

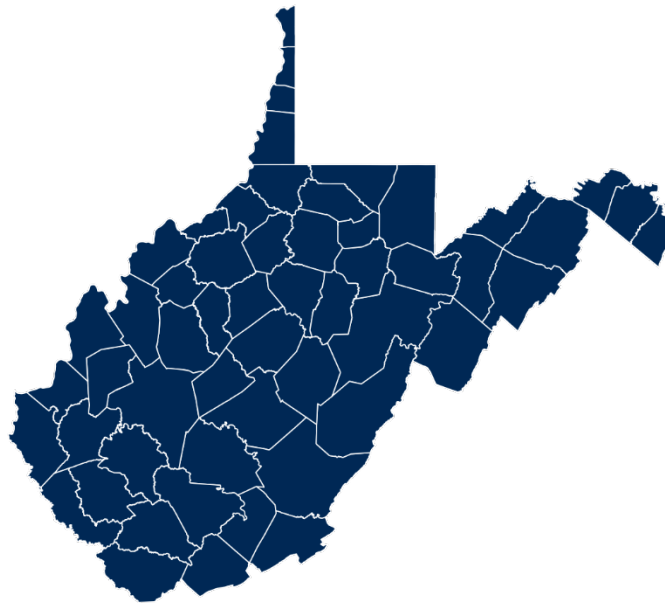
EDUCATION, DATA, HEALTH

IDEA Data Reporting in West Virginia

Challenges and Opportunities

Samuel Workman and Oreoluwa Runsewe

February 28, 2024



© 2024 – Institute for Policy Research and Public Affairs

📍 Contact Information:

Institute for Policy Research and Public Affairs
Rockefeller School of Policy and Politics
West Virginia University
1515 University Ave.
Woodburn Hall, 221D
PO Box 6286
Morgantown, WV 26506

📞 +01 (304) 293-9306

✉ policyresearch@mail.wvu.edu

🐦 [@WVUPolicyRes](https://twitter.com/WVUPolicyRes)

🌐 [wvu-policy-research](https://www.wvu-policy-research)

🌐 policyresearch.wvu.edu

1 Contents

- 1 Contents 3
- 2 List of Figures..... 4
- 3 Executive Summary..... 5
 - 3.1 Key Findings..... 5
 - 3.2 IPRPA Recommendations 6
 - 3.3 Building Data Infrastructure 6
- 4 What is IDEA? 6
 - 4.1 Public Resources..... 7
- 5 Understanding Data Organization and Curation 8
 - 5.1 Two Types of Data Curation 8
 - 5.2 Principles of Data Collection and Organization..... 9
- 6 West Virginia’s IDEA Data 10
 - 6.1 Data Delivery..... 11
 - 6.2 Delivery Recommendations 12
 - 6.3 Delivery Logistics..... 12
 - 6.4 Data Organization in Spreadsheets..... 13
 - 6.5 Data Curation 15
- 7 Overview of Relevant Federal Statutes..... 17
- 8 Institute for Policy Research and Public Affairs Partnerships 20
 - 8.1 How to Cite This Report..... 20
 - 8.2 Our Visualizations–Media 20
- 9 Acknowledgements and Disclaimers 20
- 10 References..... 21

2 List of Figures

Figure 1: IDEA Data in WVDE's Spreadsheet Form..... 13



3 Executive Summary

The *Institute for Policy Research and Public Affairs* at West Virginia University (WVU) completed a study of the quality and usability of West Virginia's IDEA data. The topic and question for the study concerned whether IDEA data were organized and delivered to inform the public sufficiently and to promote uptake of the data in scientific communities working on issues surrounding education for students with developmental disabilities. IDEA, like many administrative data sets, is optimized for accountability and not for use.

3.1 Key Findings

- The primary mode of delivery is a set of fragmented spreadsheets, which fully meet the requirements for reporting under federal statutes.

- The mode of delivery in spreadsheets is limiting for informing the broader public; spreadsheets generally are the medium of experts.
- Within the spreadsheets, the organization and curation of data do not meet agreed-upon standards for social scientific data.

3.2 IPRPA Recommendations

- The West Virginia Department of Education (WVDE) should, at a minimum, structure the spreadsheet data to be consistent with standard practices in the social sciences.
- **General Public:** Data tables for delivery to the general public should be in HTML, text, or PDF format, eliminating the need for navigating spreadsheets by the public, the non-profit sector, and private sector firms working on service delivery for these families. Tables embedded on the website for current years would be optimal.
- **Scientific/Expert Community:** Ideally, a master CSV or other text-compatible file would correct the fragmentation for delivery to scientific and other expert communities.

3.3 Building Data Infrastructure

These recommendations require more input and consultation but would revolutionize the IDEA data for uptake in relevant communities. The recommendations above are prerequisites for those that follow.

- **General Public:** IPRPA recommends that WVDE develop and embed a simple dashboard for the public on its IDEA landing page. In this era, the public is well-versed in navigating dashboards for various governmental interactions (e.g., property taxes). Currently, the WV State Auditor's Office provides a great example with its local spending data.
- **Scientific/Expert Community:** IPRPA recommends that WVDE commission an R-package that bundles the data with an API for delivery to the scientific community and promotes uptake among researchers and analysts in the public and private sectors.

4 What is IDEA?

The U.S. Department of Education's IDEA data refers to resources associated with the Individuals with Disabilities Education Act (IDEA). This U.S. federal law codified as the Individuals with Disabilities Education Act of 2004, ensures that students with disabilities receive appropriate public education tailored to their needs.¹

The IDEA data encompasses information about:

- **Child Count:** The number of children and students served under IDEA.
- **Environment:** Where children and students with disabilities receive their education.
- **Assessment:** The participation and performance of students with disabilities on statewide assessments.
- **Dispute Resolution:** Resolutions of disagreements between parents and educational agencies.

¹ 20 U.S.C. §§ 1400 et seq.

- **Discipline:** Discipline incidents for children and students with disabilities, such as out-of-school suspensions or expulsions.
- **Exiting:** Information about students with disabilities who drop out or graduate from high school.
- **Personnel:** The number of personnel serving children with disabilities.

IDEA data holds significant value beyond administrative purposes. For interdisciplinary research, it provides invaluable insights into the educational experiences, challenges, and outcomes for students with disabilities. This data can be combined with other data sets from the physical or built environment, social and governmental systems, and socio-economic, demographic, and community characteristics to paint a comprehensive picture of the multifaceted challenges faced by these students. Moreover, in public policy, IDEA data informs policymakers about the effectiveness of existing interventions, helping guide resource allocation and shaping future policies to ensure better outcomes for students with disabilities.

A notable aspect of IDEA is the federal financial assistance provided to states and localities to support the education of students with disabilities. The exact amount varies each year based on Congressional appropriations, and the funding primarily arrives through IDEA Part B grants, assisting the education of children with disabilities aged 3-21. Additional provisions, like IDEA Part C, offer funds for early intervention services for infants and toddlers with disabilities. This federal assistance underscores the government's commitment to ensuring equitable education for all students, regardless of their developmental challenges.

This data helps the federal government, state, and local educational agencies monitor IDEA's implementation, allocate resources, and ensure that students with disabilities receive the services they need to succeed in school. It also provides transparency and allows for assessing the effectiveness of special education programs across the country.

4.1 Public Resources

- *U.S. Department of Education (USED), Office of Special Education Programs (OSEP)*
 - [OSEP's Official Website](#)
 - [Official Website for IDEA](#)
 - Phone: +01 (202) 245-7459
 - The OSEP and IDC websites usually provide annual reports and other publications related to the IDEA program for each state.
- *West Virginia Department of Education (WVDE)*
 - [WVDE's Official Website](#)
 - [WVDE Office of Special Education](#)
 - [WVDE's New \(2023-2024\) Data Dashboard](#)
 - Phone: +01 (304) 558-2681
 - The state's official education website should have sections dedicated to special education, where you can find specific reports, data, and updates related to IDEA implementation in West Virginia.
- *West Virginia Developmental Disabilities Council (WVDDC)*
 - [WVDDC's Official Website](#)
 - Phone: +01 (304) 558-0416
 - The WVDDC is an advocacy organization that ensures that individuals with developmental disabilities in West Virginia receive the support, care, and opportunities they need to thrive. The council works on policy

recommendations, funding initiatives, and public awareness campaigns to promote the inclusion, independence, and well-being of people with developmental disabilities in the state.

5 Understanding Data Organization and Curation

Analysts divide the processes by which data are available and ready for use into data collection, organization, and curation (Workman 2020). For West Virginia's IDEA data, data collection is largely determined by the state and federal government as a condition of assistance (see Section \ref{statutes}) for developing state plans and targets for children with developmental disabilities on individual education plans. In our review of the data organization and curation, West Virginia's IDEA data would benefit from an overhaul in how the data is stored, delivered to the public, and organized for use by analysts.

Data organization refers to the data structure within the spreadsheets made available to the public, officials, and analysts hoping to better understand issues surrounding developmental disabilities in young people. Proper data organization allows public, private, and non-profit analysts to import and integrate IDEA data with other databases seamlessly. This is especially important given that most of our major public health concerns are multidimensional problems that relate to other factors in society in complex ways.

Data curation implies that the organization is crafted intentionally to facilitate addressing substantive questions with specific empirical methodologies or analytical approaches. Substantively, data curation enables answering questions within the data and linking it to external, related areas of concern (e.g., health or social services). In other words, it fosters the integration of data on developmental disabilities with other associated data that may shed additional light on how education, health, and social systems, at the state and non-profit level, may best serve these individuals. Data organization and curation are of utmost concern for issues like developmental disabilities.

Data organization is an issue best fostered by a set of mostly objective, logistical practices and norms common across social and natural sciences. Put simply, there is an objective best organization of data within a spreadsheet (Broman and Woo 2018; Wickham 2014). Data curation differs in that it is intentional and problem-oriented---what problems does the data help us frame, define, and understand better, and what questions do we wish to answer with that data? Data organization should be common, while data curation varies with the questions we want to answer.

5.1 Two Types of Data Curation

The emergent field of administrative informatics delves into pivotal concerns regarding the accessibility, organization, and utilization of data within administrative and policy systems (Overton et al. 2023). A significant portion of discourse surrounding administrative informatics contemplates the structuring of data. The objective is not merely to satisfy reporting mandates but to cater to policymakers and analysts striving to address pertinent public issues.

An intrinsic tension between two primary objectives lies at the heart of all administrative systems. The first emphasizes government accountability to its citizenry. Such accountability often leans towards "look-up" data infrastructures, focusing on specific data points and

emphasizing individual trees rather than the forest. While valuable for pinpoint transparency, this approach can overshadow the significance of understanding comprehensive data distributions. Gaining insights into these distributions is essential. It allows us to discern broader patterns that manifest across time and space, providing a more holistic understanding of complex issues and guiding more effective policymaking.

Like most data generated and reported under federal and state statutory requirements, IDEA data are organized and curated to meet the first goal—government accountability to the public. States collect and report IDEA data according to federal statutes mandating the provision of information as a condition of continued federal assistance for programs attached to the law. While the data is pivotal in maintaining transparency and ensuring that states meet federally set standards, it may not always be structured optimally for in-depth analytical exploration or problem-solving. Nevertheless, the inherent value of IDEA data transcends accountability. When used effectively, it can offer insights into the experiences and outcomes of students with disabilities, informing more robust policies and interventions. The challenge lies in harmonizing the demands of compliance and accountability with the equally vital objective of drawing meaningful insights for impactful policy decisions that account for the substantively complex nature of the issue.

The second objective underscores the imperative to address and, where feasible, rectify public challenges. These challenges, often termed "wicked problems" due to their multidimensional and interconnected nature (Rittel and Webber 1973), necessitate a deep analytical understanding of data across temporal and spatial dimensions. In today's world, these wicked problems intricately intertwine human systems with natural and physical ones, emphasizing the intricacies faced when integrating traditional policy data with other essential data sets, such as public health disparities or environmental justice indicators (Head 2022). Addressing these wicked problems often requires innovative approaches like dialogue mapping to foster shared understanding and collaborative solutions (Conklin 2005).

5.2 Principles of Data Collection and Organization

Before diving into data collection and organization principles, it is important to clarify the distinction between data curation and data organization. Data curation is an intentional, problem-oriented process that involves managing and maintaining data to ensure its long-term availability and usability. It requires tailoring the data handling practices to specific challenges or questions, encompassing tasks like data annotation, enrichment, and archiving. On the other hand, data organization is a more standardized process that focuses on the immediate structuring and arrangement of data, aiming to make it easily retrievable and understandable.

At the foundation of scientific inquiry and analysis lies the principle of being faithful to the data-generating process. This involves acknowledging and preserving the inherent characteristics of data as they emerge from various social, biological, or physical processes. In the era of digital technology, the way data is stored can greatly affect its integrity and usability. Keeping data in text or text-compatible formats like .txt or .csv ensures that the original form of the data is preserved. This method avoids hidden structures or content alterations in complex file formats, such as leading or trailing white space, numbers stored as text, and character strings read as factor variables. Preserving the data in its raw, textual form allows analysts to avoid these pitfalls and maintain fidelity to the data-generating process. Furthermore, preserving data is about maintaining its original form and ensuring its long-term availability. Backing up

data in multiple locations, including physical hard drives, safeguards against potential losses due to technical glitches, hardware failures, or other unexpected events.

Data curation is about more than just placing data into an objective structure. It involves creating an accessible, intuitive system that reflects the complexity and relationships within the data. Observations are usually placed in rows, variables in columns, and the corresponding values in the matrix of cells between them. This standardization ensures compatibility with most analytical tools and provides a logical, readable format that mirrors how we naturally think about data. When dealing with complex data that often contains nested or grouped structures, representing these groupings as separate columns ensures that the data's hierarchical relationships are maintained without complicating the basic structure. Labels are another crucial aspect of data organization. Simple, descriptive labels without special characters or spaces help to avoid misunderstandings or errors during analysis.

Separating calculations from data collection is another principle that must be considered. This emphasizes the preservation of the original, raw form of the data, ensuring that any transformations or calculations are transparent and reversible. Performing calculations such as sums, averages, or statistical analyses outside the primary data repository ensures that the original data remains untouched. This separation allows for transparency in how numbers were derived and enables easy updates or changes without altering the underlying data. It also facilitates automated reporting and analysis, making ongoing work more efficient. Furthermore, summaries are reductions of data, and when done prematurely, they can limit the analyses or insights that might be drawn later. The ability to explore different aspects of the data later is preserved by avoiding summaries like totals or averages within the data structure.

Finally, how data is classified or categorized can profoundly impact the insights drawn from it. In the early stages of data classification, consistency is more important than accuracy in each instance. If data are recorded consistently wrong, it is easy to correct en masse later. However, if data is recorded inconsistently, the inconsistency means data must be repeated, costing time, resources, and personnel hours. Classifications should also be mutually exclusive, meaning each observation falls into one and only one category. Ambiguous or double categorization can lead to confusion and errors in analysis.

By adhering to these principles and practices, analysts can ensure that data is accurate, consistent, adaptable, and robust. This solid foundation supports various analytical endeavors, from scientific research to private and non-profit sector strategizing. It also makes life easier for the analyst by simplifying data management and enhancing their findings' credibility, reproducibility, and transparency. Care should also be taken when using pre-existing data sets, as their original purpose might not align with the current objective. We can begin assessing the quality, challenges, and opportunities West Virginia's IDEA data provides from these general guidelines.

6 West Virginia's IDEA Data

West Virginia's IDEA data is comprehensive and meets statutory requirements per the Department of Education's (ED) determination letters issued June 23rd (Part B) and 21st (Part

C).² Yet, how the data are delivered presents significant obstacles to incorporating insights into health and education policy, informing the public, and promoting data uptake by researchers working in the area and adjacent areas. Moreover, surrounding states deliver data in a similar format, offering West Virginia an opportunity to lead in this area as in the upgrades to the State Auditor's Office tracking public spending in localities. After reviewing the challenges with the data, we offer recommendations for significant advancement in organization, curation, and delivery to the public and data analysts working in the public, private, and non-profit sectors.

In what follows, we discuss the delivery of the IDEA data alongside issues of organization and curation. We identify three target populations for the data and discuss data products designed to reach each of these audiences. Our recommendations nudge the IDEA data past the current “look-up” system and designs and towards more engagement with the public, private sector firms and non-profit communities, and researchers working in education policy and adjacent issue areas like health and public services.

6.1 Data Delivery

The primary audience for IDEA is the USED, whose administrators use the data to make evaluations of state compliance with federal statutes. Nothing in our data review should be taken to obscure the federal government's primacy as a stakeholder and primary audience for the data. However, the state and local educational districts are collecting this data anyway, and some reorganization and intentional structuring of the data would allow both private sector firms (including service providers) and the non-profit community to understand better and strategize service delivery for these children. Additionally, and importantly, it would allow researchers in the state's higher education institutions working in health, education, and public policy to understand better the complex issues surrounding disability in the educational system.

Each community requires different types of delivery for the IDEA data. While meeting the bar for federal reporting, the spreadsheets impede uptake for the more diverse set of non-profit and private stakeholders, the broader public, and researchers and analysts working in the area. Our findings are as follows:

- The main mode of delivery for the IDEA data is a set of spreadsheets that address Section 618 reporting and separate spreadsheets for reporting the data on indicators used for participation and determinations made by ED.
- However, spreadsheets are of limited use for informing the public, with little exposure and capacity for understanding complex data structures contained in them. This impedes public engagement with the data and the public's understanding of the general trends in state performance.
- Simultaneously, the “look-up” organization of the data in the spreadsheets also impedes data uptake in scientific communities and among community leaders such as private sector firms and non-profits working in the area. We address this organization in detail in the following section.

² See ED's webpage at <https://sites.ed.gov/idea/spp-apr-letters> for state reports and resulting determination letters. This report was completed in early Fall 2023.

- These issues are common to both the state and federal data repositories for section 618. Both state and federal attention to these issues is warranted to increase usability and data uptake into relevant communities.

In short, the mode of delivery for the IDEA takes the form usually served for expert communities but then attempts to make this form amenable to informing the general public. In doing so, the spreadsheets are not optimized for informing the public nor their use by researchers and expert communities, including service providers.

6.2 Delivery Recommendations

Because of this finding, we recommend that the spreadsheets be optimized for data use in these expert communities and that a simpler mode of data delivery for the public be developed for the website for state IDEA data. In particular, we have the following specific recommendations for delivering the IDEA data to the public, researchers, and relevant expert communities:

- The WVDE should, at a minimum, structure the spreadsheet data to be consistent with standard practices in the social sciences.
- **General Public:** Data tables for delivery to the general public should be in HTML, text, or PDF format, eliminating the need for navigating spreadsheets by the public, the non-profit sector, and private sector firms working on service delivery for these families.
 - Tables or dashboards embedded on the website for current years would be optimal. Note that as of early 2024, WVDE has a dashboard available.
 - We would also recommend very basic mapping facilities for the data delivered to the public to convey the geographic characteristics of the data.
- **Scientific/Expert Community:** Ideally, a master CSV or other text-compatible file would correct the fragmentation for delivery to scientific and other expert communities.

6.3 Delivery Logistics

The data across years appears in a set of fragmented spreadsheets. As noted above, a comprehensive spreadsheet in .CSV or other .TXT-compatible format has considerable logistical benefits regarding the recommendations above.³ The WV IDEA data is largely stored as MS Excel files, easily converted to the more portable .CSV file type. However, the fragmentation of the files by year makes the construction of embedded tables and dashboards logistically more difficult. The fragmented delivery impedes uptake in expert communities, including research and strategizing about service delivery, public health concerns, and understanding educational challenges for children with independent educational plans.

Building a comprehensive spreadsheet across annual reports and 618 data is made simpler if the general organization of the spreadsheets conforms to social science guidelines for organizing data in spreadsheets. This would allow the same statistical procedures for data cleaning to be applied across all spreadsheets and given the uniform reporting forms, render

³ CSV files are files containing comma-separated-values (CSV) and .TXT files refer to files stored as text. CSV files are .TXT-compatible and possible in standard spreadsheet applications such as MS Excel, OpenOffice, etc.

the process automatic for each additional year, ensuring the spreadsheets yield a living data set.

In our review of data availability as of Fall 2022, the section 618 data is easily found on the state's site for IDEA data; however, these data need the raw Local Education Organization (LEO) data for past reporting on this issue. While still informative, the table marginals presented in the most recent data convey much less information and offer less opportunity for uptake in expert communities. We recommend that the original raw data be kept on the site or otherwise scraped from ED's website for compilation and to serve as the back end of a set of dashboards for delivery to the public with proper control for privacy considerations. As currently constructed, the data's organization mitigates understanding IDEA relative to geospatial variation and temporal trends.

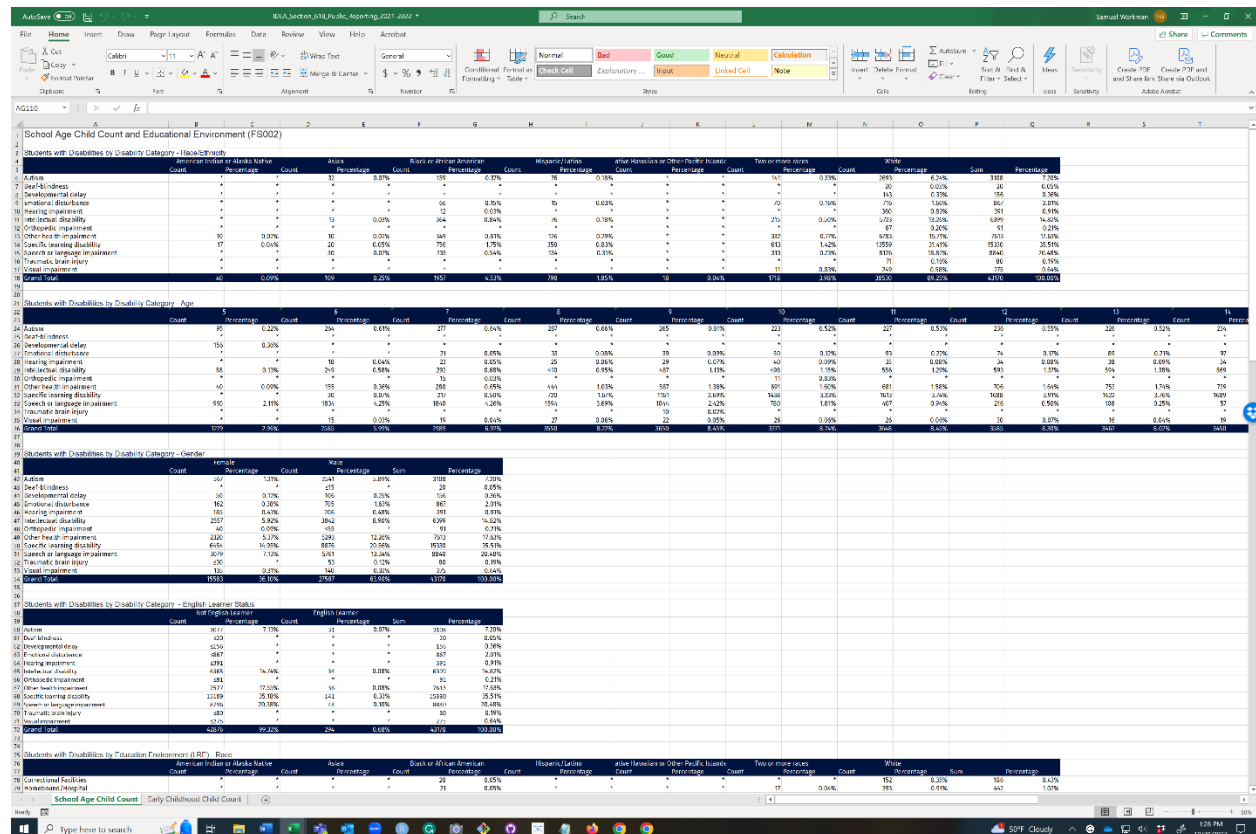


FIGURE 1 : IDEA DATA IN WVDE'S SPREADSHEET FORM

6.4 Data Organization in Spreadsheets

Unlike data curation, data organization in spreadsheets is subject to a set of mostly objective standards. The central difficulty in WV's IDEA data surrounds using spreadsheets to deliver a final product that should be informative to the public and useful to expert communities. Spreadsheets for public information should be avoided, considering most of the public has little familiarity with navigating the cells, tabs, and other features of a typical workbook like MS Excel. However, our recommendations here are useful as they build spreadsheets amenable to delivering polished data products to the public and allowing for uptake in scientific communities. Figure 1 displays a screenshot of the data for Section 618, Part B, dealing with

school-age children aged 3–21.⁴ Note that data for Part C, dealing with toddlers and those not old enough for school, is in a separate tab in the same spreadsheet. Our discussion will be limited to illustrative examples of changes needed in the organization of the spreadsheets.

First, the file names conform nicely to data science guidelines for naming files. White space should be avoided when naming files and headers within the spreadsheet. The Section 618 data is organized by year and mostly presents marginals derived from the raw data. Logistically, it would be preferable to construct end-user data products for the spreadsheets to be compiled from the raw data and presented as a comprehensive file. In general, the data should also come with a data dictionary—a document that serves as a road map and set of definitions for understanding data structures and values within the spreadsheet. For instance, in Figure 1, note the asterisks used for some cell values. Whether these represent small quantities, missing data, or unobserved data is unclear. Section 618 files should be accompanied by a data dictionary similar to that for the spreadsheets on the various indicators for IDEA plans identified in federal legislation, which is exceptional in its clarity.

The use of coloring and multiple-column headers is present throughout the spreadsheet. In general, this should be avoided. If reorienting the spreadsheets to fuel the development of data products meant to inform the public, back-end statistics and programming languages have difficulty understanding the layering in multiple overlapping column headers. These tend to be misread as textual data, providing the analyst or programmer with myriad problems reorienting the data. These overlapping or nested headers are meant to replace column variables in a more structured spreadsheet organization. Similarly, coloring in the overlapping, nested headers tend to represent variables, though here are for display purposes only. However, it should be noted that using color as a stand-in for variables is not recommended. Cells bearing color must be processed via specialized functions in the typical programming software for analyzing data.

Figure 1 also shows that the tables bear several percentage calculations—the key data are counts. We applaud the attention to detail and understanding that percentages without a baseline for understanding population data are of limited use. However, we recommend that no data cleaning or calculation be contained in the spreadsheets. Calculations made on the data should occur outside the delivery of the spreadsheets. The reasons for this are twofold. First, spreadsheets, as noted throughout this report, are not ideal for delivering the types of end-user data products for public consumption. Second, calculations inside a data spreadsheet may lead to logistical problems when they must be executed on multiple, fragmented spreadsheets. The problem is even more pernicious if the format of the spreadsheets should change from year to year—meaning the statistical routines and programming code must change for each additional year of data (or tab, as in the case of the Part C data).

The set of tables in each spreadsheet file pertains to marginals calculated from raw data not currently available on the website. These tables are intended to be end products for public consumption. Given the file's structure, they default to delimiting sub-populations in isolation and do not allow for unions or intersections of the data. As a case in point, we can understand the numbers of children with disabilities by gender, race, and age in isolation. However, most social science on health, learning, and physical and emotional welfare points to

⁴ Obtained from the state's website at <https://wvde.us/special-education/data-and-public-reporting/>.

intersectionality in observable outcomes (e.g., counts and percentages for Asian females under ten). Given public interest is likely to fall on children much like their own, these spreadsheets have difficulty delivering this type of intersectional information that more fully describes a child (the same for the rest of us). Restructuring the spreadsheet organization would not get to this type of intersectionality. Still, it would render the spreadsheet such that these could be easily calculated from the public-facing end-product, such as an embedded table or, better still, an interactive dashboard.

We reason that the public, especially parents, are interested in knowing the intersectional table margins, how the system performs over time with increased attention to issues around education and disabilities, and how their system may compare with others in the state. With a comprehensive spreadsheet built from raw data, these intersectionality and temporal performance information would be easily calculated statewide. However, the spreadsheets are mostly devoid of spatial information. This means the public cannot examine how a given system operates compared to peers and neighbors. This will become increasingly problematic as school choice becomes more widespread, and parents need information to fuel these decisions. If typical quantities for counts are low enough to raise concern for privacy issues associated with disability, at a minimum, the data could contain spatial identifiers for counties (e.g., FIPS codes), typically carried in other federal data such as the Census, the National Center for Education Statistics (NCES), the Appalachian Regional Commission (ARC), and public health agencies like the Centers for Disease Control (CDC) or the National Institutes of Health (NIH). Other states (e.g., Michigan) make the LEO data accessible as part of the state websites devoted to IDEA. These items are important for at least the following reasons:

- Spatial data would allow the construction of maps---one of the more easily understood data products for the general public when interested in spatial comparisons.
- Spatial data would allow the IDEA a broader uptake as data could be merged with data across the education, public health, and general research sectors.
- Spatial data would also better facilitate strategizing around further federal investment in the state across these sectors.

This concern with integrating IDEA data with other data in public health and education gets at the goal of broader uptake of the data in expert communities. The structural deficiencies we identify here mitigate against data curation in these other realms, where data is organized with questions that go beyond IDEA. Though we strongly recommend delivering these data in a comprehensive spreadsheet, the data organization can be improved file by file. The assessment above should not be viewed as a complete road map for data structure in the spreadsheets but as illustrative of the problems encountered from public and expert community perspectives. The structural recommendations above move the data towards fostering analysis and not merely “look-up,” as noted in previous sections. IPRPA has considerable experience in organizing, cleaning, and structuring such data. Please get in touch with us with questions about logistics and technical assistance.

6.5 Data Curation

The key distinction between data organization and curation is that data curation is intentional and guided by the problem or question at hand. The current data delivery mitigates against this broader uptake that exceeds public information and IDEA accountability standards. We

reason that the state is interested in at least three objectives with the IDEA, all outside federal reporting.

- IDEA data should inform the public about the education system and allow temporal and spatial comparisons to inform parental educational decisions.
- IDEA data should allow researchers and analysts at the state's higher education institutions to inform policymakers on spatial and temporal trends better and provide analysis when requested by the state and its partners in the private and non-profit sectors.
- IDEA data should help drive federal investment in West Virginia in adjacent and related areas of public health, capacity-building, workforce development, and education.

The recommendations here assume that the WVDE has undertaken the recommendation for data organization above, including attaching geospatial identifiers to the data. From the three objectives outlined above, we make two major recommendations for data infrastructure. These assume a comprehensive data infrastructure that integrates across years and spatial characteristics, even if continued delivery of the fragmented spreadsheets is desired or required by the USED. Our major recommendations for improving data infrastructure, delivery, and uptake are as follows:

- **General Public:** IPRPA recommends that WVDE develop and embed a simple dashboard for the public on its IDEA landing page. In this era, the public is well-versed in navigating dashboards for various governmental interactions (e.g., property taxes). Currently, the WV State Auditor's Office provides a great example with its local spending data.
- **Scientific/Expert Community:** IPRPA recommends that WVDE commission an R-package that bundles the data with an API for delivery to the scientific community and promotes uptake among researchers and analysts in the public and private sectors.

First, data should be delivered in a form amenable to public consumption and understanding. This means that the set of spreadsheets should not stand alone. We recommend WVDE build a shared dashboard that delivers the IDEA data in a format amenable to questions about intersectionality (demographic and otherwise). This dashboard should form the basis for customizable tables embedded within the WVDE website for its IDEA data. Facility for this dashboard is increasingly available in open-source code such that WVDE is not dependent on proprietary source coding for its maintenance. We recommend the open-source programming language **R** and its `shiny` and `flexdashboard` packages for delivery.

The data's general audience will have encountered dashboards in many other realms of interaction with state and local government (e.g., paying county property taxes). In addition, delivering the data in a dashboard outputting embedded tables does not require access to MS Excel or other proprietary software to examine and reason through the data. Finally, the data would be more easily accessible to mobile users on the WVDE's site, potentially driving more traffic. Increasingly, users interact on mobile platforms over and above traditional desktops. Using such a strategy, the data would be accessible from a web browser, reaching more users, parents, and families. An embedded dashboard could provide the fodder for simple mapping on WVDE's site—perhaps the easiest way to make relative comparisons across the state's educational systems.

Our second recommendation is that WVDE commission the development of an **R** package to deliver the IDEA data, including API access. Such a package would allow researchers across a broad spectrum of related issues and sectors to obtain the most recent data and foster uptake in related issue areas, including public health, education research, community development, and workforce issues. The issue of developmental disabilities requiring independent educational plans touches on many issues common in the Appalachian region, including economic and workforce development, public health, and community resilience. These issues need IDEA data that is amenable to temporal and spatial analysis.

These comparisons go beyond the educational system for children and families and are key to understanding and taking advantage of the many social and economic transitions underway in Appalachia. An accessible database for the IDEA data would allow research and strategizing around these adjacent issues and provide a basis for policymaking and investment in the region.

7 Overview of Relevant Federal Statutes

The major federal statute governing education plans for children with developmental disabilities is 20 USC 33. This statute contains performance goals and indicators for plans, directives for state performance plans, reporting requirements, and implementation assessments for state plans. The following summarizes the different provisions and requirements, especially for data reporting, and assesses West Virginia’s compliance with the general guidelines in 20 USC 33.

20 USC 33 § 1412 Performance goals and indicators

- (a) has established goals for the performance of children with disabilities in the State that–
 - (i) promote the purposes of this chapter, as stated in \S 1400(d) of this title;
 - (ii) are the same as the State's long-term goals and measurements of interim progress for children with disabilities under section 6311(c)(4)(A)(i) of this title;
 - (iii) address graduation rates and dropout rates, as well as such other factors as the State may determine; and
 - (iv) are consistent, to the extent appropriate, with any other goals and standards for children established by the State;
- (b) has established performance indicators the State will use to assess progress toward achieving the goals described in subparagraph (A), including measurements of interim progress for children with disabilities under section 6311(c)(4)(A)(i) of this title; and
- (c) will annually report to the Secretary and the public on the progress of the State, and of children with disabilities in the State, toward meeting the goals established under subparagraph (A), which may include elements of the reports required under section 6311(h) of this title.

Summary: Each state is required to establish goals for the performance of children with disabilities in compliance with IDEA, establish performance indicators (metrics) for measuring progress toward achieving outlined goals, and annually report to the Secretary of Education and the public on the progress of the state toward meeting the goals based on established indicators.

20 USC 33 § 1416 State performance plans

- (a) Plans
 - (i) In general

- 1) Not later than 1 year after December 3, 2004, each State shall have in place a performance plan that evaluates that State's efforts to implement the requirements and purposes of this subchapter and describes how the State will improve such implementation.
- (ii) Submission for approval
 - 1) Each State shall submit the State's performance plan to the Secretary for approval in accordance with the approval process described in subsection (c).
- (iii) Review
 - 1) Each State shall review its State performance plan at least once every 6 years and submit any amendments to the Secretary.
- (b) Targets
 - (i) In general
 - 1) As a part of the State performance plan described under paragraph (1), each State shall establish measurable and rigorous targets for the indicators established under the priority areas described in subsection (a)(3).
 - (ii) Data collection
 - 1) In general, Each State shall collect valid and reliable information as needed to report annually to the Secretary on the priority areas described in subsection (a)(3).
 - 2) Rule of construction: Nothing in this chapter shall be construed to authorize the development of a nationwide database of personally identifiable information on individuals involved in studies or other collections of data under this subchapter.
 - (iii) Public reporting and privacy
 - 1) In general: The State shall use the targets established in the plan and priority areas described in subsection (a)(3) to analyze the performance of each local educational agency in the State in implementing this subchapter.
 - 2) Report
 - a) **Public Report**—The State shall report annually to the public on each local educational agency's performance on the targets in the State's performance plan. The State shall make the State's performance plan available through public means, including by posting on the website of the State educational agency, distribution to the media, and distribution through public agencies.
 - b) **State Report**—The State shall report annually to the Secretary on the performance of the State under the State's performance plan.
 - c) **Privacy**—The State shall not report to the public or the Secretary any information on performance that would result in the disclosure of personally identifiable information about individual children or where the available data is insufficient to yield statistically reliable information.

Summary: Before December 3, 2005, each state was instructed to formally establish a “state performance plan” outlining compliance goals and evaluating state progress toward meeting these goals to be submitted to the Secretary of Education for review. An updated “SPP” must be submitted every six years for review and approval. Part of the state performance plan stipulates annual data collection and reporting on progress to the Secretary of Education and the public toward meeting these goals using established and approved indicators.

20 USC 33 § 1418 Reporting for federal assistance

- (a) In general

- (i) Each State that receives assistance under this subchapter, and the Secretary of the Interior, shall provide data each year to the Secretary of Education and the public on the following:
 - 1) The number and percentage of children with disabilities, by race, ethnicity, limited English proficiency status, gender, and disability category, who are in each of the following separate categories:
 - a) Receiving a free appropriate public education.
 - b) Participating in regular education.
 - c) In separate classes, schools, facilities, or public or private residential facilities.
 - d) For each year of age from age 14 through 21, stopped receiving special education and related services because of program completion (including graduation with a regular secondary school diploma), or other reasons, and the reasons why those children stopped receiving special education and related services.
 - e) Removed to an interim alternative educational setting under section 1415(k)(1) of this title. (II) The acts or items precipitating those removals. (III) The number of children with disabilities who are subject to long-term suspensions or expulsions.
 - 2) The number and percentage of children with disabilities, by race, gender, and ethnicity, who are receiving early intervention services.
 - 3) The number and percentage of children with disabilities, by race, gender, and ethnicity, who stopped receiving early intervention services from birth through age 2 because of program completion or for other reasons.
 - 4) The incidence and duration of disciplinary actions by race, ethnicity, limited English proficiency status, gender, and disability category, of children with disabilities, including suspensions of 1 day or more.
 - 5) The number and percentage of children with disabilities who are removed to alternative educational settings or expelled as compared to children without disabilities who are removed to alternative educational settings or expelled.
 - 6) The number of due process complaints filed under section 1415 of this title and the number of hearings conducted.
 - 7) The number of hearings requested under section 1415(k) of this title and the number of changes in placements ordered because of those hearings.
 - 8) The number of mediations held, and the number of settlement agreements reached through such mediations.
- (ii) The number and percentage of infants and toddlers, by race and ethnicity, who are at risk of having substantial developmental delays (as defined in section 1432 of this title) and receiving early intervention services under subchapter III.
- (iii) Any other information that may be required by the Secretary.

Summary: These constitute the boilerplate state reporting requirements for continued state assistance and funding by the Departments of Education and Interior.

8 Institute for Policy Research and Public Affairs Partnerships

The Rockefeller Institute for Policy Research and Public Affairs (IPRPA) in Eberly College at WVU is a non-partisan source of research, data, and analysis for state and local officials in West Virginia and the broader Appalachian region. IPRPA conducts basic and applied research on various problems important to the region and related to public policy and broader social, political, and economic transitions.

We employ the full range of qualitative and quantitative research methodologies to societal problems to help public, non-profit, and private sector partners plan, strategize, and adapt to changes and challenges in the region, be they physical or biological, economic, governmental, or societal. We have extensive expertise in developing and maintaining large-scale data infrastructures to answer fundamental questions about public policy and its effects on communities. Our partnerships span the public, nonprofit, and private sectors and develop research co-designed with our stakeholders, clients, and communities, improving the use of research and data-driven decision-making in tackling important policy problems. The Institute aims to spur evidence-based policymaking and uptake of research in the state and region.

Please contact the [Institute for Policy Research and Public Affairs \(IPRPA\)](#) for questions and technical assistance in pursuing any recommendations outlined here.

8.1 How to Cite This Report

Workman, Samuel, and Oreoluwa Runsewe. 2024. "IDEA Data Reporting in West Virginia: Challenges and Opportunities." Institute for Policy Research and Public Affairs (IPRPA), West Virginia University (WVU). IPRPA-2024-005.

8.2 Our Visualizations—Media

All tables and data visualizations in this report are available. Please get in touch with us here: policyresearch.wvu.edu/contact-us.

9 Acknowledgements and Disclaimers

This project is made possible by funding from the WV Developmental Disabilities Council (WVDDC) and partly supported by a grant from the U.S. Administration for Community Living (ACL), Department of Health and Human Services, Washington D.C. 20201.

While we acknowledge the support provided by WVDDC and Think Kids, errors in content, style, or judgment in the report remain with the Institute for Policy Research and Public Affairs (IPRPA) at West Virginia University.

10 References

- Broman, Karl W., and Kara H. Woo. 2018. "Data Organization in Spreadsheets." *The American Statistician* 72 (1): 2-10. <https://doi.org/10.1080/00031305.2017.1375989>.
- Conklin, Jeff. 2005. *Dialogue Mapping: Building Shared Understanding of Wicked Problems*. John Wiley & Sons, Inc.
- Head, Brian W. 2022. *Wicked Problems in Public Policy: Understanding and Responding to Complex Challenges*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-94580-0>.
- Overton, Michael Overton, Stephen Kleinschmit, Mary Feeney, Federica Fusi, Nick Hart, Spiro Maroulis, Kayla Schwoerer, Eric Stokan, Herchel Thomas, and Samuel Workman. 2023. "Administrative Informatics: A Roundtable on the Conceptual Foundations of a Public Administration-Centered Data Science Subfield." *Journal of Behavioral Public Administration* 6 (June). <https://doi.org/10.30636/jbpa.61.330>.
- Rittel, Horst W. J., and Melvin M. Webber. 1973. "Dilemmas in a General Theory of Planning." *Policy Sciences* 4 (2): 155-69. <https://doi.org/10.1007/BF01405730>.
- Wickham, Hadley. 2014. "Tidy Data." *Journal of Statistical Software* 59 (10). <https://doi.org/10.18637/jss.v059.i10>.
- Workman, Samuel. 2020. "Four Principles of Data Collection." *Towards Data Science*. <https://towardsdatascience.com/four-principles-of-data-collection-82ad07938ac1>.