Kurzweil
Intelligent Scanning System

Intelligence: The essential difference in text and data entry.

Kurzweil 4000: the only system that can enter virtually any typed or printed material into computerised files, without manual keying.

Using Intelligent Character Recognition (ICR) technology developed by Kurzweil, the 4000 overcomes the scanning limitations of conventional OCR systems. It can scan virtually any typeset or typewritten document, including proportionally spaced copy. It reads mathematical, foreign language and scientific symbols. It differentiates between multiple fonts within a single document. And through the use of the system's electronic tablet, portions of text can be separated from surrounding material for selective automatic entry.

In addition, specialised features make the 4000 uniquely suited to the demanding needs of printing and publishing firms. These features include the ability to identify and code font changes, and distinguish such characters as left and right quotation marks, hyphens and dashes, and horizontal spacing. Taken together, they provide a decisive edge for the 4000 over any other system.

A price/performance breakthrough.

Now virtually any organisation can afford the benefits of intelligent scanning. The 4000's purchase price is less than half that of previous systems, yet its performance is significantly enhanced. It includes a higher-speed processor. It can train itself in the recognition of simple formats and type fonts. Operator prompts and menus reduce training and minimise human error. And its more compact size enables the 4000 to fit easily into office environments.

And yet, for all of this, the 4000 uses the same design approach that has stood the test of time in over 700 user installations. Even in high-volume, high-demand production environments, the 4000 can be depended on for round-the-clock performance. And it's backed by nationwide support and service.

How the 4000 works.

The 4000 uses Kurzweil's exclusive artificial intelligence software to recognise and read copy. As it scans a document, the 4000's artificial intelligence analyses the shape of each printed character, and using feature extractions, identifies it as a letter, number, or other symbol. This process — Intelligent Character Recognition — enables the 4000 to enter text up to 10 times faster than the best human keyboard operators.

Operation is simple and straightforward. The system's ergonomically designed workstation includes an adjustable screen which minimises eye fatigue and enhances productivity. And the 4000's powerful software includes a comprehensive set of menus and prompts that guide the user through each operation.

System startup is accomplished quickly, as the 4000 scans the first few lines of a new document and — together with the operator — establishes guidelines for the interpretation of specific characters.

When broken characters, smudged type, or other problems make recognition uncertain, the system flags the character for operator assistance, displays an enlarged version of the character and asks for an identification decision.

With reasonably clean copy, this training process is completed in a matter of minutes — at which point production scanning begins. During the production phase, operator intervention is far less frequent, and mostly concerned with guiding the system through columns of text or other graphic formats.

The 4000 enables you to scan the thousands of inevitable variations that occur in most documents. For example, the 4000 "knows" to split two characters that are touching, or combine the parts of a broken character. It can make decisions, in context, between numbers and letters, opening and closing quotes, hyphens and dashes, and many other ambiguities.

Options

A wide range of peripheral and software options assure maximum throughput and adaptability to specific text or data entry applications. Standard or customised output communications provide links to typesetting and publishing systems, word processing systems, office automation networks, information retrieval systems, and other computer-based systems.

- Bisynchronous communications controller.
- Second programmable asynchronous communication port.
- Electronic tablet — For specifying material to be scanned on a document. Available in U.S. and international document size.
- Automatic document feeder — Handles up to one-half inch stack of pages in one feeding, document size of 8½" wide and up to 14" long, and paper stock from 20 to 50 lbs.
- Licence for custom format interfaces — Automatically translates and transmits format codes for paragraph, page, underline, superscript and subscript, margins, and tabs (indent, left/right, and decimal), centered lines, "soft hyphens," page headers and footers. Custom interfaces available for word processing systems, office network systems, and front-end typesetting systems.
- Licence for additional language packages — To facilitate proofreading of text during production and to decrease actual training time. Language packages are available in English, German, Dutch, Swedish, Italian, Norwegian and Danish.
- Magnetic tape drive — Dual density (800/1600 bpi) 9-track magnetic tape drive with 45 ips mechanical arm transport.

(Left) Conventional OCR is severely limited in the number and variety of type fonts and formats it can handle. ICR technology "reads" virtually all type styles and formats, including proportionally spaced type — radically enlarging the possibilities for automated text entry.

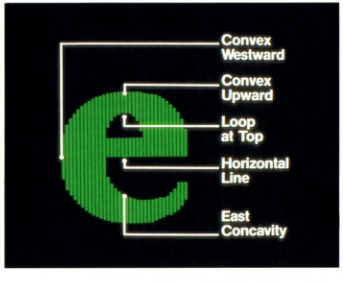
(Right) Kurzweil's proprietary artificial intelligence software enables the 4000 to "read" each character by analysing its shape. The 4000 can recognise any type font in sizes ranging from 6 to 24 point, including multiple fonts within a single document.

Kurzweil

Kurzweil
13 11 7 9 14 10 5 5

Kurzweil
13 10 7 9 15 9 5 5

Kurzweil
14 11 8 9 15 9 5 5



(Left, Top) Optional foreign language lexicons greatly enhance the 4000's ability to read most languages using the Roman alphabet.

(Left, Centre) Broken or irregularly formed characters are highlighted on the system's screen to assist the operator in verifying identification. Users can also edit text as it scrolls on the screen.

(Left, Bottom) Ligatures, or joined images, can be split to output the individual characters quickly and easily.

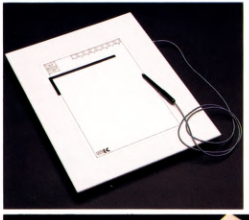
(Right) The Kurzweil 4000 features an ergonomically designed workstation for maximum operator efficiency and productivity.

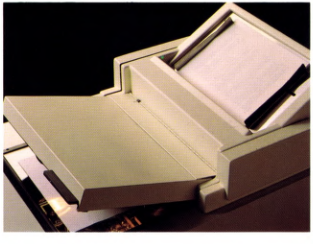
Kurzweil 4000 Product description

(Left, Top) With optional electronic tablet, portions of text can be separated from surrounding material for selective automatic entry.

(Left, Bottom) Optional automatic document feeder handles sheets up to 8½″ × 14″

(Right) Optional tape drive enables additional flexibility in communication with mainframes.





Recognition capability

- Most serif and sans-serif fonts (script excluded) 6–24 point* including Roman and Italic.
- Uniformly or proportionally spaced material (typed or typeset).
- Any line spacing providing lines do not touch.
- Accents, multi-part characters and ligatures.
- Letter quality dot-matrix (9 × 14 dots per character).*

* *Recognition of various point sizes is dependent upon character quality and spacing of material being scanned.*

Input

- Light-coloured paper of any weight.
- Pages up to 11″ wide by 14″ long.
- Black printing ink or carbon-based colours.
- Bound books, manuals, publications.
- Clean photocopies.
- Maximum skew of 3/8″ over 7″ on 10 point type.
- Multi-column or complex page formats (with electronic tablet option).

Operational features

- Flags questionable characters for operator attention.
- Learns most fonts in most document formats.
- Preserves horizontal spacing.
- Automatically learns fonts in single-font document.
- Operates in production mode while learning.
- Learns on horizontally fragmented characters.
- Recognises true superscripts and subscripts.
- Allows user-defined 2-3 character mnemonics for special symbols.

On-line editing features

- Highlighting of questionable characters.
- Scanner image display for operator verification.
- Insertion, deletion, and replacement of characters, words, lines.

Special features

- On-line storage of multiple training sets.
- Over 400 unique character definitions per training set.
- Editing of character definitions within training set.
- Library of training sets can be saved off-line on floppy disk.
- Operator adjustable contrast and background control.
- Ambiguous character analysis.
- On-line lexicon to reduce training time and manual editing.
- File management system for creating, copying, editing, appending, and deleting files, with throughput log, back-up utilities, and help feature.

Standard communications/ output tailoring

- Operator-generated character tables (ASCII, EBCDIC, TTS tables supplied with system).
- Character string translations/replacements.
- On-line storage of output tailors.
- Optional flagging of paragraphs.
- Protocols supported:
  - –IBM 2780 standard (bisynchronous) — transmit only, point-to-point (transparent or non-transparent)
  - –Asynchronous –TTY/Echoplex
  - –800/1600 bpi magnetic tape
- Ability to review text file on CRT prior to output.
- Automatic transmission of prefix and suffix files.
- Character count at output.
- Transmission of single files or batch mode.
- Tape output options include: variable or fixed length records, block size options and provision for IBM standard labels.
- Optional flagging of questionable characters.
- Maintains or suppresses line and page breaks.
- Optional flagging of indentations.

Standard system components

- Workstation including high-performance 384K processor scanner subsystem, 10mb hard disk drive and floppy disk drive.
- Operator terminal.
- Licence for use of recognition operating system.
- Licence for use of English language package.
- Programmable asynchronous communications port.
- Text editor — Enables operators to inspect and edit text on the 4000 before it is communicated to another device.

Environmental specifications

- Electrical requirements: 240V, 50Hz, 6A
- Dimensions: 24″ × 22″ × 29⅜″
- Weight: 375 pounds
- Temperature: recommended operating, 60-80°F
- Humidity: 50%–80% non-condensing

KURZWEIL COMPUTER PRODUCTS
A Xerox Company

Kurzweil Computer Products, Inc.
First Floor, 2-10 Whitchurch Road, Pangbourne, READING RG8 7BP
Telephone Pangbourne (07357) 5258 Telex 849910 BEBECE G