one or the other category, as crowd management techniques can be quite different in open and (semi-)closed spaces.

2.3 Data collection

For each event, we collected the most recent precise number of deaths, or the most recent interval. For graphics and calculation, we retained the highest number in the interval, or the highest precise number (if the most recent source stated 'at least 70 deaths', we retained the number 70 for calculation).

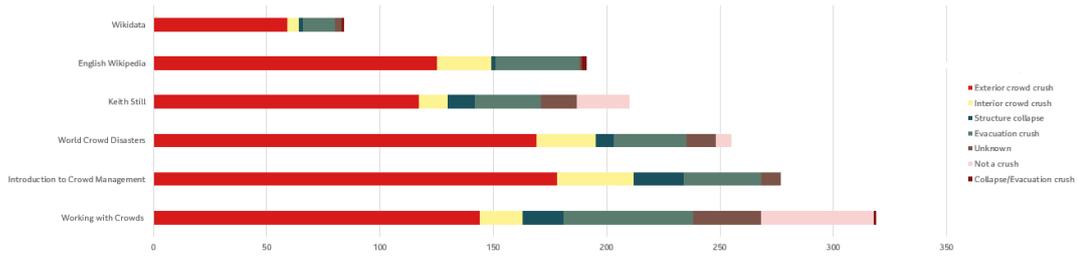


Figure 1 Number of events per category for each of the 6 sources (from year 27 to 2022)

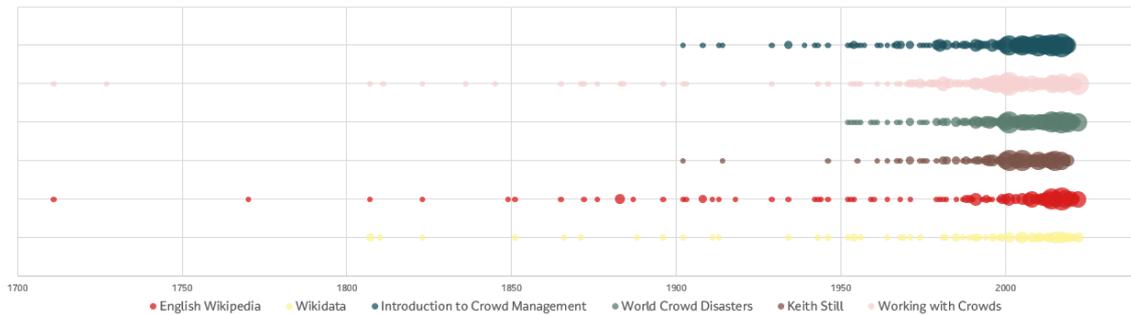


Figure 2 Time repartition of the events per source

We listed the current country of the location where the crowd crush took place, wrote a short description, and listed the number of references available for the event on English Wikipedia and in Working with Crowds. We published the final list on [9].

3 Results and discussion

3.1 No source is totally complete

As we can see in Fig. 1, sources have very different numbers of events. Wikidata is the less complete source, with 84 events. Working the Crowds looks the more complete source, with 319; however, we classified only 269 of them as crowd crushes, so with 277 events (all of them being crowd crushes), Introduction to Crowd Management is the largest source.

It is even more true when we look at the period covered by the book, from 1900 to 2019, in Fig. 3. Despite starting only in 1952, the World Crowd Disasters Web App has a very good coverage, with 234 crowd crushes (compared with 277 for the Introduction to Crowd Management). Despite being crowd-sourced, Wikipedia lacks coverage compared to the other sources; it stills offers a list of 157 events.

In Fig. 2, it is visible that the time repartition of the events listed by source is relatively similar, although there are some peaks in some sources that are not visible in the others.

Despite some sources being more complete than others, no source has a complete coverage of all the crowd crushes listed in the compilation of the 6 sources. As it is visible in Tab. 1, Introduction to Crowd Management and Working With Crowds have the best

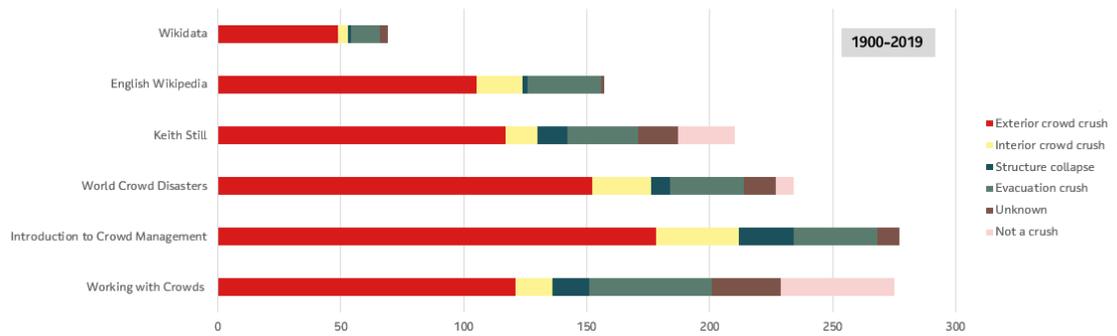


Figure 3 Number of events per category from year 1900 to 2019

Source	Coverage 27-2022	Coverage 1900-2019
Working with Crowds	59%	57%
Introduction to Crowd Management	51%	57%
World Crowd Disasters	47%	49%
Keith Still	39%	44%
English Wikipedia	35%	33%
Wikidata	15%	14%

Table 1 Coverage of the crowd crashes by source (percentage of the compiled list of crowd crashes covered by the source)

coverage for the period 1900-2019 with 57%, which is still relatively low. When we look at the global picture (from year 27 to 2022), Working With Crowds also has a coverage of 57%. This source, Wikipedia and Wikidata have a better coverage globally than for the period 1900-2019, especially for Wikipedia, which could mean there is no bias in favour of the recent period in crowdsourced sources.

Each of the sources has its own added value: even Wikidata contains crowd crashes that are not listed in any of the other sources (mainly for events that have an article in another linguistic version of Wikipedia than the English one). Tab. 2 presents the number of crowd crashes that are only listed in one of the sources. Working With Crowds and Introduction to Crowd Management are the biggest original contributors to the list.

Some cases are particularly well-covered, with 38 events being described by all the six

Period	English Wikipedia	Keith Still	Working with crowds	World Crowd Disasters	Introduction to Crowd Management	Wikidata
Until 1899	6		7			4
1900-2019	4	11	61	11	59	3
All (until 2022)	10	11	74	14	59	7

Table 2 For each source, number of crowd crashes that have only been described by this source

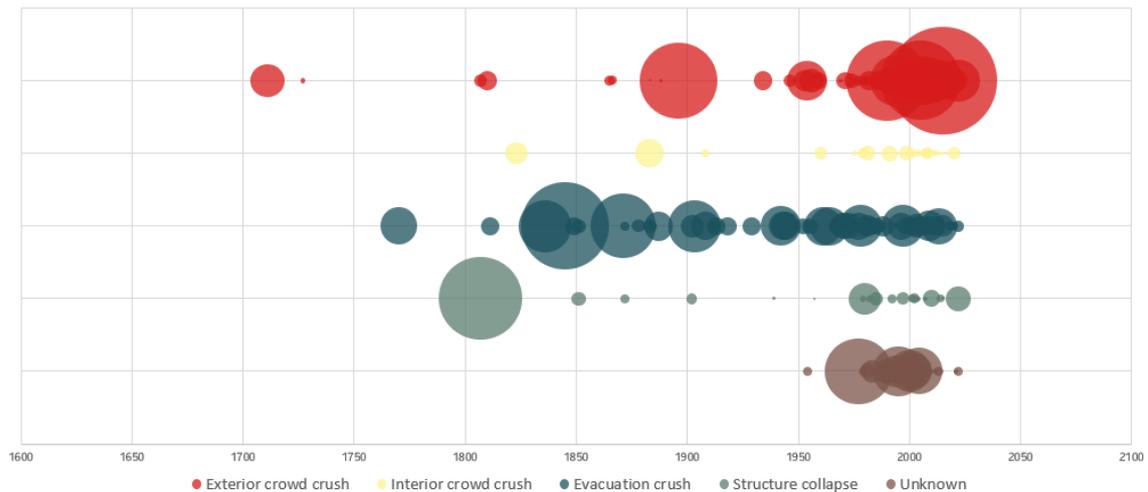


Figure 4 Number of deaths per year and per category

sources. These events have generally an excellent press coverage and sometimes reports of their causes.

3.2 There is no clear pattern of evolution by category

The number of deaths per year and per category is visible on Fig. 4. As we said previously, these categories are not perfect and do not presume a causality. However, we can still observe the importance that a single collapse event can have in this graph, such as the Eitai Bridge collapse in 1810; this is an additional reason to be careful with trends statistics. Additionally, we can observe that the number of deaths linked to evacuation processes tends to decline, which could be linked to better fire norms. On the opposite, the number of deaths in exterior crowd crushes tends to increase, but this rise could probably be explained by only the Hajj crowd crushes.

Globally, we prefer to be careful and assume there is no clear visible pattern by category in our statistics. Not only the categories could be refined, but an approach by causality and context, such as developed in [7], would be more useful in analysing the trends.

3.3 We may have a very partial view of crowd crushes

As we can see with its coverage, even a source like Introduction to Crowd Management, using a systematic approach of press articles plus existing sources has a relatively low coverage of 58% in its period. Authors of this source state that 'it probably contains all accidents which are accessible through the internet using mostly English as the main working language', however our analysis cannot confirm this assertion. We even think that our own consolidated list do not contain all these accidents accessible in English language, and is only a small subset of all the crowd accidents that occurred.

This can be explained by a succession of possible events after an accident:

- The crowd crush may have not been reported to the authorities nor the press.
- The crowd crush may have been known but considered as relatively normal and then not reported (especially with smaller events with only injuries).
- The event may have been hidden by the authorities (especially by authoritarian regimes) or not covered by the press due to insufficient means.
- The publication may have only done in a local language, and therefore not available in English or an other dominant language
- The publication may have been done only on paper, and not having digitised (or being only available as an image, or in a pay-walled database)
- The vocabulary used may be uncommon, different than the 'stampede' incorrectly but generally used, making it difficult to run an automated detection.

As the appendix C in [1] explains, when press sources are available, a number of biases still persist:

- The number of deaths for a single event may be different depending to the source (authorities, hospitals, newspapers, witnesses...)
- The definition of death (linked to an accident) is different depending in the location.
- There are biases for injuries and crowd size (that we did not collect in our own data set).
- Some accidents, such as football stadiums crushes in the UK or school accidents in China, may be over-reported compared to the others, being already the focus of press and sometimes researchers.

The latest point suggests that we still have a very partial view of crowd crushes: if some type of events can be over-represented in some countries, it means that these events are underrepresented in others, which means we could find more about similar events in future research.

For most of the cases, the press description of the events is concise and not precise. It often relies on a few first-and testimonies (generally having a very incomplete view of the events) and first findings by medical and police services. If the causes are listed, they can be different, are never analysed as a causal tree. Court causes could be a great source to help determine this causal tree, but they are rarely available, even if some court decisions are covered by short press reports.

This analysis confirms there is a big potential for analysis of past crowd crushes, using the methods of historians, accident specialists, jurists, and potentially modelling some of these accidents.

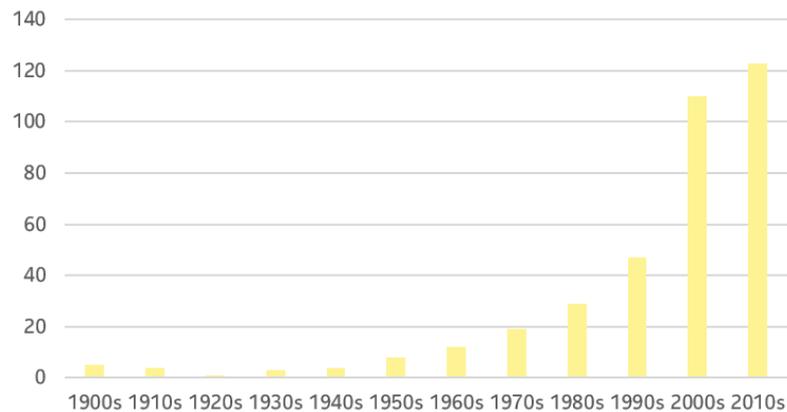


Figure 5 Number of crowd crushes by decade between 1900 and 2019

3.4 Crowd crushes may not be a growing tendency in regard to global population

In this section, we reproduce some of the analysis made in [1] on the 1900-2019 period, notably on the number of crowd disasters and crowd deaths. We excluded the events categorised as 'Not a crush' but also as 'Unknown' from this category.

In Fig. 5, we show the number of crowd crushes by decade between 1900 and 2019. Compared to [1], the number of events is higher in each decade, but the trend is very similar to the one observed in [1], with a progressive rise that suddenly accelerates in the 2000s, potentially with the increased accessibility of information.

In Fig. 6, we show the number of deaths due to these crowd crushes. We can see a big spike in the 1910s, linked to very deadly fires (Iroquois Theater Fire, Collinwood school in 1908, both in the USA, with respectively 602 and 175 deaths). In the following decades, both the number of deaths and events are very low, which could also be linked to an increased propaganda during and after WW1 (with some deadly events being hidden). The number of deaths is then higher but stable in the 1940s to the 1980s, then knows a boom from the 1990s.

This global shape looks relatively similar to [1], but we still have notable differences. The spike in the 1910s only appears in our data set, as the two deadly fires are not listed in the data set used in [1]. Similarly, many accidents are not listed in the 1960s to the 1980s, so we can observe a decrease of the deaths in [1], while this number of fatalities is stable in our data. In both data sets, the 1990s have a lower number of events that the 2000s and 2010s but are almost as deadly: this may be the rise of a good information availability about very deadly events, before a generalisation of information about crowd crushes.

We used the World Population Data from the UN to compute the number of deaths linked to crowd crushes per billion capita per decade. Results are shown in Fig. 7. It shows that in the recent decades, the phenomena of crowd crushes may not be growing faster than the world population. The lower rates from the 1940s to the 1980s and especially from the 1910s to the 1930s could be explained by under reporting and lack of availability of the information.

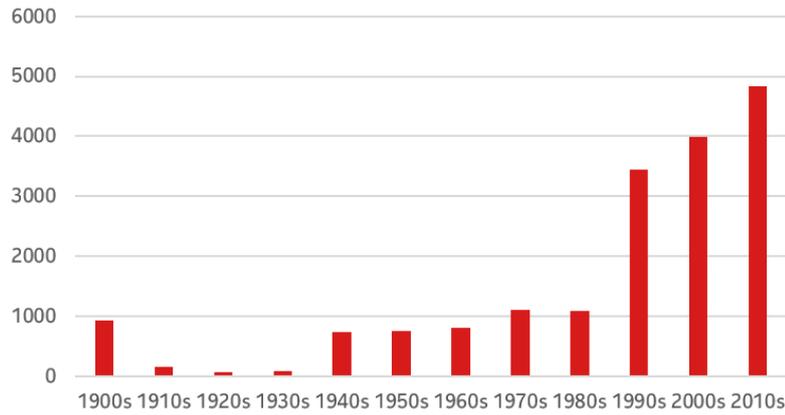


Figure 6 Number of deaths due to crowd crushes by decade between 1900 and 2019

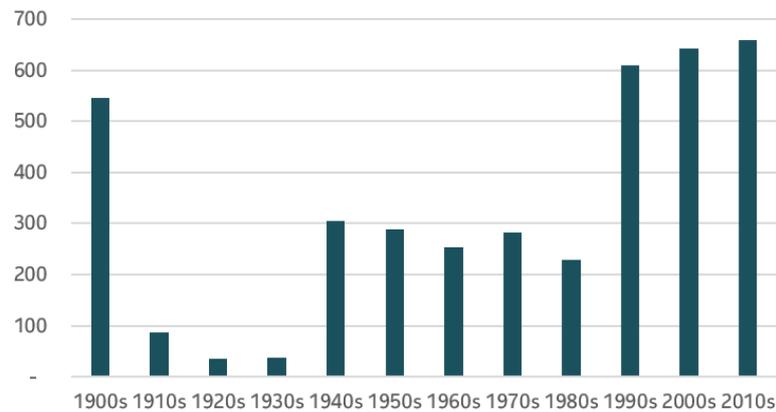


Figure 7 Deaths per billion capita due to crowd crushes by decade between 1900 and 2019

However, as we stated previously, we may have a very partial view of the crowd crushed, even in the recent years. We then prefer to be very cautious and not to conclude about a possible tendency about crowd crushes. We then prefer not to try to replicates the correlations made in [1], as we think it would be better to build an even more complete data base before.

4 Future work

We found in this paper that it was likely that the data about crowd crushes was still very incomplete, despite a huge effort from individuals, researchers and online communities. Compared to the merged data base, each source had a very incomplete view of the crowd crushes. Still, each source had new information that could have been shared broadly.

We then want to suggest to the community the use of Wikidata to compile information about crowd crushes. Despite being the least complete source of the six considered, Wikidata can be easily improved by using data of the other sources. It offers a structure that

makes it easy to gather the most possibly complete information about the crowd crushes, offering the possibility to precise the number of deaths and injuries for each source, but also the gender divide of these deaths, the causality, etc. Wikidata also offers the possibility to run queries to extract this data easily and visualise it.

In a future work, we will start this completion work on Wikidata, as a part of a Rail Crowds Observatory. We aim to better prevent crowd-related accidents in transportation context by understanding the causes of crowd accidents and possibilities to mitigate them. One important part will also be to look at near-misses and not only accidents, as those could also teach important lessons for prevention.

5 Conclusion

By compiling every event from six existing sources of crowd-related accidents, we showed that none of these sources had a global coverage of these events. Nevertheless, each of these sources contained information about crowd crushes that were contained in none of the other sources. Therefore, we recognise the value of each of these sources but also the interest in comparing them and building a more complete source.

The fact that any source, even those built with the willingness of being complete, fails to do so, suggests that even with this compilation, we still have a very partial view of crowd crushes. This may be linked with linguistic, reporting and information availability biases.

Even with this merged data, we cannot prove that the number of crowd crushes per capita is stable nor decreasing, but we can suggest that it is not greatly increasing.

The study of past crowd crushes is still a very worthy subjects, for which we may need new methods and supplementary means.

Acknowledgements We would like to thank Rémi Coulaud (SNCF Transilien) and Antoine Leroux (AREP Flux et Mobilités) for their contribution to the Rail Crowds Observatory project.

We also would like to thank AREP for funding this research, especially Matthieu Goudeau (AREP Flux et Mobilités) and Hiba Debouk (AREP Territoires).

Our very warm thanks go to the reviewers for their precious and valuable feedback.

We finally would like to thank all the individuals, resarchers, institutions and online communities creating, maintaining and supporting the precious sources used in this paper.

Ethics Statement Competing interests: Capucine-Marin Dubroca-Voisin is the volunteer chairperson of Wikimédia France (from 2021) and the treasurer of Wikimedia Europe (from 2022). Both organisations support the Wikimedia communities and projects, including Wikipedia and Wikidata.

Author Contributions Capucine-Marin Dubroca-Voisin: Conceptualization, Data Curation, Formal Analysis, Methodology, Visualization, Writing - Original Draft

References

- [1] Feliciani, C., Corbetta, A., Haghani, M., Nishinari, K.: Trends in crowd accidents based on an analysis of press reports. *Safety Science* **164**, 106174 (2023). doi:10.1016/j.ssci.2023.106174
- [2] Wikipedia contributors: List of fatal crowd crushes (2023). URL https://en.wikipedia.org/w/index.php?title=List_of_fatal_crowd_crushes&oldid=1133987934. Page Version ID: 1133987934
- [3] Wikidata contributors: Stampedes repertoried on Wikidata (2023). URL <https://github.com/Mind-the-Cap/Observatory/tree/main/Wikidata>
- [4] Still, G.K.: Crowd Disasters | Prof. Dr. G. Keith Still (2019). URL <https://www.gkstill.com/ExpertWitness/CrowdDisasters.html>
- [5] Asgary, A.: World Crowd Disasters Web App V1 (2022). URL <https://yorku.maps.arcgis.com/apps/webappviewer/index.html?id=e7c52856187642e19bd227865393432c>
- [6] Working With Crowds (2022). URL <https://www.workingwithcrowds.com/crowd-safety-home/>
- [7] Feliciani, C., Shimura, K., Nishinari, K.: Introduction to Crowd Management Managing Crowds in the Digital Era: Theory and Practice. Springer International Publishing AG, Cham (2022). doi:10.1007/978-3-030-90012-0. OCLC: 1305842513
- [8] Feliciani, C.: List of crowd accidents from 1900 to 2019 (2023). doi:10.5281/zenodo.7523480
- [9] Dubroca-Voisin, C.M.: List of past crowd crushes (2023). doi:10.5281/zenodo.8186684