A Primer on Source Code and Its Role in Elections
By David Wagner

What is source code? Source code is the human-readable representation of the instructions that control the operation of a computer. Computers are composed of hardware (the physical devices themselves) and software (which controls the operation of the hardware). The software instructs the computer how to operate; without software, the computer is useless. Source code is the human- readable form in which software is written by computer programmers. Source code is usually written in a programming language that is arcane and incomprehensible to non-specialists but, to a computer programmer, the source code is the master blueprint that reveals and determines how the machine will behave.

Source code could be compared to a recipe: just as a cook follows the instructions in a recipe step-by-step, so a computer executes the sequence of instructions found in the software source code. This is a reasonable analogy, but it is also imperfect. While a good cook will use her discretion and common sense in following a recipe, a computer follows the instructions in the source code in a mechanical and unfailingly literal way; thus, while errors in a recipe might be noticed and corrected by the cook, errors in source code can be disastrous, because the code is executed by the computer exactly as written, whether that was what the programmer intended or not. Also, computer software is vastly more complex than most recipes: while a typical recipe may contain perhaps a dozen steps and fits onto a single 3x5” index card, computer source code often contains hundreds of thousands of steps which, if printed, would fill up thousands of single-spaced 8.5x11” sheets of paper. What does source code have to do with elections? Over the past several decades, as we have automated more and more of elections operations, elections have become increasingly reliant upon computing technology. For instance, touchscreen voting machines use computers to capture votes; paper ballots are scanned using computer-driven scanning machines; and computers tabulate and tally the votes to determine the winner. This makes the software that controls these machines of critical importance to our elections.

The source code in voting machines is in some ways analogous to the procedures provided to election workers. Procedures are instructions that are provided to people; for instance, the procedures provided to poll workers list a sequence of steps that poll workers should follow to open the polls on election morning. Source code contains instructions, not for people, but for the computers running the election; for instance, the source code for a voting machine determines the steps the machine will take when the polls are opened on election morning.

Who writes election-related software? Today, counties and states buy voting equipment from commercial vendors. These voting system vendors write most of the software in their machines. However, voting system vendors also incorporate software from third-party software vendors into their products. For instance, a voting system vendor like Diebold might license software from Microsoft for use in their touchscreen voting machine. The voting vendor might or might not receive source code to the third-party software; if they do, they normally would not have permission to re-distribute this third-party source code to others. Third-party software is sometimes called COTS (commercial off-the-shelf) software.

Who sees election-related source code? Today, most voting system vendors treat any source code they write as confidential and proprietary. The vendors tightly control access to this source code. Election officials use the equipment, but they are normally not given access to its source code. Candidates, political parties, technical experts, and interested citizens are normally not given access to voting system source code, either. Federal voting standards require voting system vendors to share their source code with a testing laboratory selected by the vendor, and the testing labs are supposed to check that the system complies with the federal standards. However, the testing labs have come under growing criticism for missing security and reliability problems in deployed voting systems, and many experts have expressed concerns about the ability of the testing labs to ensure that voting systems are fit for use[1] [2]. Most states do not receive or require access to voting source code. However, there are some exceptions[3]. Five states appear to require source code for certified voting systems prior to their use (FL, NY, TX, UT) or have the authority to demand source code at their discretion (CA). Two states go farther and require that the vendor provide source code to representatives of the major parties upon request (NC, MN).

[1] This work was supported by the National Science Foundation under Grant No. CNS-052431 (ACCURATE). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

[2] I do not speak for UC Berkeley, ACCURATE, the California Secretary of State, the EAC, the TGDC, or any other organization. Affiliations are provided for identification purposes only.