# Building an Accident Report Data Base for Local Agencies 

James L. Lubkin, Thomas L. Maleck, David P. Relyea, and Kenneth H. Tsuchiyama


#### Abstract

A microcomputer accident report input and analysis system was adapted for the use of local police and engineering agencies. The Accident Report Entry System (ARES) conforms to current state rules for coding, internal consistency, and edit checks. As adapted, it permits agencies to enter their own traffic accident data from Michigan's standard accident report form, the UD-10. For the local jurisdiction, the ARES software also locates the accident on the Michigan public roadway system and determines a road number and milepoint. This task was previously done only on the Michigan State Police (MSP) mainframe computer. Historically, local agencies have had limited access to state-entered accident information. Using ARES, they can now enter their own UD-10 data to the same standards as those of the MSP. The information is available for study immediately, whereas mainframe data have recently been delayed for long periods and are not always in a form convenient for local users. Output reports equivalent to MSP standard analytical reports are already available locally. More advanced reporting capabilities similar to the sophisticated methods now in use at the Illinois Department of Transportation are being developed.


Michigan has long had a mainframe-based system for recording traffic crash data (1). This system is able to locate properly described accidents on all of its approximately 115,000 mi of public roads. The roadway data base is called the Michigan Accident Location Index (MALI) and is maintained on a Michigan State Police (MSP) mainframe computer. The index has a unique number for every road segment, as well as a milepoint for every intersection. The Michigan Office of Highway Safety Planning (OHSP) has supported the creation of microcomputer software to deal with several aspects of accident records. One group of programs permits the extraction of the subset for a county or city road system from the mainframe's statewide road index. This subset is almost always small enough to be manipulated effectively on a microcomputer.

A second series of programs for the Accident Report Entry System (ARES) allows a local police or engineering agency to create a local data base of traffic crash records for their own jurisdiction. These programs have been described elsewhere $(2,3)$. Extensive details of the software appear in two manuals $(4,5)$, one for the operator and another for the programmer. The operator sits at a personal computer (PC) and enters data from the UD-10 report, using screen forms similar to those used by the MSP. Some of the data are taken directly

[^0]from the UD-10, and other information is heavily encoded into a compact format. Encoding was needed formerly to save mainframe memory and disk storage, but it is hardly necessary today, even on small machines. Encoding has been retained primarily to conform to state procedures. In the future, the state might benefit from paying local agencies for each accident report entered according to proper standards; local police and agencies, of course, benefit from the immediacy of the data. The disadvantages are that operators need training in coding and have to enter data that have little value for local agencies, but they could be offset by state payments.

ARES abides by MSP rules and could replace the existing system if this were desired. The replacement would probably be best implemented as a local-area PC network, usually more economical today than a minicomputer. A study of the present system indicates that a considerable amount of paper handling could be avoided if ARES became a state system. However, all state institutions have adjusted to the present system, and much dislocation would result. From the viewpoint of standardization, quality, and training, having both the central data entry system and a local agency work with exactly the same software would be desirable. But combining the software may place an extra burden on local agencies that want to use ARES and may not need all of the data and cross-checks now incorporated for state purposes. The situation will probably become clearer after further experience.

One advantage of ARES is that the mainframe no longer has to be used at all. ARES makes location determinations right on the microcomputer. Accident reports are very clean before they are added to the local data base, because ARES furnishes on-line diagnostics and many checks before it accepts a record. Before the creation of ARES, the inflexibility of the prevailing system limited timely and convenient access by local agencies to their own accident data, and inhibited systematic analysis by these agencies. A plan for modifying the state system is being studied, in response to questions raised at Michigan's 1987 Governor's Conference on Highway Safety.

Local users must have a good reason to create their own data base, or the effort is not worthwhile. In addition to improving public safety, good data analysis can facilitate effective engineering and law enforcement countermeasures, and thus is a prime defense against tort liability. Local groups must learn how to use the information buried in accident data. Until now, the available accident data often arrived too late to be of much value in law enforcement or roadway improvement. The current pressure of tort liability judgments has become a powerful inducement for local agencies to learn how to use accident analysis in self-defense, and interest in
doing so is growing. Accordingly, a training program in accident data analysis is considered to be a vital follow-up to current project developments.

## PROJECT OBJECTIVES

The original project was altered after a reorganization of the MSP data processing unit and was refocused toward the needs of local users. Originally, the ARES system was intended to replace the central data entry procedures of MSP, and that objective can still be seen in the details following. At this time, many but not all of the following objectives have been achieved:

- Create a data input system for a local agency that conforms to the current state format, with all existing cross-checks and consistency checks. Create software that is as independent as possible of the particular microcomputer hardware used. Try to use the inexpensive equipment that is often all that local agencies can afford.
- Use commercially available data base management software, because it is relatively easy to use and good programmers are normally available for such packages. Furnish online operator assistance. Nonprogrammers and relatively inexperienced computer users should be able to learn the system quickly.
- Furnish analysis software so that timely action can be taken by local engineering and police personnel to make optimum use of their limited resources. Create a report system guided by the practices of major state agencies and other large users, but one that is adapted to the needs of local agencies.
- Explore graphical representations of accident data, such as collision diagrams, which make it easier for unsophisticated users to comprehend the significance of the data. Extend the system to integrate related functions, such as citations and court actions.
- Search out and use related projects in other states and incorporate their ideas where practical.
- Encourage the creation of training programs in the use of data entry and data analysis software.


## EARLY RESULTS

Tests of several objectives were carried out at the MSP central data entry location, at a time when a microcomputer network was being given feasibility tests as a replacement for the existing minicomputer system. At present, an assembly line of personnel go through processes called locating, coding, and data entry, accompanied by much paper handling. Only the data entry clerk actually sits at a computer terminal. The ARES system was designed to have all these functions performed by a single person at a single machine, entering data for a single county or city. Location was designed to take place on the microcomputer, to avoid the 20 to 25 percent report rejection rate on the mainframe, caused almost entirely by location difficulties. The early prototype software worked on the MS-DOS operating system using a version of the R:BASE commercial data base package. The microcomputers were IBM PC ATs or compatibles.

Roadway data bases are normally too large to fit into the 640 kilobytes of MS-DOS memory, and searching for a road in a data base on a hard disk takes ton much time for a production environment. The practical solution is to add large amounts of memory and place the entire data base in a virtual or random access memory (RAM) disk. Typically, 5 to 7 megabytes of RAM disk is needed for all but the largest counties or cities, to have room for both a roadway description network and a reasonable amount of accident data. The data base program itself sits in normal DOS memory as usual. The data base must be saved frequently on the hard disk (which is just a menu selection for the operator). As a conservative practice, the hard disk actually contains the current date's data base, another that is 1 production day old, and a third that is 2 production days old. In addition, each successfully added record is written out as pure for the American National Standard Code for Information Interchange (ASCII) data in another region of the hard disk.

In the prototype feasibility test, six operators were selected by the MSP supervisor from the existing MSP data entry staff. The operators had a range of experience with the present state system and were carefully selected to avoid using unusually poor or good performers. Over 1,000 accident reports were processed on the experimental system during a period of about a month, with the system being debugged as necessary. The average entry time for all operators over all reports was about 20 percent shorter on the prototype than on the existing system. Not much time was expected to be saved, because the same amount of diagnosis is needed for the existing and the experimental systems, and the same average number of keystrokes is needed for both. On the basis of the results, it was estimated that an experienced, fulltime operator could enter 18,000 to 20,000 Michigan-style reports per year, even with the early, slow, bug-infested software. Operator response was positive.

These results established the feasibility of carrying out the location process on a single PC (rather than exclusively on a mainframe), and of using a single person at a single machine to perform the necessary work. Current, bug-free versions of the software for data input are faster, more reliable, and kinder to the operator.

Part of the project was to test the software at a cooperating locality. The police and traffic engineering groups of the city of Battle Creek agreed to examine the value of entering and analyzing their own data. Battle Creek personnel were trained by the MSP in state data entry methods. A staff member of traffic engineering then entered about 1,000 accident reports over a period of several months. The software was continually updated in response to operator experience. Locating the accidents on a microcomputer produced the same results as on the mainframe, when the same input data were entered and the same roadway description network was used. Every discrepancy in location between the MSP mainframe and the Battle Creek microcomputer was accounted for. The software has now been adjusted to reduce those discrepancies, which can be attributed to operator choices. Further checking is in progress, but the project proceeds with confidence.

Practical application of ARES outside Michigan may be possible, depending upon how other locating systems for accidents (if any) function. ARES can be adapted to a variety of locating schemes. The real question is: "What data should be

```
        Michigan State University
Department of Civil and Environmental Engineering
Accident Report Entry System (ARES Version 1.0) 26 April }198
```

Please enter your name, or RETURN to quit: David

```
(1) Enter a new record
(2) Resume editing an incomplete record
(3) Correct an existing complete record
(4) List unlocated reports [Followup on operator override]
(5) Change operator name [Machine left on between operators]
(5) Save database (coffee break time)
(6) Quit production for the day (save database, exit to DOS)
```

FIGURE 1 R:BASE startup screen (top) and main menu (bottom).

| Accident Location Information | [Screen 1] |
| :---: | :---: |
| Reel 93 Frame 3112 Lein 11 | Compl 11-3692-87 |
| Accident Number 333333 Date 12/04/87 | Time 3 pm |
| County 33 Township 9 |  |
| Primary Road E I96 |  |
| Distance 0.2 Units (mi/ft) mi | Direction E |
| Intersecting Okemos |  |

REEL and FRAME refer to microfilm storage; COMPLaint is the local police report No.; ACCIDENT NUMBER is a State Police No.; the accident occurred on Eastbound I96, 0.2 miles EAST of the intersection with Okemos (Road) in Ingham County (33) Township No. 9.

| Accident Conditions | 333333 |  | [Rept. No.] | [Screen 2] |
| :--- | :--- | :--- | :--- | :--- |
| Area 999 | Pseudo | Rd Align 6 | Traffic 0 | Road Loc 3 |
| Acc.Type 3 Where 4 | How 0 | Tags 9 | Road Defect 0 |  |
| \#Units 2 | \#Parked 0 | \#Drivers 2 | \#Injuries 1 | \#Killed 0 |
| Weather 3 | Light 1 | Surface 2 | Const.Zone N | At. Scene Y |

AREA is "highway area type"; PSEUDO is used for failure to locate in the stated district; RD ALIGNment is "curved" vs. "straight" etc; TRAFFIC is for type of traffic control device or person; ROAD LOC is the road location code; ACC.TYPE WHERE HOW is a complex of codes describing and classifying the accident; TAGS is a special State Police accident designation; \#UNITS means No. of vehicles or cyclists or pedestrians; SURFACE refers to "wet" or "dry" etc.; AT SCENE means "investigated at accident scene".

FIGURE 2 Screens showing accident location information and conditions from Michigan accident form UD-10.


RESIDENCE (of driver); driver INTENT and OBJect FiIT; SITUATION and CONTRibuting CIRCumstances; SOS item is for Secretary of State; and the LICense has been SOUNDEX-encoded from the driver name and $\mathrm{DOB}=$ date of birth.


HAZardous ACTion is first classified by a police officer and perhaps differently by a coder (as SOS HAZ ACT); HBD = had been drinking; TEST gives alcohol test data; INJ O is "uninjured"; MAKE/TYPE/TRAILER classifies the vehicle; HAZ CITE = citation by officer; DRIVEABLE refers to the vehicle; IMPACT locates the point; SEVERITY classifies the accident; and RESTRAINTS encodes belt usage at up to 6 positions.

FIGURE 3 Vehicle and driver information screens.
entered?" The answer is probably different for different localities, and it is here that major departures from state and national schemes may be advisable. If ne appropriate scheme exists already, it would be helpful if NHTSA could develop a suggested scheme for nationwide data collection at the local level, focusing the ideas in the original Highway Safety Improvement Program (HSIP) of 1981 (6) on local needs.

## DETAILS OF ARES FOR INPUT

The ARES system uses the data base program called R:BASE for DOS on fast microcomputers based on the Intel 80286 CPU. This program is easily adapted to most local jurisdictions in Michigan, but an 80386 machine may be needed for the largest regions. Some operator training will always be necessary if data entry is to conform to state rules and standards. Local units of government may or may not actually
need to enter precisely the same data that the state as a whole may wish to collect.

Figures 1 through 12 show a subset of the screens and menus that the operator sees, just to give the flavor of the system. Many have been omitted to save space. The displays normally fill the video screen, but they have been reduced in size. Any items in square brackets are for clarifying comments. Figure 1 shows the first screen seen by the operator after R:BASE starts and the main menu screen.

Selecting Item 1 from this menu results in a series of screens. Screens 1 and 2 appear once per accident report form (see Figure 2). All items come directly from Michigan accident form UD - 10. Most of the top half of Screen 2 contains heavily encoded information for the Michigan Department of Transportation (MDOT) and the MSP, requiring operator training and judgment to enter it.

Screens 3 and 4 in Figure 3 are for vehicle and driver information, and Screen 5 in Figure 4 is for details of nondriver

|  | Injury | Data |  | 333333 |  | [Screen 5] |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Unit | Pos | Age | Sex | Inj | Helmet |
| Injury 1 | 1 | 3 | 21 | F | A | N |
| Injury 2 | 0 | 0 | 0 |  |  |  |
| Injury 3 | 0 | 0 | 0 |  |  |  |
| Injury 4 | 0 | 0 | 0 |  |  |  |
| Injury 5 | 0 | 0 | 0 |  |  |  |
| Injury 6 | 0 | 0 | 0 |  |  |  |
| Injury 7 | 0 | 0 | 0 |  |  |  |
| Injury 8 | 0 | 0 | 0 |  |  |  |

The injured person is in POSition 3, with disabling INJury, type "A".

Choose report to RESUME, or press ESC for the main menu.

| 3688 | Unable to locate |
| ---: | :--- |
| 27564 | UD-10 sent back for correction |

User moves cursor keys and presses Return/Enter to select a highlighted report (cannot be simulated here) for RESUMEd editing.


FIGURE 4 Nondriver injury data (top) and resume editing (bottom) screens.

LOCATE Version 2.0 of 30 January 1990
--------------------------------------------1
Locating WASHINGTON intersecting with MAIN
Searching for primary road: Found
crossroad: Found

| Mile- <br> point | Primary |  | Road |  |
| :--- | :--- | :--- | :--- | :--- |
| PRN | Name | Crossroad |  |  |
| $\mathbf{3 . 2 1}$ | 1234567 | WRN | Name |  |
| 0.17 | 7654321 | WASHINGTON DR | 4444444 | MAIN AVE |

Cursor highlights a row in the menu above (not visible in this simulation). Operator moves highlighted row with cursor keys and selects with Return/Enter key: resolves roadway name ambiguity.

| SOUNDEX | Entered |  | Computed |  |
| :--- | :--- | :--- | :--- | :--- |
| First Letter | B |  | B | OK |
| Last Name | BLOW | 350 | 350 | OK |
| First Name | JOE | 450 | 450 | OK |
| Middle Name | QUINCY | 546 | 548 | BAD |
| Birth Date | 05/10/66 | 357 | 357 | OK |

FIGURE 5 Example of a correctable Soundex check that did not work the first time.


## ARES Analysis 1.0 -- MSP Standard Reports

(1) Intersection Search
(2) Segment Search
(3) Intersection High Rank
(4) Return to the Main Menu
[Abbreviation PRN below is for the official MALI road number.]

```
Michigan State Police Standard Report
Intersection Search
```

```
SELECTING ROADS BY NAME OR PRN
When selecting a road, you may either enter the road name,
or the PRN preceded by '#'. For example, if you wish to
select by the name, at the prompt you would enter:
    COLUMBIA
To select by PRN, at the prompt you would enter:
    #1297108
For Absolute Township, use 1000 * (county no.) + (twp no.)
```

Enter Absolute Township (RETURN to include all townships): 13080
Enter Main Road: \#1298109
Enter Crossroad (or RETURN for menu): DICKMAN
Enter Intersection Search Radius (RETURN for none): 125

FIGURE 6 ARES startup (top), MSP standard reports (middle), and intersection search (bottom) menus.
injuries. In Screens 3 and 4, the data entry operator must again encode many items. Here Screen 5 is for an injury in Unit 1.

Selecting Item 2 in the main menu of Figure 1 results in the resume screen of Figure 4. If main menu Item 3 is selected (see Figures 1 and 4), the "Select an option" menu displayed there appears.
Figure 5 shows the screen seen during the ordinary processes of location, after Screen 1. This form of display helps resolve roadway name ambiguity.

Soundex checking (Figure 5), which is part of the ARES system, strictly for the U.S. Department of State, verifies whether the driver license numbers agree with the date of birth and driver name. While the Soundex routine is calculating what the numbers should be from the name and birthdate, these numbers and the original entries appear on the screen. If no problem is discovered and all items are the same,
the screen clears and the program continues. If a problem occurs and one or more of the numbers do not match, the numbers stay on the screen until the operator presses a key. Thus, the numbers can be examined and the difficulty can be corrected. Figure 5 shows an example of a (correctable) Soundex check that did not work the first time.

The many other screens and error messages must be omitted here. Operators and users have a reference and training manual (4), which has more than 220 pages and includes most of the instructions for MSP coders. The Programmers Manual (5) is of comparable size.

## FEATURES OF ARES FOR OUTPUT

At present, MSP and MDOT furnish various standard reports to interested users, based on data in MSP's master accident
Michigan State Police Standard Report Intersection Search Select road for DICKMAN

R E P O R T PARAM E T E R S
ARES Analysis Version 1.0 15 May 1989


FIGURE 7 Intersection search pop-up menu (top) and report parameters (bottom).
tapes. These tapes mainly contain the data entered by the central-site operators, but they also include the location data created by the MSP mainframe. A few other calculations are performed, but the output is almost the same as the input. User groups who have examined the matter believe that only input data (plus computed location) should be stored. Only then can users know the precise meaning to ascribe to every data field. The philosophy of ARES output follows this idea of storing and analyzing raw input data alone.

ARES input is first converted to ARES output by a program devised for that purpose. It changes many numerical codes, such as 2 or 14 , into intelligible items like Rain or Snow, or into abbreviations like FaiYi, for Failure to Yield. The output data base omits the tables that are needed for input but have no purpose during output.

The user who starts ARES output sees a startup menu (Figure 6). Here, Option 1 delivers reports adapted from
those now provided by MSP or MDOT. These reports are being used at the cooperating agencies in Battle Creek. They consist of regular R:BASE reports, speeded up by interfacing R:BASE with C-language utilities for which R:BASE cannot do what is necessary quickly enough. Option 2, still undergoing development, is much more advanced. These reports incorporate some of the ideas used by the Illinois Department of Transportation (DOT) but require additional tables of data, which have only just been obtained. These extra data consist of traffic volumes and elements of roadway geometry. Custom reports, Option 3, must be crafted by those who know the R:BASE system well and can create ad hoc reports for themselves. After choosing Option 1 of Figure 6, the user receives the second menu shown.

If the user selects the intersection or segment searches from this menu, the last screen of Figure 6 appears and the user is prompted for data. When entering a road name, such as Dick-



FIGURE 9 Page two of actual intersection search and summary from partial-year Battle Creek data.

| DRIVER PROFILE |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Drinking |  |  |  | RESIDENCE |  |  |  |
| Age | or | Drugs | Cou- | In- | Out of | Driver |  |  |
| Group | No | Yes | nty | State | State | less | Other | TOTAL |
| 16-25 | 5 | 0 | 4 | 1 | 0 | 0 | 0 | 5 |
| 26-55 | 9 | 1 | 8 | 0 | 0 | 0 | 0 | 10 |
| 56-98 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |
| 99 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| TOTAL | 16 | 1 | 13 | 1 | 0 | 0 | 0 | 0 |

FIGURE 10 Driver profile of actual intersection search and summary.
man in the sample screen, a pop-up menu displays all the Dickmans in the data base of the townships specified, for the user to select. This list is shown in Figure 7. The user can now edit report parameters, as in the segment search in the lower part of Figure 7. A highlighted cursor (not visible here) permits selection among the choices. EDIT allows changes in the data, CLEAR removes all entries, QUIT or ESC quits, and PRINT pops up the printer control menu (not reproduced here). An actual intersection search and summary using partial-year Battle Creek data is shown in Figures 8-10.

A sample segment search is shown in Figure 11, with the summary portions shown in Figures 8-10 omitted. Figure 12 shows the leading portion of the intersection high rank report, which does not change often or need to be run frequently. This last standard report does not allow for traffic volumes and other factors. Unfortunately, this version of ranking is quite common and is often all that is available; it will be improved as soon as the Illinois DOT software can be introduced.

Option 4 of the main menu (examine road network, Figure 6) makes use of the local roadway network, which is already in the data bases for both input and output. Using this branch, all road names corresponding to a given road number, such as 1297108 , or all road numbers that go with a given name (like Main St) can be retrieved. For every road number, the milepoint, road name, and road number corresponding to every intersection on that road can be inspected. The other main menu options are self-explanatory.

## FUTURE PLANS

The City of Battle Creek continues detailed testing of input and output software. The sophisticated analysis scheme used by the Illinois DOT for its trunklines is being incorporated into the Battle Creek project. This incorporation requires traffic volume and geometry data, only recently obtained. As soon as good software has been thoroughly tested, other local agencies in Michigan will be able to use the entire ARES system. A significant number of agencies are known to be
interested. Considerable thought will be given to reducing the data entered to that really needed by local agencies. This reduction may offer savings of entry time and operator training and should be explored if it does not impinge on data quality and adequacy.

Other aspects of a complete Highway Safety Improvement Program will be integrated gradually into the ARES system. Police citations, court records, and other factors that affect traffic safety probably have a place in the overall picture. If the project proves to be successful, short courses and training will also need to be devised as a follow-up.

## CONCLUSION

An accident data system like that now used by the MSP can be implemented on a microcomputer, for a jurisdiction as large as an above-average county in the state, or a city of comparable size. The ARES software is able to extract an appropriate subset of the statewide road network (MALI Index). This permits ARES to locate accidents on public roads without requiring a mainframe or even a minicomputer. Thus, local engineering or police agencies can use an inexpensive machine to enter their own accident data and to analyze it immediately with the software now being refined. In turn, this should help to improve both public safety and the tort liability problems of many local jurisdictions.

## ACKNOWLEDGMENTS

This paper was prepared in cooperation with OHSP and NHTSA. The authors wish to acknowledge the guidance provided by James Rojeski and Lloyd Stoner, formerly of the MSP Traffic Services Division, and their staff; by the staff of the MSP Criminal Justice Data Center; and by Jack Benac, Don Mercer, Jim Neve, and Joe Meszaros of MDOT. The cooperation and advice of the Battle Creek Police and Traffic Engineering departments are also hereby gratefully acknowledged.


FIGURE 11 Sample segment search.


FIGURE 12 Leading portion of intersection high-rank report.

## REFERENCES

1. R. L. Blost. Michigan Accident Location Index (MALI). In National Safety Council Forum on Traffic Records Systems, March 1976.
2. J. L. Lubkin and T. L. Maleck. ARES: an Accident Report Entry System for Local Agencies. Proc., 3rd International Conference on Microcomputers in Transportation, San Francisco, Calif., ASCE, New York, 1989.
3. T. L. Maleck and J. L. Lubkin. ARES: an Accident Report Entry System for Local Agencies. Presented at 15th International Forum on Traffic Records Systems, National Safety Council, El Paso, Tex., 1989.
4. J. L. Lubkin and T. L. Maleck. Operators/Users Manual For ARES (Accident Report Entry System). Department of Civil and Envi-
ronmental Engineering, Michigan State University, East Lansing, 1989.
5. J. L. Lubkin and T. L. Maleck. Programmers Manual For ARES (Accident Report Entry System). Department of Civil and Environmental Engineering, Michigan State University, East Lansing, 1989.
6. Highway Safety Improvement Program. Report FHWA-TS-81-218. FHWA, U.S. Department of Transportation, 1981.

The opinions, findings, and conclusions expressed in this paper are those of the authors and not necessarily those of OHSP and NHTSA.

Publication of this paper sponsored by Committee on Traffic Records and Accident Analysis.


[^0]:    J. L. Lubkin, T. L. Maleck, and D. P. Relyea, Department of Civil and Environmental Engineering, Michigan State University, East Lansing, Mich. 48824. K. H. Tsuchiyama, Department of Public Works, City of Battle Creek, P.O. Box 1717, Battle Creek, Mich. 49016.

