IDOT Work Zone Activity Data (WZAD)

Work Zone Data Hub (WZDH) Technical Memorandum

Prepared for: Illinois Department of Transportation

Prepared by: University of Wisconsin-Madison and Lakeside Engineers

Date: October 8, 2026

Introduction

This technical memorandum documents the data processing framework of the Illinois Department of Transportation (IDOT) Work Zone Data Hub (WZDH). It catalogs the work zone activity data (WZAD) sources ingested by the platform, the processes used to transform and harmonize those data into a common model, and the mechanisms used to provision the data for analysis and dissemination.

The WZDH was developed as a foundational component of the 2022 IDOT SMART Grant to provide a scalable, centralized platform for IDOT work zone activity data management. As a primary outcome of the project, the WZDH serves as the operational engine behind IDOT's Connected Work Zone (CWZ) 1.0 data feed, published to the USDOT Work Zone Data Registry on May 1, 2026. Additional objectives include enabling work zone analysis, performance management, and road user information through standardized data sharing, while accommodating future data sources, technologies, and operational requirements.

A conceptual architecture of the WZDH is illustrated in Figure 1. The hub decouples three primary data processing functions - data acquisition, transformation, and dissemination - enabling scalable integration of multiple WZAD sources and consistent delivery of harmonized outputs. The WZAD sources (inputs) and users (outputs) are described in subsequent sections.

The internal framework consists of three primary components:

- *Ingestion* – The ingestion layer captures WZAD from multiple external source systems using source-specific interfaces and update frequencies. Source data are archived in their original formats to preserve provenance and support traceability. Processes execute on a recurring 5-minute cycle, during which incoming records are compared against previously ingested data to identify new or modified records for downstream processing.
- *Harmonization* – The harmonization layer transforms and integrates ingested WZAD into a unified internal representation that supports consistent dissemination, analysis, and operational use. This layer normalizes attributes across disparate sources, aligns spatial and temporal information, and consolidates related records into a single, statewide view of work zone activity. The harmonized representation is managed through a common data model based on the Connected Work Zone (CWZ) 1.0 standard, extended to support broader

analytical, operational, and historical WZAD needs within IDOT. By decoupling source-specific formats from internal data management and dissemination requirements, the harmonization layer provides a normalized data foundation for downstream processes.

- *Provisioning* – The provisioning layer prepares harmonized WZAD for consumption by external systems and applications. This includes formatting data for real-time APIs, map services, and analytical tools, as well as exporting selected data elements into tabular representations for reporting and analysis. Provisioning services currently operate on a 5-minute cycle to support the two primary dissemination endpoints: the IDOT CWZ 1.0 data feed and the AGO WZAD Analytics Tool.

The WZDH is designed as a lightweight, modular platform that supports incremental addition of new data sources, evolution of the data model, and future operational and analytical applications.

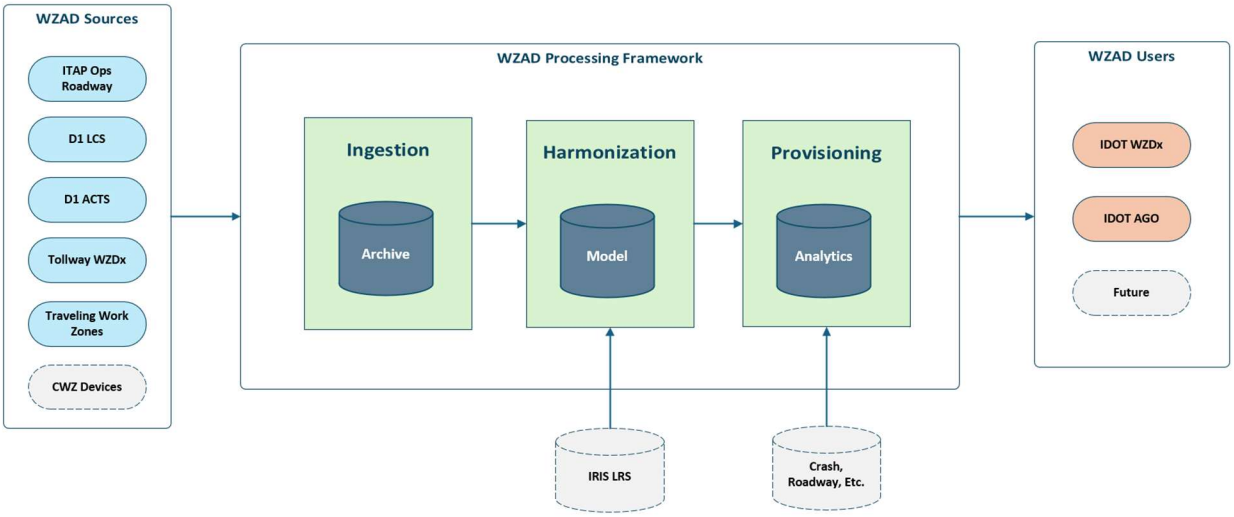


Figure 1. IDOT Work Zone Activity Data Hub (WZDH) conceptual architecture. Components drawn with dashed outlines represent future (planned) development.

The remainder of this document describes the data sources and the processes by which they are retrieved, archived, and transformed to a common data model. Transforming all incoming data into a common representation with known required and optional properties facilitates further processing used to generate exported records suitable for use to generate the CWZ and for other purposes. All processes described are scheduled to run at 5-minute intervals.

WZDH Data Sources

A primary goal of the WZDH was to provide a platform by which different data sources could be integrated into the common model framework without resorting to code changes or direct edits to a database. This is accomplished by a user interface that lets non-technical users define a new data source and connection configuration. The configuration contains the URL and desired update check

frequency as well as credential information dependent on the source including an optional user name and password, API key, and token generation endpoints as shown in Figure 2.

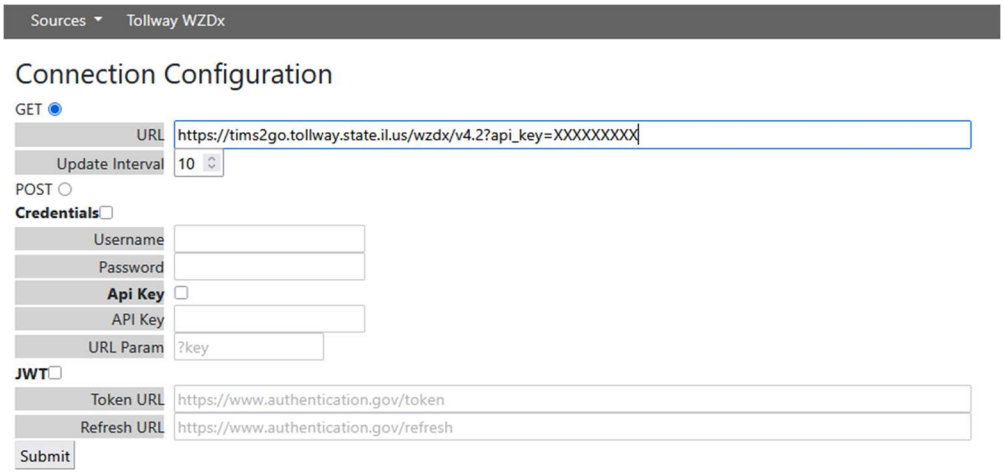


Figure 2. Data source connection configuration user interface.

More complicated cases such as XML sources and sources without an accessible URL require configuration beyond that which can be reasonably accommodated in a self-service application. For these cases, additional customization is required to extend the retrieval interface which, as a modular component, plugs directly into the subsequent archiving and harmonization processes.

The WZDH currently ingests five primary work zone activity data sources with the names and access method given in Table 1. A general description of these datasets is provided in the IDOT Work Zone Data Summary Technical Memorandum, developed under Task 2 of the SMART Grant, which documents the available data and their characteristics.

This memorandum complements the Data Summary by documenting how those data are implemented within the WZDH. The descriptions that follow focus on how each source is acquired, structured, and used within the platform to support standardized data sharing and operational use. Data dictionaries for each source are included in the Appendix.

ITAP OPS Roadway Data

The ITAP OPS Roadway Data is maintained by the IDOT Bureau of Information Processing (BIP). This dataset is based on IDOT's electronic OPER 2410 process, which supplies manually approved construction, utility, and maintenance related work zone restriction information to IDOT's oversize / overweight freight permitting system. This data is post-processed through the ITAP geo-spatial database system and is made available to the WZDH through two IDOT AGO data interfaces representing linear and point features. The AGO feature layers are typically updated nightly. The OPS Roadway has similar content to the Getting Around Illinois Dataset – both are derived from the ITAP database, however, OPS Roadway has been created specifically for operational use by the WZDH. BIP has made significant improvements to its quality, consistency, and completeness over the course of the SMART Grant project. It has been observed in practice that the data feed updates daily

Monday through Friday with approximately 1000 features in the points layer and of those, 10 or fewer updates in a given day. The points and lines are complementary and are associated by the "submissionId" which serves as a unique identifier for a given work zone. Where an associated linear feature exists for a work zone represented as intermediary vertices in the points layer, only the linear feature is propagated to subsequent processes. Otherwise, if a work zone event exists in the points layer with no linear feature associated by submission Id, the point event feature is passed for further processing.

D1 Lane Closure System (LCS) Data

The District 1 (D1) LCS data source tracks lane closures on the D1 expressway network. The LCS system is maintained by the University of Illinois – Chicago (UIC) on behalf of D1 and serves as a data source for both operational needs for D1 and information dissemination through the Travel Midwest regional traveler information system. The D1 LCS data is provided to the WZDH in XML format via an *rsync* data exchange interface that was configured by UIC for this project. UIC has plans to eventually provide the XML data over a web service interface. In some cases, D1 LCS entries will prompt OPER 2410 submissions, leading to potential overlap between the D1 LCS and ITAP OPS Roadway data sources. The data retrieval script checks the D1 LCS source hourly and when a new record is detected, an XML to JSON transformer is applied to facilitate ingestion into the WZDH by the same means as the other sources. The data is relatively constant when compared to other sources, with fewer than 10 changes typically occurring over the course of several days.

D1 Arterial Construction Tracking System (ACTS) Data

The D1 ACTS data source tracks construction and utility lane closure requests for arterial roadways in D1. The ACTS systems is maintained by the University of Illinois – Chicago (UIC) on behalf of D1 and serves a similar, complementary role to D1 LCS as a data source for both operational needs and information dissemination through the Travel Midwest regional traveler information system. The D1 ACTS data is provided to the WZDH in XML format via an *rsync* data exchange interface that was established by UIC for this project. UIC has plans to eventually provide the XML data over a web service interface. The XML data is typically updated daily. In some cases, D1 LCS entries will prompt OPER 2410 submissions, leading to potential overlap between the D1 LCS and ITAP OPS Roadway data sources. Both ACTS and LCS are derived from the same source XML structure. The same XML to JSON process as a precursor to further processing described regarding the D1 LCS source applies. Similar to the D1 Source, there may be few or no changes over the course of a day.

Illinois Tollway WZDx Data

The Illinois Tollway maintains a WZDx v4.2 compliant work zone data feed covering tollway operations in D1 and D2. The feed endpoint is published in the USDOT Work Zone Data Registry and is accessible via an API key provided by the Tollway. The Tollway WZDx data feed is incorporated in its entirety into the WZDH, with minor changes to convert the representation into a CWZ 1.0 aligned format. Work zone records in the Tollway WZDx feed are regularly updated. Based on the update times provided in the Tollway WZDx feed, it is observed to update at ten minute intervals. Within the feed, fewer than ten events typically have an update in a given iteration.

Traveling Work Zones Data

Traveling work zone data provide real-time, geo-located activity status from IDOT's highway maintenance fleet. This data source consists of GPS point locations, provided a fixed intervals, that are activated by fleet drivers when commencing maintenance activities. The transmission of points is terminating algorithmically when the maintenance vehicle exceeds 40 mph. This data source is provided to the WZDH by IDOT BIPS via an access restricted AGO layer. As with the other sources, the source feed is polled at 5-minute intervals and any change will be conditionally saved. Different from the other sources, no start- or end-time is provided in the source feed. The start and end time for the traveling work zone event is conservatively set as two hours before or after the feed acquisition time.

CWZ Devices Data (Future)

Connected work zone devices, such as connected arrow boards and location markers, provide an opportunity to implement time and location field verification of scheduled work zones, along with the potential identification of work zone activity not captured by existing data sources. Integrating CWZ devices data into the WZDH is planned for future work.

Data Source	Maintainer	Coverage	Update Frequency	Format	Access Method
ITAP OPS Roadway	IDOT BIP	Statewide	Daily M-F	ESRI Feature Layer	IDOT AGO
D1 Lane Closure System (LCS)	D1 / UIC	District 1	Variable (typically daily)	Custom XML	UIC rsync
D1 Arterial Construction Tracking System (ACTS)	D1 / UIC	District 1	Variable (typically daily)	Custom XML	UIC rsync
Illinois Tollway WZDx	Tollway	District 1	Near real-time	WZDx 4.2	Tollway API
Traveling Work Zones	IDOT BIP	Statewide	Real-time	ESRI Feature Layer	IDOT AGO

Table 1. Summary of the five implemented WZDH data sources.

Data Retrieval

A scheduled task iterates over the configured data sources and connection information to retrieve updates from the source and conditionally creates a new archival record. The first step in this process is to retrieve from the database the most recent archive for a given source. The process exits early if the source specific desired update interval has not yet passed based on a comparison of the system time and the most recent archive record time. The data is then retrieved from the source typically as an HTTP request though the D1 LCS and ACTS sources go through an intermediary step to convert XML sources retrieved by rsync into a JSON structure compatible with subsequent processes used for other sources. The newly retrieved data feed is compared against that of the most

recent record. In cases where a difference is detected, a new archive record is created with a column recording the update time as represented in the feed when available or otherwise the system time. The conditional archive creation prevents unneeded work by subsequent processes during times that source data has not changed.

Common Model Transformation

A key step in the data workflow is to take the work zone entities represented in the source data in disparate formats and convert all to a common structure. While not completed in the first stage of the WZDH, a similar approach will be applied to incoming connected work zone device data under a future project phase. Due to its flexible nature, JavaScript Object Notation (JSON) was selected to describe the structure of the input data, the transformation logic, and the container for the common model output. The JSON schema convention permits definition of the data structure, property types (strings, numbers, dates, etc.), and the distinction between required and optional properties. Tooling allows validation of data against a defined schema and if data is found to be invalid, a readily available report to the reason why should it be a missing child data structure or incorrect property type.

The common model leans heavily on the established schemas available at:

https://github.com/ite-org/cwz/tree/main/schemas/1.0

Rather than build custom data structure to contain work zone and connected device information, the CWZ specification provides ready made and vetted structures for each form the basis of our common data model. Complementary data structures capture additional information such as text based locations and scheduling as this does not have a place in the CWZ specification.

Figure 3 shows a high level of the data structure represented a block diagram and a subset of the properties defined in the JSON Schema.

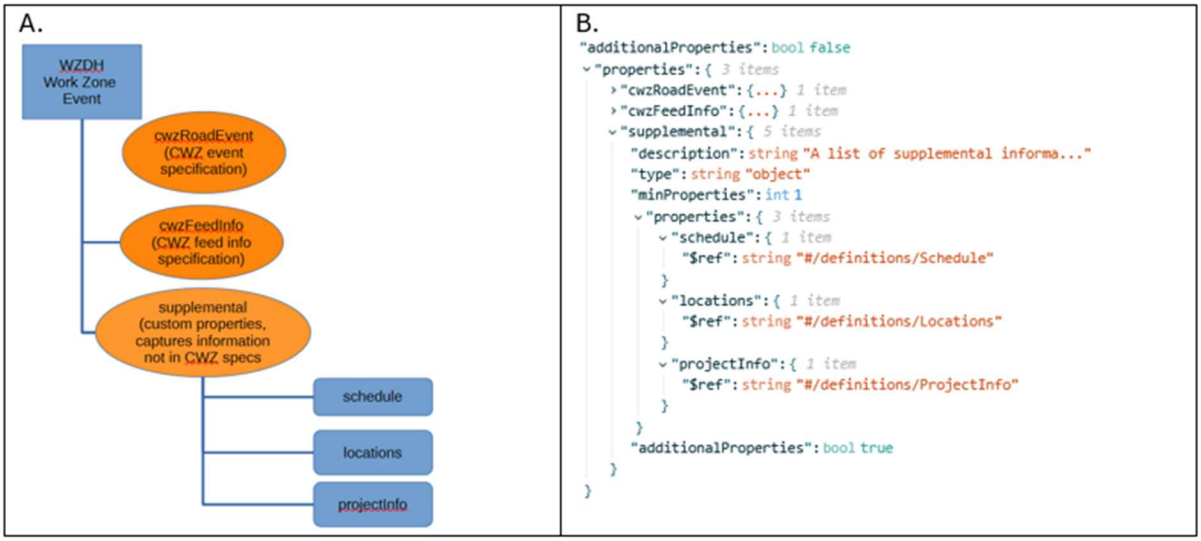


Figure 3. Minimal representation of the common data model represented as A. a block diagram and B. the JSON data schema

The structure of 'cwzRoadEvent' matches that of the CWZ specification. Optional supplemental properties such as the Schedule fully defined in the JSON specification document provide a space for properties not captured by the CWZ specification such as recurring schedules or construction activity by time of day. Such a structure stores and updates a work zone with a recurring schedule as a single entity from which a sequence of CWZ conforming events can be derived specific to the recurrence pattern.

As there is no expectation that the incoming data will have a schema provided, the schema of the source data can be inferred from the data itself. Automated tools inspect a sample of the data for a given source and deduce the structure and property types. Similarly, the common model is represented as a JSON data structure with the schema provided in Appendix 1. Finally the, transformation is defined for each source as a JSON rule set of key-value pairs where the keys correspond to the output structure and properties of the common object model and the values are the mapping to the properties in the source data using a custom syntax to allow descending into nested object structures. Documented post processing functions are triggered by a $<suffix> flag. For example, a property mapping followed by <property>edata will trigger a custom catch all date processor on the property value as it exists in the data feed such that any date representation (timestamp, text, …) can be appropriately converted to the format expected by later processes. Documentation for the translation and helper functions is available within the online user interface and also included in Appendix 2.

An interactive user interface has been developed to iteratively test and assure valid common model output of the translation representation. The interface has four columns to assist the user define the translation representation. From left to right in the interface shown in Figure 4 the columns show the common data model definition, either the data source schema an individual feature from the latest retrieval from the corresponding source, the translation definition as JSON key value pairs, and the

output of the translation which updates interactively as the user makes adjustments. A final step allows validation of the defined translation definition before persisting it to the database.

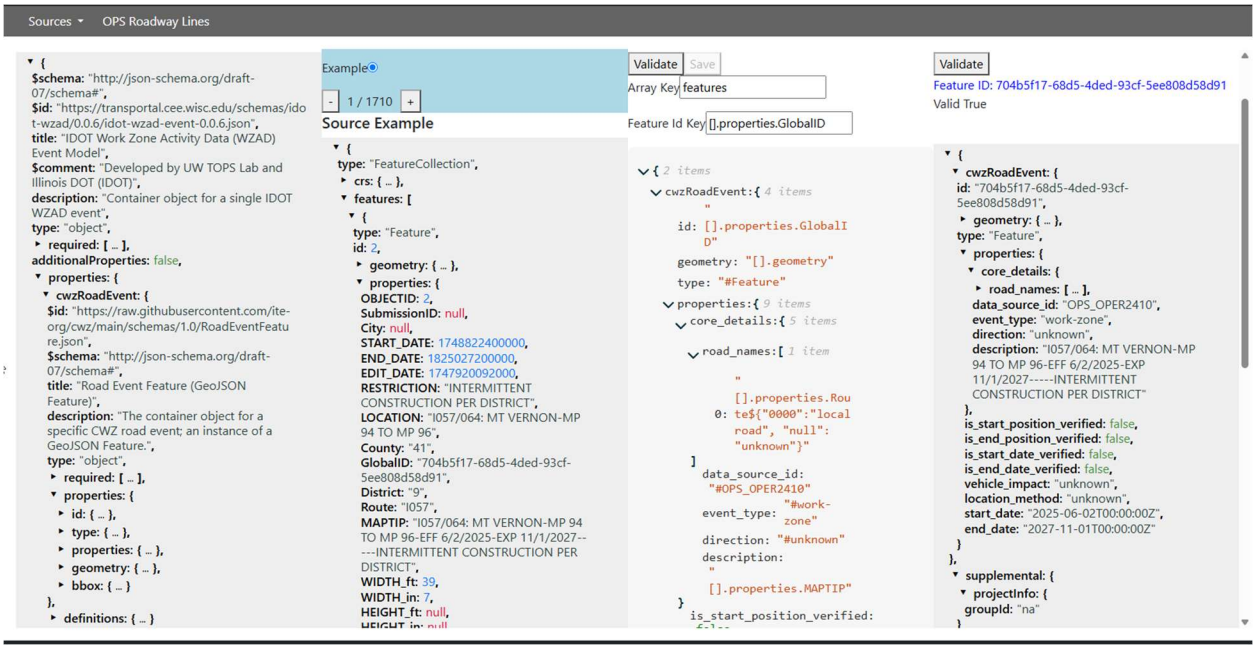


Figure 4. WZDA user interface for development and testing of source to common data model translation definitions.

As with conditional saving of archive records in cases where there has been a change, the same conditional approach is applied to individual road construction events. While the conditional archiving of data feed pulls applies to the source data pull as a whole – it will be different if any of the records has changed – conditional saving of the transformed records applies to the individual road work events. The common model transformation is applied to all records in the source data and saved to a temporary table. A process checks if a given roadway event exists in the transformed records table uniquely identified by an identifier in the source data feed. If it does and is identical in both geometry and properties to a record created by a previously seen data feed pull, no new transformed record is created. Otherwise – if the entity is new or has been updated – a new transformed record row is created uniquely identified by the construction event feature id and an integer record number which is zero for a newly seen feature or incremented by one if an existing feature was updated. The row contains the event information as the JSON common model and lifecycle tracking information such as the origin archive id and a reference to the transformation model that was applied.

Record Export

The flexibility of JSON makes it an ideal container for the common object model with its abundance of optional properties. However, its loose structure does not lend well to performance where a limited subset of required properties are needed or in cases where JSON data is not permitted. An export step was created to facilitate such use cases where a typical relational database table is more appropriate. An export table with a structure defined in the data dictionary in Appendix 3 is populated

by extracting the properties from the JSON common object model. The rows retain the lifecycle tracking information as with the transformed records table along with the columns that can be used by desktop GIS, ESRI map services supporting external applications, and in the interests of synchronization the same table is used to populate the CWZ feed described in the following step.

CWZ Generation

The export table is queried for rows matching the following criteria:

- CWZ flag is complete – this flag is populated during the transformation step and true indicates that the source data contains all required properties of the common data model such as start and end date and the work zone geometry
- The work zone is current – the system time is between the start and end datetimes for the work zone from the source data
- Current – the row is the most recent update for a given work zone event

The CWZ feed is populated from the relevant rows with the data shape conforming to the CWZ specification and ensuring that all required properties have valid values.

CWZ Source Data Snapshot

A snapshot of the IDOT WZDH data sources is provided in Figure 5, taken from June18, 2026 at 12:00 PM CDT. There are potential overlapping records between the various data sources, in particular the D1 LCS and ACTS data with respect to the OPS Roadway. Methods to identify and merge overlapping WZAD records are still under development.

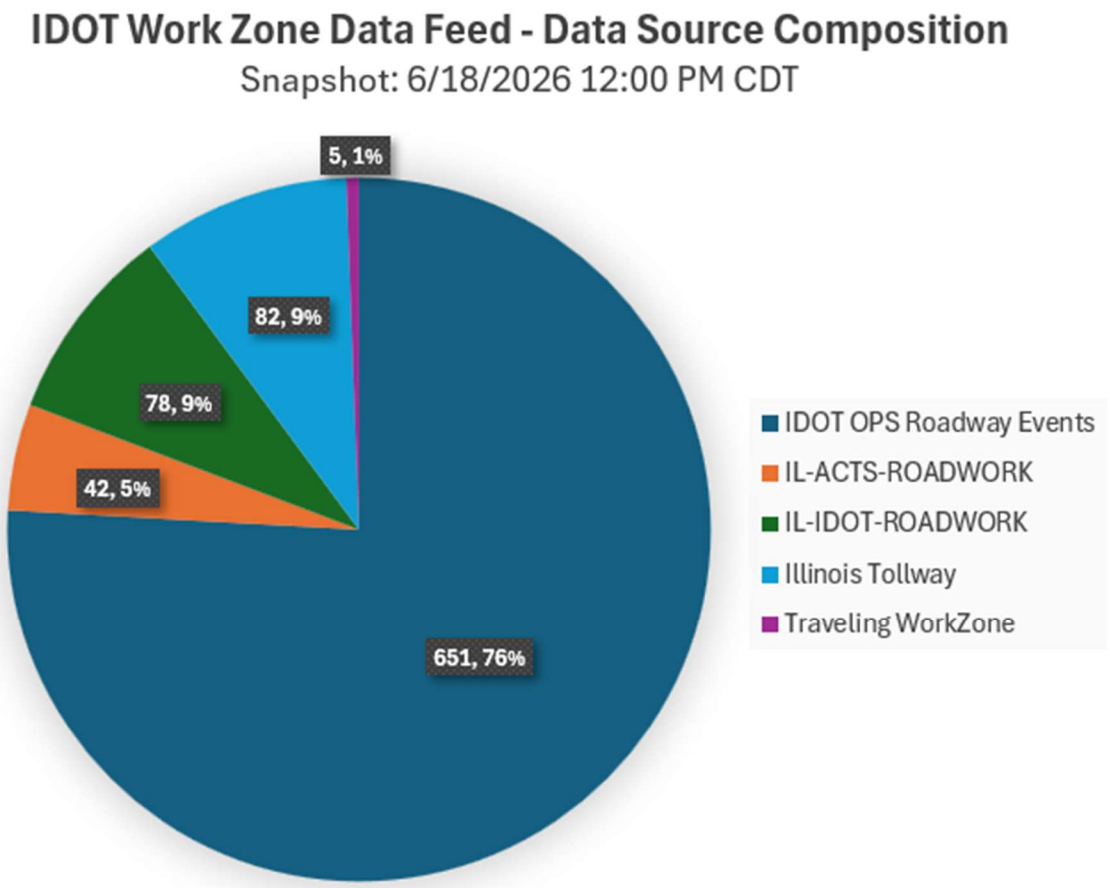


Figure 25. Snapshot of the IDOT WZDH data sources illustrating the overall contribution.

1. Data Model JSON schema – available in the ROSA P archive upon completion of the project.
2. Source to common model transformation documentation - available in the ROSA P archive upon completion of the project.
3. Record export table data dictionary - available in the ROSA P archive upon completion of the project.