a paragraph sign, ¶, I used a dollar sign, $.

My method of making a modern computer transcription of item 1609 was as follows. The Marshall Library provided me with two xerographic copies of all of 1609, one of which I supplied to a volunteer, Jacques Guy. Working independently on separate continents, by different methods, we prepared two computer transcriptions of 1609. Then the two transcriptions were compared and all discrepancies resolved.

One of us (Guy), typed about two thirds of the text into his computer the ‘‘old fashioned’’ way.

The other (Reeds) tried to use modern optical character recognition (OCR) methods to scan his copy into the computer. In consultation with Henry Baird, an OCR expert, it became clear that the variations in inking of the original and in the quality of the photocopies, as well as the extensive pencil markings on some of the pages of the original, would make the output of an automatic OCR run unusably inaccurate, even if the OCR algorithm [Ba] were trained on the document itself. Instead, we devised a semi-automatic scheme, where the computer assigned a preliminary guess at the values of each of the printed characters on a page. Then the computer displayed images of all characters assigned value ‘‘A,’’ say, and the human operator could quickly spot and correct misassignments by using mouse clicks. As bogus As were reassigned, their images disappeared from the screen, leaving behind a ‘‘purer’’ field of As in which it was ever easier to spot misassignments. When all the putative As were indeed visibly As the computer then displayed all putative Bs, and the process repeated. At the end, the computer (which of course remembered where on the page each of the characters was printed) wrote a file showing the resulting reading of the given page of 1609. By using this technique, I did not see characters in context, which had both good and bad consequences. On the one hand, contextual clues for guessing at faintly printed letters were absent, so the error rate for such letters was elevated. On the other hand, errors which might have been introduced by paying less attention to a character than to its neighbors (analogous to unconscious correction of spelling mistakes in transcribing ordinary text, as well as transposition errors) were presumably avoided.

The results of the comparison of the two transcriptions are are as follows. In all, 113,366