ALLIANCE FOR INNOVATION AND INFRASTRUCTURE

Improving Upon Our Dig Laws:

Why Data Must Take Center Stage to Reform Damage Prevention

January 2023

Executive Summary

The latest Damage Information Reporting Tool (DIRT) Report for 2021 was released in the past fall.[1] The Report, prepared by the Common Ground Alliance (CGA), is a continuation of 18 consecutive annual reports. Though imperfect, the DIRT Report remains the only source for comprehensive data and trends on excavation damage to underground facilities across the country. This year's report stands out in that the Report was based on significantly less reported data than in previous years.

Despite the dearth of reporting, from the data that is available, CGA notes that "damages have plateaued or slightly increased" in recent years. The Report only uses a three-year model, which sets 2019 as the baseline year. If damages are higher than 2019 – the highest estimated total on record – that would mean **excavation damage in the United States reached a new all-time high in 2021**. This is an issue not specifically pointed out by CGA in the Report.

The Report shows a precipitous decline in damage events **reported** to DIRT for 2021, falling to the **lowest level since 2013**. This is not the result of a significant reduction in the number of damages that occurred but appears to be the result of a decision of some industry participants that previously reported damages to stop doing so in 2021. The drop off in reporting was so significant that CGA was **unable to determine an estimated total damage level** for the first time since 2006.

U.S. and Canada	**2018**	**2019**	**2020**	**2021**
Total Events entered into DIRT	440,749	534,151	475,770	**228,393**
Damages (unique events)	341,609	453,766	385,381	**203,618**
Total Estimated Damages	509,000	532,000	468,000	–

While not producing a total damage estimate, CGA did still evaluate a three-year trend. To compensate for the lack of data, CGA was required to form a slimmed down "comparable dataset" for its trend analysis. This approach removes nearly 70 percent of the reported data from 2019 and 2020 in order to match them with the reduced 2021 reporting figures. This dataset therefore includes only entities that submitted reports in all three years. This also led to the creation of a new figure for damages per 1,000 One-Call Transmissions that is no longer comparable to previous DIRT Reports.

Whereas previous years reported a ratio of *estimated total damages* to total transmissions, CGA now uses a ratio of *unique reported damages* to total transmissions. This sets up a confusing measure of the ratio of damages to transmissions relative to previous reports, making comparisons beyond the three presented years impossible. To see how significantly this contorts the statistic, we display older data from the 2020 Report next to new data for 2021.[2] Critically,

the new ratio for 2021 cannot be used to compare with the previous reports, but displaying the new ratio alongside the old illustrates the difference based on how the ratios are calculated.

U.S. only	**2018**	**2019**	**2020**		**2021**
Total Estimated One Call Transmissions	244.3 M	267.6 M	273.9 M		**288.3 M**
Damages per 1,000 One Call Transmissions	2.08	1.99	1.71	⟸ Calculated using *Total Estimated Damages* Calculated using *Unique Reported Damages* ⟹	**0.570**

Given the reduced reporting in 2021, there is a question about whether excavators, locators, facility operators, and other participants in the damage prevention process themselves value the DIRT Report. However, even setting aside the reduced data, the Report does indicate that year over year, damages continue to increase, and the root causes of those damages have generally remained the same over time.

In addition to our discussion regarding the quantity of data provided in 2021 we discuss the quality of the data submitted to CGA and the implications of less data being reported on data quality.

Introduction

We have focused on two areas for discussion with respect to this year's DIRT Report: data quantity and data quality. While these topics are not new, and in fact have been a concern of CGA for years, they present themselves as more significant than ever this year.

As background to our discussion of data quantity and quality, it appears that damages to underground facilities may continue to be increasing. The Report includes metrics that support this conclusion, such as a one percent year-over-year increase in damages per 1,000 transmissions and a three percent increase in damages relative to construction spending. Both ratios are also higher than 2019, which was this Report's baseline year.

This is disappointing primarily because we believe the technology and best practices are available to address this trend of increasing damages, yet for the better part of a decade, excavators, locators, facility operators, one-call centers, policymakers, and observers have seemed to simply point to rising damage numbers and call for the voluntary adherence to new and existing best practices. Given the year-over-year trend, this approach seems to have little effect on reducing facility damages. Over the years, there has been little discussion of new regulatory requirements and how they may be beneficial in reversing the current trend. This is understandable with participants preferring to self-regulate. Clearly there has not been enough of a change in practices to have a positive impact on the damage trend despite new best practice initiatives and available new technology[3] discussed in prior DIRT Reports.

Our first point of discussion is the reduction in the volume of data available to analyze for 2021. Given the voluntary nature of reporting, both quantity and quality have always been difficult to

guarantee despite the emphasis on their importance by CGA. With few exceptions, prior to the current year's Report, every year saw more records submitted to DIRT than the previous year. What is unprecedented is that the level of reporting declined by over half in 2021. Not only did over 250,000 expected event reports fail to be reported in the DIRT platform, but CGA received the lowest level of event reporting since 2013.

It appears that one or more stakeholders with significant market share or notable presence decided not to submit data.[4] For a member organization that presents itself as offering processes to reduce damages, this does not bode well if CGA cannot get their own members to report the data that the Report has historically stated is necessary to improve damage results. Additionally, this certainly raises concern around the effectiveness of improvement through voluntary self-regulation and implementation of best practices.

There is little discussion in the Report of what happened this year that caused a drastic reduction of reported data. In fact, if readers do not have previous DIRT Reports open alongside the 2021 Report, they would have no way of knowing the magnitude and significance of the reduced reporting. CGA merely presents the numbers, without providing readers the context that they were significantly lower than reported in previous years, instead making vague statements about "fluctuation" or "change in the makeup" of the datasets.

Following our concern regarding the lack of data, is a concern about data quality. This remains a struggle, with no clear pathway for how it will be improved. For years, CGA has focused on data quality and the need for quality data to be able to drill down to the specific root causes of damages. We believe CGA's desire for data quality to be earnest, but we would hope to see more action and greater use of CGA's influence to bring it about, rather than continually recommending better reporting. Perhaps due to contractual and legal concerns, getting data that allows the DIRT Report to pinpoint root cause issues is a pipedream.

Data Quantity

In statistical analysis, more data typically provides higher confidence and better clarity. Of course, the quality of the data is integral to this, but when more data is available, analysts are better able to filter or remove bad data. That is a luxury afforded by a robust dataset; but when less data is available, there can be less confidence in the inferences that can be made.

> "Voluntary data submission allows for a change in the makeup of the dataset year over year and can complicate annual trending."
> -CGA

The DIRT Report for 2021 takes a remarkable step backward with a significantly lower level of data compared to prior years. Curiously, while CGA recommends multiple times throughout its Report that data *quality* is important, in the first year since 2012[5] to see less reporting than the previous year, there is virtually no meaningful mention of the importance of data *quantity*.[6] The level of data reported to DIRT fell by 52% percent year-over-year and by 57 percent since 2019. This merits more than a passing mention in the Report, but a major call out and an explanation by CGA steps they are taking to increase, or at a minimum *resume*,[7,8] reporting.

DIRT Data[9]

	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Total Reported Events	232,717	224,616	273,599	363,176	390,366	411,867	440,749	534,151	475,770	228,393
Change from previous year reporting	…	-3.48%	21.81%	32.74%	7.49%	5.51%	7.01%	21.19%	-10.93%	**-51.99%**
Estimated Total Damage	350,000	335,000	349,000	378,000	416,000	439,000	509,000	532,000	468,000	–

With the lowest level of reporting in eight years, CGA declined to offer an estimated total damage number. Not since 2006 has the overall level of data been such that CGA did not produce such an estimate. The current Report does indicate that damages may be moving upward, but that finding is largely relegated to the Report's Appendix B and couched in technical statistical language.

The hidden story here is that – according to CGA's own claim that "Statistical models used for three-year trend analysis point to an overall plateau or slight increase in damages since 2019"[10] – **damages may be at the highest level ever recorded**. Readers would not know this from the Report, which uses only a three-year analysis. Unstated is that the baseline year (2019) for this report was the highest year for damage since CGA began producing damage totals and estimates.[11]

In 2019, estimated total damages peaked at 532,000. The same year, CGA estimated annual economic harm from damage cost the U.S. economy over $30 billion. In 2020, economic disruptions in response to COVID-19 diminished both construction activity and excavation damage numbers slightly, but this was by all accounts an aberration. Now, CGA notes that "there is some weak evidence that counts in 2021 differ (upward) significantly[12] from those in 2019 after accounting for key driving factors."[13] If true, it is an unprecedented new high in excavation damages to underground facilities and costs, yet the DIRT Report does not point out this key context. Moreover, as the Report cautions, we cannot know this with great confidence because of the limited data. Using only a three-year trend is already limiting, but starting a trend analysis at a peak and not making that clear to readers is misleading. Even if it were true that damages had plateaued, that means they have leveled off at a nationwide historical high.

To address the limited reported data, CGA created a new "comparable dataset." To build this new dataset, CGA revisited the previous two years and stripped out data from reporters who did not submit damage reports in 2021 (so that data for 2019, 2020, and 2021 all featured consistent reporters). The number of reported events deleted to create this new comparable dataset appear to be 222,329 for 2020 and 296,763 for 2019. In other words, more data was *deleted* for each of the previous years to create the comparable dataset than was *submitted* in 2021.

An implication of the lower data level is the generation of a new damages per 1,000 transmissions ratio. In previous years, **total estimated damages** was divided by **total transmissions**. This produced a ratio of between 1.5 and 2.0 damages per 1,000 transmissions. For 2021, with no estimated total damages to use, the Report uses **unique reported damages** divided by **total transmissions** to produce a statistic in the range of 0.5 damage per 1,000 transmissions. [14]

U.S. only	2018	2019	2020	2021
Total Estimated One Call Transmissions	244.3 M	267.6 M	273.9 M	288.3 M
Total Estimated Damages	509,000	532,000	468,000	
Total estimated damages per 1,000 One Call Transmissions (2020)	**2.08**	**1.99**	**1.71**	
Total Damage Number used to produce ratio[15]		149,588	154,753	164,331
Damages per 1,000 One Call Transmissions (2021 comparable dataset)		**0.559**	**0.565**	**0.570**

Obviously, this handicaps Report-to-Report comparisons, and even though the comparable dataset revised the past two years into the same terms using fewer total reports, the statistic may now be misconstrued by those may believe it conveys progress rather than a new formula. *Functionally, by calling it the same thing, CGA has managed to reduce the damage ratio by having less reporting.*

Any reader familiar with the previous estimation may be confused with what to make of the new ratio. Worst of all, there does not appear to be an acknowledgement of this difference, and the Report merely relies on telling the reader that the comparable dataset is used. When it comes to this statistic, CGA reports it as "the standard DIRT metrics" being reported. This is odd, if not misleading, and contributes to the confusion that stems from this year's Report.

Another curiosity is that the damage to one-call transmission ratio was declining from 2018 to 2020 despite CGA reporting that damages were increasing. With the new comparable dataset, the damage to one-call transmission ratio is now rising year over year. This inconsistency adds to confusion and readers cannot easily discern whether the new ratio is more trustworthy or valuable than the previously reported ratio, nor whether the damages per 1,000 one-call transmissions is a worthwhile metric to assess in the first place.

The comparable dataset also led to the creation of a new statistic for damages per million dollars of construction spending.[16] With the lower level of data reported to DIRT, CGA had to reform its datasets and revise standard ratios they have reported for years. As a stand-alone report, the 2021 DIRT Report does present a multi-year trend with a comparable dataset.[17] However, it violates a basic tenet of statistics by starting a trend at a peak, it fails to disclose or comment on that and other notable issues, and the Report cannot be compared side-by-side with any previous DIRT

Reports because the new metrics contort the long-reported ratios like damage per 1,000 transmissions.

While it is confusing that that the metric changes from year to year, the data nonetheless shows that the trend is the same – **damages have continued to increase**. Maybe the real conclusion to be drawn here is the lack of influence that CGA is able to exert on its members. If they cannot even get members to report data and, after years of pursuing better data, are not able to improve data quality, is it pointless to think that industry participants will voluntarily adopt new technology and best practices promulgated by CGA in an effort to reduce damages?

Data Quality

In 2006, CGA recommended the creation of a data quality index (DQI). The index was adopted in 2007 and has been used to analyze data quality in each of the 14 reports since. In 2007 the weighted DQI was 50 percent. No year has broken the ideal 70 percent threshold.[18]

Data Quality Index[19]

	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Avg. DQI before 2021	50	57	59	57	~60	~60	64	65	67	67	63	62	59	-	
DQI from "comparable dataset"													67.6 *	66.5 *	65.6 *

The use of a new comparable dataset offers some insight into the quality of reporting. For 2019, the higher number of reports submitted resulted in an average DQI of 59, but when reducing that data set to only the consistent reporters for 2019 through 2021, the data quality improved. It seems then that the industry participants who did not submit data for 2021 provide lower quality of data and may have been dragging down the data quality.[20] Thus, the comparable dataset, then may overstate the quality of the data, because it simply removes poor quality reporters.

> "Improving the quality of DIRT reports would give the industry a much clearer picture of how and why damages occur." -CGA

Data quality has long plagued the preparation of the DIRT Report. Indeed, there is no shortage of emphasis by CGA that data quality improvements are needed.

> Improving the quality of damage prevention data comprises an entire category of recommendations in our 2021 DIRT Report and is echoed by many of our other programs and initiatives, including the industry-advancing work of CGA's Next Practices Initiative. As CGA's work with Next Practices, Best Practices, DIRT and our newest arm – the Damage Prevention Institute – highlights, the industry needs more high-quality data as well as data-driven methods for evaluating the effectiveness of education and training programs, the efficacy of practices such as electronic white-lining, and the impact of investments in GIS mapping and other technologies in order to take the next steps toward our long-term goal of zero damages.[21]

"Better data means we can more accurately identify areas in which we need to focus our work…"
-CGA

Paired with the consistent nature of broadly stated root causes, the stagnated quality of data should be construed as a major impediment to reducing damage. We take statements like, "the root causes of these incidents have remained remarkably consistent" to highlight a real problem. The very fact that data quality has not improved may have limited CGA's ability to see nuanced and detailed issues within those broad root cause categories. In other words, it is not good to have the same top root causes year after year because it means we have not been able to address them or reduce them. Better data will provide the roadmap to eliminating certain root cause.[22]

With better quality data, root causes can be evaluated with greater nuance so that public policy, technology, best practices, education, and training can all be calibrated more precisely to address the real underlying dynamics of the consistent root causes. We are not sure how CGA will be able to achieve better and more nuanced data if DIRT participants do not feel compelled to even contribute to the Report, as was the case for 2021.

Conclusion

The DIRT Report has the potential to be an invaluable resource for shedding light on causes of excavation damages, but it is only as valuable as the data it manages to collect and present. This year was an unmistakable step backward for the data and the Report. However, notwithstanding the dearth of data, from what we can discern it looks like facility excavation damages have increased, continuing a years-long upward trend.

Is the DIRT Report providing as much value to industry participants as it could, given that root causes have not changed over the years and industry participants do not seem compelled to supply necessary data? Further, previous years carried significant recommendations, and little progress seems to be occurring as damages continue to rise.

While recommendations this year seem to weakly ask members to improve their data quality, CGA left almost wholly unaddressed the major decline in reported data – will this continue or is it a one-year aberration?[23] The Report should have told readers up front that it is built on fewer than half the reports as previous years, noted how significantly that hamstrung analysis, and issued a forceful call to members and industry participants to submit damage reports next year.

This year's DIRT Report continues our concern that self-regulation and voluntary member organization standards may not be enough. Not only have CGA and the industry not counteracted rising damage trends, but now appear to be unable to adequately collect reliable data (struggling in both quantity and quality). If meaningful changes in damage prevention are to occur, it seems that they must be at least partially driven by incentive and enforcement from public bodies. Legislation or regulatory requirement may be needed as much for data reporting as for the implementation of recommended best practices and technology. If so, state-level policy is likely the best, with federal policy focusing most on incentives and on promoting consistency in state law and policy.

Questions

In lieu of recommendations, this paper concludes with questions that we believe need to be answered before the next DIRT Report. These questions are addressed to damage prevention stakeholders, the Common Ground Alliance, and public policymakers concerned with the damage prevention system.

- How can data quantity and quality be meaningfully improved and consistently reported?
    - Specifically, what requirements or best practices are needed to improve data reporting without sacrificing confidentiality, creating liability, or imposing disproportionate costs? Is it possible that given these constrains, better data quality is an unrealistic expectation?
    - At what point is it time for regulators to mandate reporting requirements at the state level?
- If limited quantity and quality of data continue to be submitted to DIRT, what is the value of the Report to the industry?
    - Will statistical models continue to extrapolate with low confidence? Will industry participants respect or derive value from such a report?
- When can the public expect to see the impact of recommendations and reforms proposed in the DIRT Reports?
    - The creation of five new CGA working groups in 2019 have been slowly reported on and with limited visibility, how can greater transparency be achieved?
    - Why are damage numbers rising even after reforms to best practices and reports from Next Practices?
    - Will the creation of the Damage Prevention Institute lead to reduction in damage numbers? On what time scale?

Appendix A

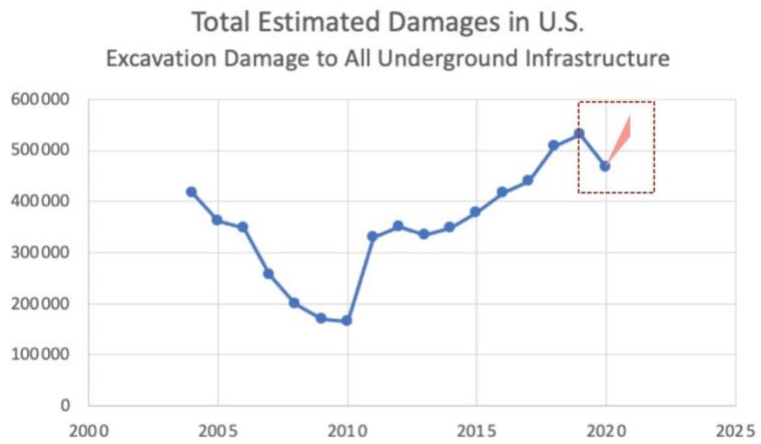


Figure 2: Data Reported to CGA DIRT

This figure is from the Alliance for Innovation and Infrastructure report, *Excavation Damage to Underground Infrastructure: A Look at the Federal Damage Prevention Approach*, April 2022. It has been updated to include a red box and cone of possible damage totals in line with CGA's conclusion that "damages have plateaued or slightly increased." We draw your attention to the overall trend and how the DIRT Report this year uses only a three-year trend with 2019 as its baseline.

The information in the box is not from CGA directly, but our framing to explain what CGA did present, namely a tight three-year trend beginning at the high-point in the overall damage trend and the statistical trend pointed out in the 2021 DIRT Report. To "plateau" means to be somewhere on the level of 2019 damages, while to "slightly increase" means to be above that level. Accordingly, the red cone presents our best approximation of visualizing the information CGA discusses.

Appendix B

Another metric this year is damages per million dollar of construction spending. In the same way, the paired down comparable dataset creates a statistic that is out of step with prior reports and clouds the issue. It is difficult to see the state of excavation damage in 2021 from this report, and the attempt to create new datasets for consistency end up creating inconsistency.

U.S. only	2018	2019	2020	**2021**
Construction Spending (millions of 2021 dollars)	1,462,365	1,489,721	1,576,142	**1,626,444**
Total Estimated Damages	509,000	532,000	468,000	
Total Estimated Damages per million dollars of construction spending (2020 dollars)	0.348	0.347	0.319	
Total Damage Number used to produce ratio[24]		148,972	154,461	**164,270**
Damage per dollar of Construction Spending (2021 comparable dataset)		.100	0.098	**0.101**

Citations and Notes

[1] Common Ground Alliance. (2022). *Damage Information Reporting Tool: 2021 Analysis & Recommendations.* Volume 18. Retrieved from https://commongroundalliance.com/Portals/0/DIRT%20Report%202021%20-%20FINAL1.pdf?ver=2022-11-30-165941-267.

[2] The 2021 DIRT Report recreated 2019 and 2020 data in proportion with 2021. We analyze the comparable dataset below.

[3] *See*, Common Ground Alliance. (2021). *Technology Report 2021. Technology Advancements & Gaps in Underground Safety*. p. 2. Retrieved from https://commongroundalliance.com/Portals/0/2021%20Technology%20Report.pdf?ver=2021-05-27-165320-157 ("What the Future Might Look Like: An Idealized Excavation Project in the Year 2030").

[4] There is a precedent for one stakeholder or group having a substantial impact on DIRT, such as a single stakeholder group (locators) increasing the reports submitted to DIRT by 100,000 in 2011.

[5] 2020 excluded due to obvious economic and governmental disruptions.

[6] The recommendation to "Improve data quality and reporting by industry" is focused on data quality and blanks left in reports, while "Increase DIRT reporting directly from excavators" is a narrow recommendation that merely points to excavators having high quality, but relatively low direct reporting. Nowhere does the Report call out the party or parties who seemingly left out over 200,000 reports in 2021, nor emphasize that data quantity is critical, instead using most references to data as an opportunity to promote quality (a needed, but secondary issue this year).

[7] In 2020, Locators accounted for 236,680 reports (68.4 percent) while in 2021 only account for 17,800 reports (8.9 percent). It is clear that the market leader in the locate industry, or a host of locating companies, stopped reporting.

[8] One interesting corner CGA seems to be in this year is that because of the action of one or more CGA members to stop reporting, CGA had to create a "comparable dataset." By definition, if that major entity resumes reporting in 2022, they will be forced out of the "comparable dataset" because they are not a consistent reporter. We are eager to

see how CGA chooses to handle this eventuality. If this is used again, to ensure data can be incorporated next year, a standard that "consistent" reporters are those submitting for at least three of the last four years seems called for.

[9] Data from DIRT Reports for 2012 through 2021.

[10] *Supra* note 1 at p. 4.

[11] *See* Appendix A for a graph.

[12] "Significant" in the statistical sense

[13] *Supra* note 1 at p 63.

[14] To produce the statistic from prior years: Total Estimated Damage / Total Transmissions * 1000. 468,000/273,900,000 * 1000 = 1.71.
To produce the statistic this year: X / Total Transmissions * 1000. X / 288,300,000 * 1000 = 0.57. X = 164,331. This number is lower than the unique reported events for 2021. So, while in previous years, the damage number used to produce "damages per 1,000 transmissions" was the estimated total damage extrapolated from data to account for unreported incidents that likely occurred – a much higher number than reported events. In 2021, CGA uses a number far lower unique reported events while still attempting to demonstrate "damages per 1,000 transmissions"

[15] It seems that the number used for this calculation is the "Reported Unique Damages (Comparable Dataset) from Table 7. However, the precise number is not given. The table produced here solves for the damage number algebraically.

[16] *See* Appendix B

[17] Although a three-year trend is somewhat minimal in comparison to the five-year trend presented in the 2020 DIRT Report.

[18] There is nothing inherently special about 70 percent besides providing a higher level of clarity. It is merely a yet-to-be achieved goal. In 2020, CGA explained in its 2019 DIRT Report that "DIRT could be greatly improved by raising the scores of those below 70…"

[19] Data sourced from DIRT Reports for 2007 through 2021.

[20] This points to a locator stakeholder, consistent with the 2020 DIRT Report finding that "The overall quality and completeness of data is not consistent across stakeholder groups. Of the top three sources of data (see Figure 1), natural gas and excavators have DQI (Data Quality Index) scores exceeding 80, while locators are at 49."

[21] *Supra* note 1 at p. 2.

[22] *See*, for example the same idea expressed by CGA in its 2020 DIRT Report: "better-quality data could improve the quality of the analysis…"

[23] And what will CGA do with the comparable dataset if reporting picks back up to a higher level?

[24] It seems that the number used for this calculation is the "Reported Unique Damages (Comparable Dataset) from Table 7. However, the precise number is not given. The table produced here solves for the damage number algebraically.

Author

Benjamin Dierker, JD, MPA
Director of *Public Policy,* Alliance for Innovation and Infrastructure

For more information or inquiries on this report, please contact the Aii Media Coordinator at comms@aii.org

Recommended Citation for this report

B. Dierker. (January, 2023). *Improving Upon Our Dig Laws: Why Data Must Take Center Stage to Reform Damage Prevention.* Alliance for Innovation and Infrastructure.

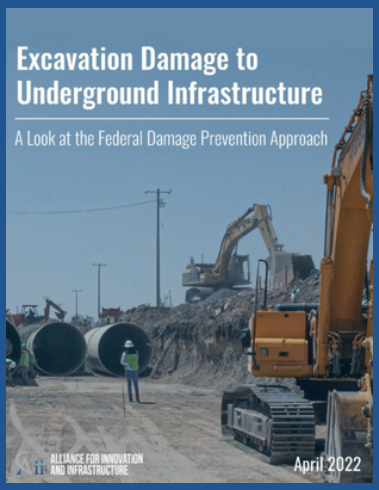
About Aii

The Alliance for Innovation and Infrastructure (Aii) is an independent, national research and educational organization that explores the intersection of economics, law, and public policy in the areas of climate, damage prevention, energy, infrastructure, innovation, technology, and transportation.

The Alliance is a think tank consisting of two non-profits: the National Infrastructure Safety Foundation (NISF), a 501(c)(4) social welfare organization, and the Public Institute for Facility Safety (PIFS), a 501(c)(3) educational organization. Both non-profits are legally governed by volunteer boards of directors. These work in conjunction with the Alliance's own volunteer Advisory Council.

Board of Directors

Related Works



The Alliance for Innovation and Infrastructure (Aii) is an independent, national research and educational organization. An innovative think tank, Aii explores the intersection of economics, law, and public policy in the areas of climate, damage prevention, energy, infrastructure, innovation, technology, and transportation.

ALLIANCE FOR INNOVATION AND INFRASTRUCTURE

Aii.org