

Readiris[™] PRO 10

MANUALE DELL'UTENTE



Readiris Corporate

© 2005 I.R.I.S. Tutti diritti riservati

Tecnologia OCR di I.R.I.S.

Tecnologia Connectionist, AutoFormat e linguistica di I.R.I.S.

Tecnologia ICR e lettura codici a barre di I.R.I.S.

Tecnologia BCR e tecnologia di analisi dei campi da I.R.I.S.

© 2005 I.R.I.S. Tutti diritti riservati

RISPARMIARE TEMPO SENZA RIDIGITARE IL TESTO

Congratulazioni per l'acquisto di Readiris. Questo pacchetto software sarà indubbiamente di grande aiuto per acquisire testi, tabelle, elementi grafici, biglietti da visita, codici a barre e anche testo scritto a mano.

Efficiente quanto un computer, richiede innanzitutto l'immissione da tastiera delle informazioni. Tutti sanno che ridigitare una relazione di 15 pagine o un'ampia tabella numerica può essere noioso e dispendioso in termini di tempo. Questo pacchetto OCR, che rappresenta lo stato dell'arte, consente di inserire automaticamente testi nelle proprie applicazioni e acquisire un livello di efficienza e produttività senza precedenti.

È sufficiente eseguire la scansione di un documento stampato o dattiloscritto, indicare le aree d'interesse o lasciare che il sistema lo faccia automaticamente, quindi effettuate il riconoscimento dei caratteri. I documenti composti da diverse pagine vengono elaborati dall'inizio alla fine in un'unica operazione. Pochi clic del mouse compensano lunghe ore di lavoro, poiché Readiris converte i documenti cartacei in file editabili, impiegando tempi fino a 40 volte inferiori a quelli necessari per la normale ribattitura.

La procedura guidata ha lo scopo di guidare l'utente lungo l'intera procedura di elaborazione OCR: rispondendo ad alcune semplici domande si otterranno facilmente e rapidamente risultati eccellenti con Readiris. Sarà possibile inviare i risultati direttamente al proprio editor di testo o a un foglio elettronico. Per il riconoscimento di fax e la conversione di documenti PDF, è sufficiente trascinare il file Esplora Risorse di Windows alla finestra di applicazione Readiris o fare clic con il pulsante destro del mouse sull'immagine per inviarla immediatamente a Readiris.

Readiris riconosce i dati tabulari e li riorganizza in fogli di lavoro o come oggetti di tabelle all'interno di un editor di testo. I dati numerici sono immediatamente pronti per un'ulteriore elaborazione.

Basato sulla tecnologia Connectionist di I.R.I.S., Readiris rappresenta quanto di meglio l'OCR può offrire. L'estrazione di caratteristiche indipendenti dei caratteri



è completata da tecniche di autoapprendimento derivanti da una rete brevettata neurale. Mediante l'analisi del contesto il sistema è in grado di apprendere nuovi caratteri: la conoscenza linguistica di sillabe e parole migliora le prestazioni di OCR.

Readiris supporta fino a 117 lingue: supporta tutte le lingue americane ed europee, incluse le lingue dell'Europa centrale, le lingue baltiche, il greco e le lingue cirilliche ("russo" e simili). In alternativa, è possibile leggere documenti ebraici e quattro lingue asiatiche: giapponese, cinese tradizionale, cinese semplificato e coreano. Readiris è in grado di gestire anche testi in più alfabeti: il software individua le parole "occidentali" che compaiono nei documenti in lingua greca, cirillica, ebraica e asiatica, nelle quali numerosi nomi propri non traslitterabili, marche ecc. vengono scritti usando le lettere dell'alfabeto latino.

Readiris utilizza criteri linguistici *durante* e non dopo la fase di riconoscimento. Come diretta conseguenza, Readiris riconosce con estrema precisione tutti i tipi di documenti, inclusi documenti di basso livello qualitativo, fax e stampe di matrice di punti. Gestisce documenti acquisiti o copiati di qualità scadente nei quali i caratteri sono troppo chiari o troppo scuri. I caratteri uniti vengono scomposti e le forme frammentate, come ad esempio i simboli di matrice di punti, vengono ricomposte.

La verifica che l'utente può eseguire mediante i messaggi a comparsa ("pop-up") non solo consente di individuare i caratteri dubbi, ma anche di aumentare la precisione del sistema. Tutte le soluzioni confermate dall'utente vengono memorizzate con velocità e sicurezza sempre maggiori. Utilizzare Readiris significa renderlo ogni volta più intelligente. Questo potente strumento di apprendimento fa in modo che Readiris apprenda caratteri particolari come, ad esempio, simboli matematici e segni grafici decorativi, ma che consenta anche di gestire caratteri distorti presenti nei documenti.

Al fine di aumentare ulteriormente la produttività, Readiris non esegue solo il riconoscimento dei testi, ma è anche in grado di *formattarli* in modo automatico. Grazie alla "formattazione automatica" Readiris ricrea una copia facsimile del documento acquisito, mantenendo la formattazione della parola, del paragrafo e della pagina del documento originale.

Durante il riconoscimento vengono utilizzati tipi di caratteri simili e mantenuti gli stili tipografici e le dimensioni dei caratteri presenti nel documento di partenza. Il posizionamento di colonne, blocchi di testo ed elementi grafici rispetta i documenti originali. Readiris è inoltre in grado di inserire nel documento riconosciuto anche la fotografia di sfondo di una pagina acquisita! Poiché Readiris supporta la scansione di immagini in scala di grigi e a colori, è possibile acquisire qualsiasi elemento grafico, sia esso un'immagine creata manualmente, una foto in bianco e nero o un'illustrazione a colori. Se un documento contiene delle tabelle, Readiris le riorganizza in celle e ricrea i bordi delle celle delle tabelle originali.

In altre parole, Readiris consente di archiviare una copia dei documenti sotto forma di file di testo compatto editabile e non più di un'immagine. Sono disponibili vari livelli di formattazione, la scelta è lasciata all'utente.

Vengono riconosciuti anche i codici a barre presenti in una pagina acquisita, così come il testo scritto a mano, sempre che le lettere maiuscole ben distaccate.

Readiris permette anche di riconoscere biglietti da visita: il programma acquisisce biglietti da visita, li riconosce e li converte trasferendoli in un database. Si pensi all'ultima fiera di settore a cui si è partecipato. Si torna sempre a casa con una pila di biglietti da visita, che richiedono alla segretaria due giorni di lavoro per essere ridigitati.

I dati del biglietto vengono estratti automaticamente dall'immagine e quelli riconosciuti vengono assegnati a voci specifiche del database. Readiris utilizza un database di confronto, quindi è in grado di distinguere il nome dal cognome, una città della nazione, un numero di telefono da un numero di fax, ecc. Il risultato delle informazioni può essere inviato direttamente al programma di gestione contatti, come ad esempio Microsoft Outlook (Express) o qualsiasi applicazione compatibile vCard.

Readiris supporta un'ampia gamma di scanner molto diffusi sul mercato - scanner da tavolo, scanner con alimentazione a foglio e periferiche multifunzione ("MFP" - "multifunctional peripheral"). In alternativa, è possibile utilizzare fotocamere digitali. Inoltre, Readiris supporta lo standard di scansione Twain e alcune piattaforme di scansione. La funzione scansione ad intervalli consente di acquisire documenti costituiti da più pagine in modo efficiente anche se il vostro



scanner non è dotato di un alimentatore automatico di fogli. (Readiris Corporate supporta gli scanner ad alta velocità ed esegue l'OCR in blocco su grandi raccolte di immagini: è possibile utilizzare pagine vuote per segmentare i blocchi acquisiti in documenti separati, mentre la lettura automatica dei codici a barre garantisce la corretta indicizzazione dei documenti "scomposti".)

SOMMARIO

Risparmiare tempo senza ridigitare il testo	III
Sommario	VI
Riconoscimenti e copyright	VII

Capitolo 1: Installazione

Requisiti di sistema	1-1
Installazione del software Readiris	1-1
Disinstallazione del software Readiris	1-3
Programma di disinstallazione di Readiris	1-3
Disinstallazione guidata di Windows	1-4
Installazione delle opzioni software	1-5
File installati	1-8
File "Leggimi" e documentazione	1-8
Modulo di scrittura manuale	1-8
Driver dello scanner	1-8
Registrarsi per ottenere vantaggi	1-9
Assistenza sul prodotto	1-11
Contattare I.R.I.S.	1-12

Capitolo 2: Informazioni guidate

Avvio del software	2-1
Primo avvio	2-2
Visualizzazione dell'interfaccia Readiris	2-3
Una prima esercitazione	2-5
Ingrandimento delle immagini	2-10
Primo, scomposizione di un'immagine acquisita	2-13
Primo e mezzo, ordinamento delle finestre	2-16
Secondo, organizzazione manuale in finestre dell'immagine acquisita	2-19

Terzo, salvataggio di modelli di organizzazione in finestre	2-23
Con Readiris, viaggi in tutto il mondo!	2-25
Lessici d'utente per ampliare la linguistica	2-30
Se necessario, Readiris cambia la lingua	2-32
Lettura di documenti con lingue miste	2-34
Definire le caratteristiche del documento	2-36
Readiris diventa semprepiù intelligente	2-37
Apprendi	2-40
Non apprendi	2-40
Elimina	2-41
Annulla	2-41
Termina	2-42
Interrompi	2-42
Il ruolo dei dizionari dei font	2-42
Invio del risultato all'applicazione	2-45
Salvataggio dei risultati in un file di testo	2-49
Creazione di un documento esportabile	2-53
... O leggerli	2-60
Riconoscimento di pagine multiple	2-63
Modifica di documenti a più pagine	2-73
Creazione di un nuovo documento	2-76
Riconoscimento di zone di testo	2-77
Organizzazione del documento di testo	2-79
Configurazione dello scanner	2-80
Come colorare le scansioni	2-82
Apparecchi diversi, diverse risoluzioni	2-85
Salvataggio delle impostazioni predefinite	2-89
Salvataggio di impostazioni specifiche	2-90
Scansione di documenti	2-91
Regolazione dell'immagine acquisita	2-95
La procedura guidata, un aiuto infallibile!	2-100
Readiris ricrea il layout del documento	2-101
Colonne, non cornice	2-106
Formattazione di testo, parte 2	2-109
Esportazione di testo ripetuta	2-110
Salvataggio separato degli elementi grafici	2-111
Salvataggio di sfondi colorati	2-114



Inserimento completo di elementi grafici	2-117
Lettura di fax e riconoscimento differito	2-119
Riconoscimento di blocchi di documenti	2-120
Creazione di una cartella controllata	2-122
Organizzazione in blocchi	2-124
Riconoscimento di tabelle	2-128
Riconoscimento della scrittura manuale	2-134
Lettura delle barre e degli spazi	2-138
Riconoscimento dei biglietti da visita	2-139
Scansione dei biglietti da visita	2-140
Attivare la modalità lettura dei biglietti	2-145
Riconoscimento dei biglietti da visita	2-149
Uso della guida in linea	2-150

RICONOSCIMENTI E COPYRIGHT

Il software Readiris è stato progettato e realizzato da I.R.I.S. Tecnologia OCR, ICR, lettura dei codici a barre, BCR, Connectionist, AutoFormat e linguistica di I.R.I.S. I.R.I.S. è titolare del copyright del software Readiris, della tecnologia OCR, della tecnologia ICR, della tecnologia di lettura dei codici a barre, della tecnologia BCR, della tecnologia linguistica, del sistema di guida in linea e del presente manuale.

AutoFormat, Connectionist, l'IBCR-II, il logo I.R.I.S., Readiris e tecnologia linguistica I.R.I.S. sono marchi di I.R.I.S.

XML parser sviluppato da Apache. Questo prodotto comprende software sviluppato da Apache Software Foundation (www.apache.org).

Acrobat e Reader sono marchi registrati di Adobe. Excel, Windows e Word sono marchi registrati di Microsoft. Intel è un marchio registrato di Intel.

Capitolo 1

INSTALLAZIONE

In questo capitolo vengono illustrati i requisiti di sistema e l'installazione del software Readiris.

REQUISITI DI SISTEMA

Di seguito vengono indicati i requisiti di sistema necessari per utilizzare Readiris:

- PC Intel 486 o compatibile. Si consiglia un PC Pentium.
- 64 MB di RAM. Si consigliano 128 MB di RAM per l'elaborazione di immagini in scala di grigi e a colori.
- 120 MB di spazio libero su disco. Sono sufficienti 105 MB di spazio su disco se le immagini d'esempio vengono lasciate sul CD-ROM.
- sistema operativo Windows XP, Windows ME, Windows 2000, Windows 98 o Windows NT 4.0.

Readiris Corporate richiede un monitor con risoluzione 1.024 x 768.

È possibile che alcuni **piloti di scanner** non siano compatibili con la ultime versioni di Windows. A questo scopo, leggere la documentazione fornita in dotazione con lo scanner.

INSTALLAZIONE DEL SOFTWARE READIRIS

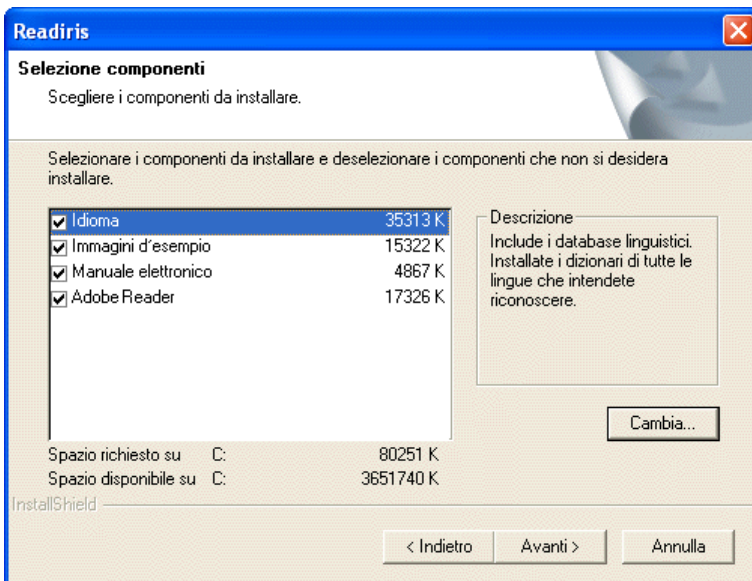
Il software Readiris viene fornito esclusivamente su **CD-ROM autoinstallante**. Per installarlo è sufficiente inserire il CD-ROM nell'apposita unità e attendere l'avvio del programma di installazione. Seguire le istruzioni visualizzate.



Nel caso in cui, dopo aver inserito il CD-ROM nell'apposita unità, l'installazione non si verifica, avviare il programma di installazione utilizzando il file MENU.EXE.

Gli utenti di Windows XP, Windows 2000 e Windows NT devono accertarsi di possedere i **diritti di accesso** appropriati contattando, se necessario, l'amministratore del sistema.

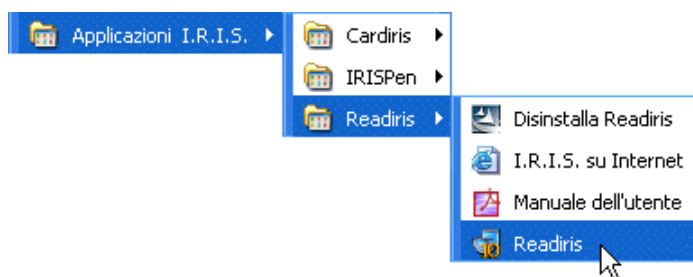
Per l'installazione sono possibili diverse opzioni. Assicurarsi di installare i **database linguistici** di tutte le lingue che si desidera riconoscere. Per impostazione predefinita sono installati tutti i dizionari. Si consiglia di installare le **immagini d'esempio** utilizzate nelle esercitazioni di questo manuale.



Allo stesso modo, se necessario, installare il software Adobe Reader per accedere alla documentazione software. Il **manuale elettronico** viene copiato

sul disco rigido per impostazione predefinita. È possibile anche consultarlo direttamente dal CD-ROM.

Il sottomenu "Applicazioni I.R.I.S. - Readiris" del menu "Programmi" viene creato automaticamente dal programma di installazione.



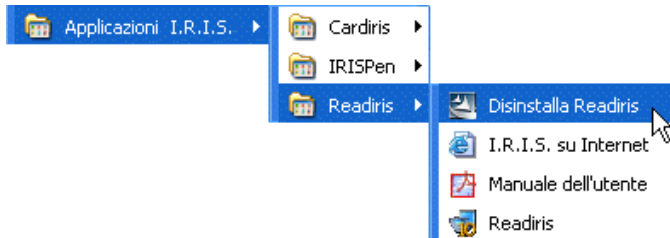
In questo modo è possibile accedere al **collegamento** di Readiris dal **desktop** e lanciarlo direttamente.

DISINSTALLAZIONE DEL SOFTWARE READIRIS

Sono disponibili due modi per disinstallare Readiris: utilizzare il programma di disinstallazione di Readiris e la disinstallazione guidata di Windows. Si consiglia di *non* disinstallare Readiris o i relativi moduli software cancellando manualmente i file del programma.

Programma di disinstallazione di Readiris

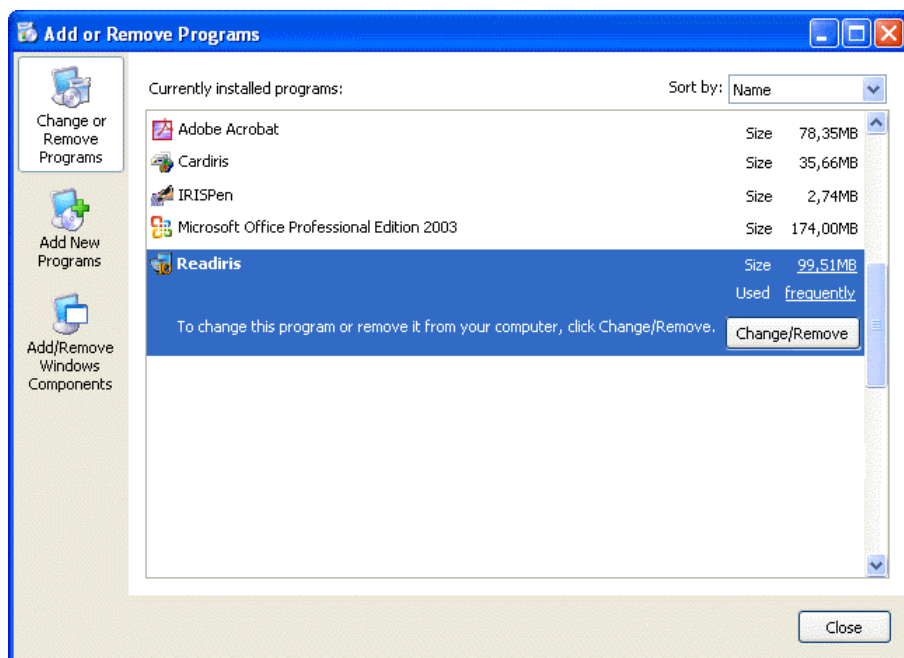
Per avviare il programma di disinstallazione di Readiris, selezionare "Disinstalla Readiris" nel sottomenu "Applicazioni I.R.I.S. - Readiris" e seguire le istruzioni visualizzate.



Disinstallazione guidata di Windows

Per utilizzare la procedura guidata di Windows, svolgere le seguenti operazioni.

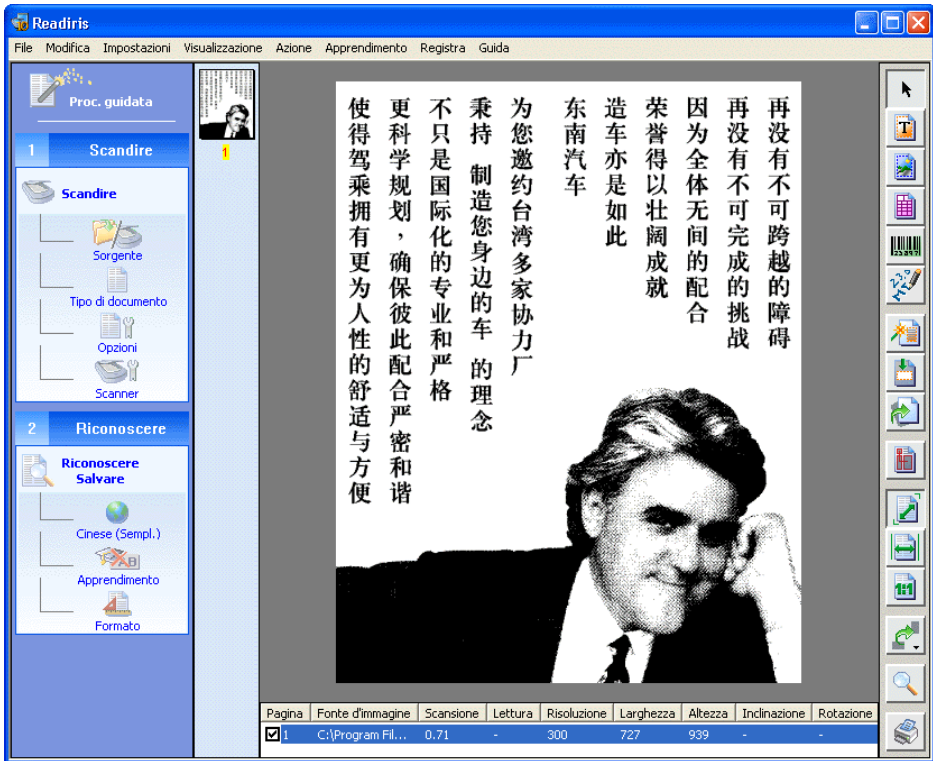
- Fare clic su "Impostazioni" nel menu di avvio di Windows e passare al "Pannello di Controllo".
- Fare doppio clic sull'icona "Installazione Applicazioni".



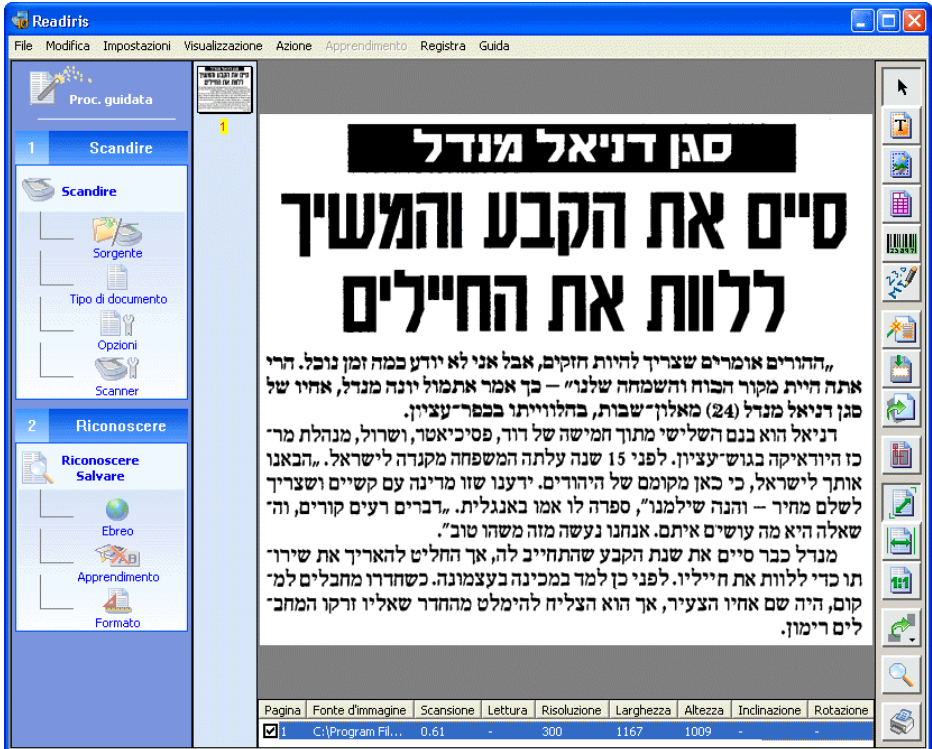
- Seguire le istruzioni visualizzate per rimuovere il software Readiris.

INSTALLAZIONE DELLE OPZIONI SOFTWARE

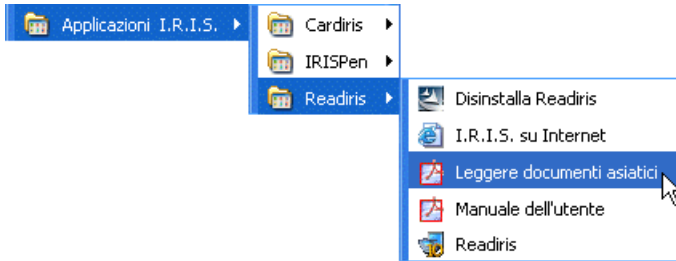
Sono disponibili due opzioni software per il software Readiris: il “Modulo OCR asiatico” e il “Modulo OCR ebraico”. Il “Modulo OCR asiatico” consente di leggere il giapponese, il cinese tradizionale, il cinese semplificato e il coreano.



Il “Modulo OCR ebraico” naturalmente consente di riconoscere documenti in ebraico.



Queste opzioni software sono disponibili sul CD-ROM autoinstallante. Installando il "Modulo OCR asiatico" sarà disponibile una documentazione specifica su come riconoscere i documenti asiatici.



FILE INSTALLATI

Il programma di installazione crea una cartella in cui vengono collocati i file di Readiris. Non tentare mai di disinstallare Readiris o alcuni dei relativi moduli cancellando manualmente i file del programma: utilizzare invece il programma di disinstallazione di Readiris o la disinstallazione guidata di Windows. Vedere i capitoli precedenti.

File “Leggimi” e documentazione

README.HTM File “Leggimi” (in formato HTML)

MANUAL.PDF Manuale dell’utente (in formato Adobe Acrobat)

Modulo di scrittura manuale

TEMPLATE.PDF Modulo vuoto di scrittura manuale per la ristampa

TEMPLATE.DOC Modulo vuoto di scrittura manuale per la modifica

Driver dello scanner

Rivolgersi al produttore o al rivenditore dello scanner se i problemi con i piloti dello scanner persistono. La maggior parte dei produttori consente di scaricare le versioni aggiornate dei piloti dai propri siti web.

REGISTRARSI PER OTTENERE VANTAGGI

Non dimenticare di registrare la propria licenza Readiris. Questo consentirà a I.R.I.S. di inviare aggiornamenti sugli sviluppi del prodotto e sui prodotti I.R.I.S. correlati. Solo gli utenti registrati usufruiscono dei vantaggi della registrazione, incluse l'**assistenza gratuita sul prodotto** e le **offerte speciali**.

È possibile effettuare la registrazione in diversi modi: inviando la scheda di registrazione o trasmettendola via fax, contattando I.R.I.S. durante l'orario di lavoro e compilando il modulo di registrazione disponibile sul sito web I.R.I.S.



Guida Readiris

Hide Back Forward Home Print Options

Contents Index Search

- Benvenuti nella guida Readiris
- Presentazione dell'OCR
- Riconoscere i documenti
- Riconoscere la scrittura manuale
- Riconoscere i codici a barre
- Come...?
- Informazioni di riferimento
- Versioni e opzioni software
- Registrazione del prodotto
- Registrazione della licenza**
- Modulo di registrazione
- Ordinare per e-mail
- Supporto del prodotto
- I.R.I.S.

Registrazione della licenza Readiris

Perché effettuare la registrazione?

- La registrazione consente di tenersi aggiornati sulle **nuove versioni del prodotto** sui **prodotti I.R.I.S. correlati**.
- La registrazione conferisce il diritto di ricevere gratuitamente **assistenza tecnica per il prodotto e offerte speciali**.
- In base al pacchetto software acquistato, riceverete un codice o **chiave software** ("softkey"), necessaria per continuare ad utilizzare Readiris dopo il mese di prova.

Come...?

Posta

Spedire la **scheda di registrazione**.

Fax

In caso di perdita della scheda di registrazione o se la documentazione relativa al software era stata fornita solo in **forma elettronica**, stampare il [modulo di registrazione](#) Readiris e spedirlo o trasmetterlo via fax a

La **registrazione in linea** di Readiris disponibile nel menu "Registra" guida agevolmente l'utente durante la procedura di registrazione.



A seconda della versione acquistata, l'utente riceve la **chiave software** necessaria per continuare a utilizzare il software Readiris dopo un mese.

ASSISTENZA SUL PRODOTTO

Il comando "Supporto del Prodotto" del menu "Guida" di Readiris fornisce informazioni su come ricevere assistenza tecnica. Descrivere chiaramente il problema verificatosi e includere tutti i dati rilevanti relativi a Readiris, allo scanner e al sistema in uso.



Guida Readiris

Hide Back Forward Home Print Options

Contents Index Search

- Benvenuti nella guida Readiris
- Presentazione dell'OCR
- Riconoscere i documenti
- Riconoscere la scrittura manuale
- Riconoscere i codici a barre
- Come...?
- Informazioni di riferimento
- Versioni e opzioni software
- Registrazione del prodotto
- Supporto del prodotto
 - Ricevere l'assistenza al prodotto**
 - Contattare I.R.I.S.
 - Contattare I.R.I.S. tramite e-mail
 - I.R.I.S.

Come ricevere assistenza per il prodotto

A tutti i **clienti registrati** viene offerta **assistenza tecnica gratuita**. (La [registrazione](#) dà diritto anche a ricevere offerte speciali.)

Europa

Numero verde: 32-10-45 13 64 (ore ufficio) (tutte le lingue principali)
Fax: 32-10-45 34 43

USA

Numero verde: 1-561-921-0847 / 800-447-4744 (ore ufficio)
Fax: 1-561-921-0854

WWW

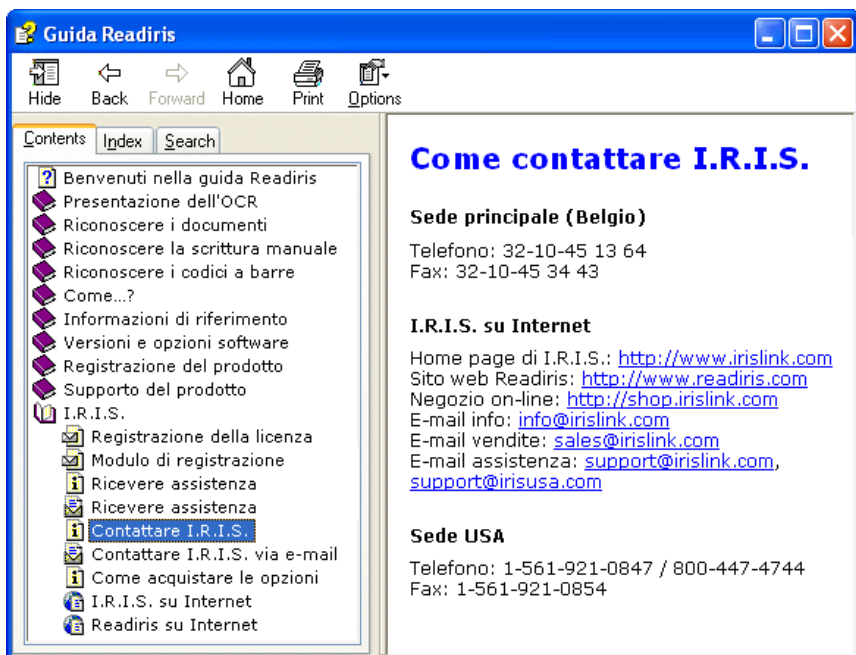
www.irislink.com/support.html (info risoluzione problemi)
Fare clic [qui](#) per accedere alle informazioni relative alla risoluzione dei problemi.

E-mail

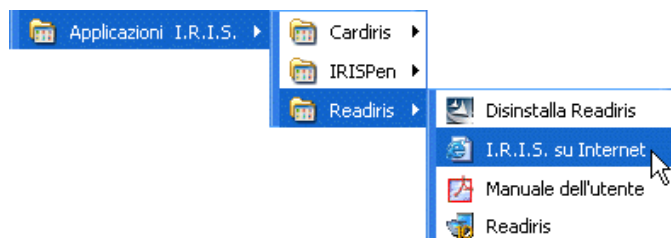
support@irislink.com, support@irisusa.com

CONTATTARE I.R.I.S.

È possibile contattare I.R.I.S. anche per ricevere maggiori informazioni su altre soluzioni della gamma di prodotti; il comando "Contatta I.R.I.S." del menu "Guida" di Readiris spiega come mettersi in contatto con I.R.I.S.



Un'icona di applicazione del sottomenu "Applicazioni I.R.I.S. - Readiris" nel menu "Programmi" porta direttamente alla **home page** I.R.I.S. In alternativa, è possibile accedervi dalla schermata di avvio di Readiris selezionando il comando "I.R.I.S. su Internet" nel menu "Guida" di Readiris.



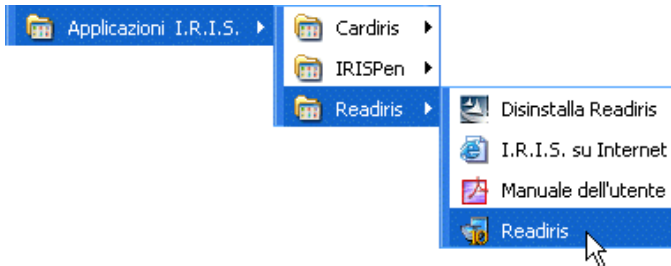
Capitolo 2

INFORMAZIONI GUIDATE

Readiris è un pacchetto OCR che rappresenta lo stato dell'arte ed è dotato di numerose funzioni avanzate. In questo capitolo vengono presentate le principali caratteristiche e le informazioni utili, nonché i consigli relativi all'uso di Readiris.

AVVIO DEL SOFTWARE

Fare clic sull'applicazione Readiris nel sottomenu "Applicazioni I.R.I.S. - Readiris" o sul collegamento dell'applicazione Readiris del desktop.



Vengono visualizzate la schermata di avvio e la finestra dell'applicazione Readiris. La schermata di avvio mostra la versione e il copyright del software Readiris. Consente inoltre di accedere direttamente alla **home page** I.R.I.S.; è sufficiente fare clic sull'URL per visitare il sito web I.R.I.S. Facendo clic in qualsiasi punto, la schermata scompare.

La finestra seguente riguarda la procedura guidata per l'installazione dell'OCR; fare clic su "Esci" per il momento.

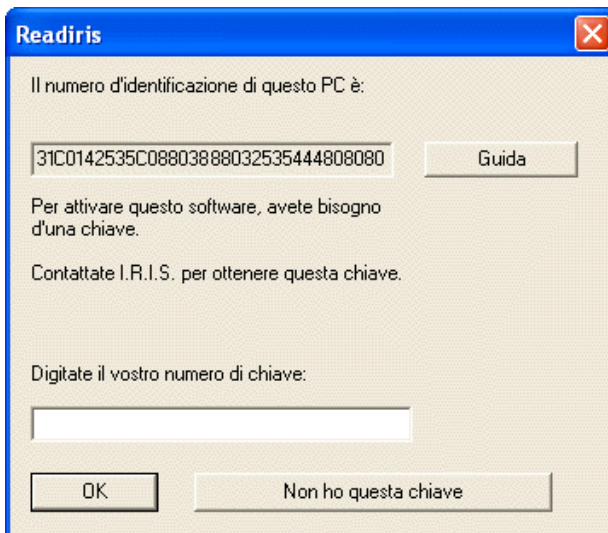


PRIMO AVVIO

A seconda del pacchetto software acquistato, per il primo avvio può essere necessaria una procedura particolare; è possibile che venga richiesta la registrazione della licenza.

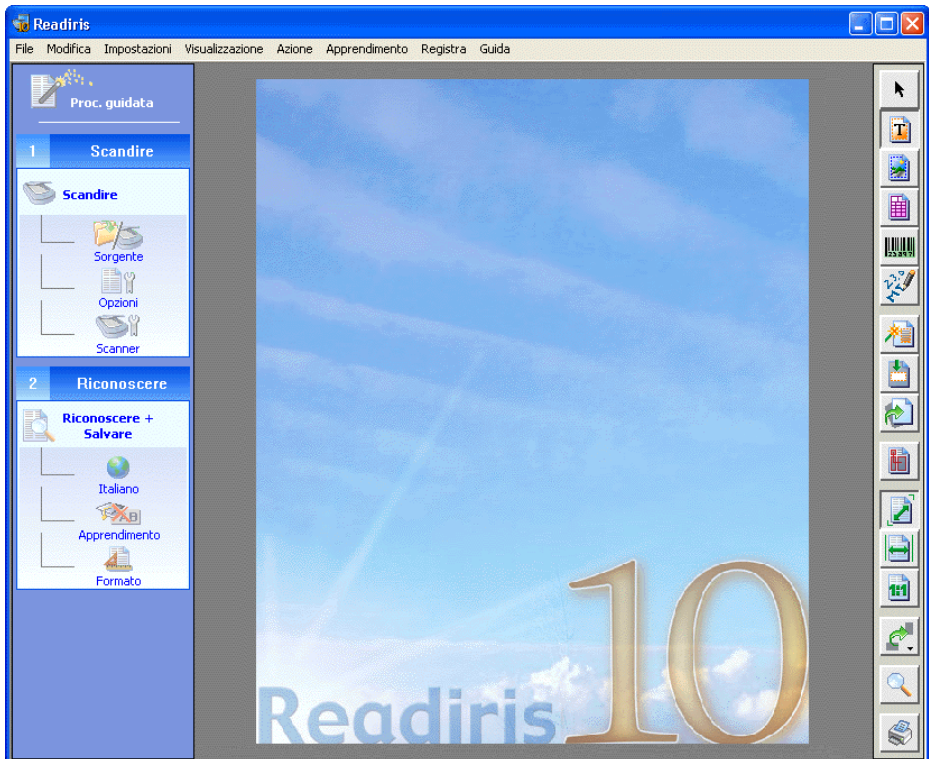
In questo caso, la possibilità di utilizzare Readiris è limitata a un mese. In seguito alla registrazione, I.R.I.S. invierà **una chiave software** che consente di continuare a utilizzare Readiris dopo il primo mese.

Per generare la chiave software è necessario il **numero di identificazione** personale; durante la registrazione della licenza assicurarsi che tale numero sia disponibile o indicato.



VISUALIZZAZIONE DELL'INTERFACCIA READIRIS

La finestra dell'applicazione Readiris non solo contiene i **menu dei comandi**, ma anche due barre di pulsanti che consentono un rapido accesso a tutti i comandi utilizzati di frequente. Inizialmente alcuni menu relativi all'anteprima non sono visibili e non risultano disponibili fino a che non viene aperta un'immagine.





Lo stesso vale per la **barra degli strumenti immagine** posta sul lato destro della finestra dell'applicazione: questa contiene tutti i comandi necessari durante l'anteprima dell'immagine. La **barra degli strumenti principale** posta sul lato sinistro consente un rapido accesso a tutti i comandi generali utilizzati di frequente.

Per scoprire quale comando corrisponde ad un determinato pulsante, tenere per qualche secondo il puntatore del mouse sul pulsante: verrà visualizzata una **descrizione** sulla funzione del comando.



Il riquadro della finestra o **zona immagine** è l'area in cui vengono visualizzate le immagini acquisite. Per riconoscerle è possibile portare i file immagine nella zona immagine e sull'icona Readiris.

Durante l'elaborazione pagine, viene aggiunta sul lato sinistro un'altra barra, la **barra strumenti della pagina**. Rappresenta le diverse pagine del documento e consente l'accesso ai comandi facendo clic con il pulsante destro del mouse (menu di scelta rapida).



In secondo luogo, il **pannello dei documenti** viene visualizzato sotto l'immagine acquisita. Mostra le **statistiche** e le informazioni riguardanti le pagine acquisite - origine e risoluzione dell'immagine, tempo necessario alla scansione e al riconoscimento ecc. Il pannello dei documenti include anche le descrizioni dei vari campi.

Pagina	Fonte d'immagine	Scansione	Lettura	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\multipage.tif	1.11	-	300	2000	2388	-	-
<input checked="" type="checkbox"/> 2	C:\Program Files\Readiris\multipage.tif	0.82	-	300	2000	1888	-	-
<input checked="" type="checkbox"/> 3	C:\Program Files\Readiris\multipage.tif	0.77	-	300	1912	2004	-	-

UNA PRIMA ESERCITAZIONE

Il modo migliore per acquisire dimestichezza con il funzionamento di Readiris è indubbiamente utilizzarlo. Unitamente al software viene fornita una serie di



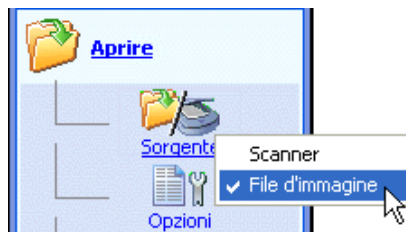
immagini preacquisite che consentono di iniziare anche se lo scanner non è collegato al computer. Segue la spiegazione.

Il tasto "Sorgente" della barra strumenti principale consente di indicare se si intende utilizzare un'immagine acquisita o preacquisita come origine delle immagini.

Immagini a colori, in scala di grigi e in bianco e nero sono supportate allo stesso modo. Readiris consente di aprire documenti PDF Adobe Acrobat, immagini fax DCX (una versione a più pagine del formato Paintbrush), immagini DjVu (*.djv, *.djvu), immagini JPEG, immagini JPEG 2000 (*.j2c, *.jp2), immagini PNG, immagini TIFF (non compresse, compresse di LZW, PackBits, Gruppo 3, Gruppo 4 e JPEG), immagini TIFF a più pagine, bitmap Windows (*.bmp) e immagini ZSoft Paintbrush (*.pcx).

Questa proprietà è particolarmente utile per convertire i **fax** in file di testo editabili.

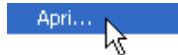
Se si desidera aprire un'immagine preacquisita, selezionare "File d'immagine" e non lo scanner come origine delle immagini, mediante il pulsante "Sorgente" della barra strumenti principale.



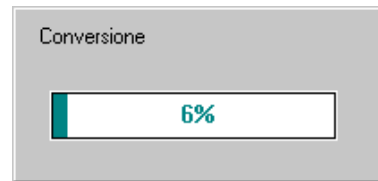
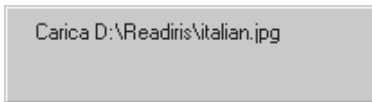
Cliccare quindi sul pulsante "Aprire". Selezionando il disco come origine dell'immagine, il tasto "Scandire" viene sostituito dal tasto "Aprire", mentre il corrispondente comando "Scansione" del menu "Azione" viene sostituito dal comando "Apri".



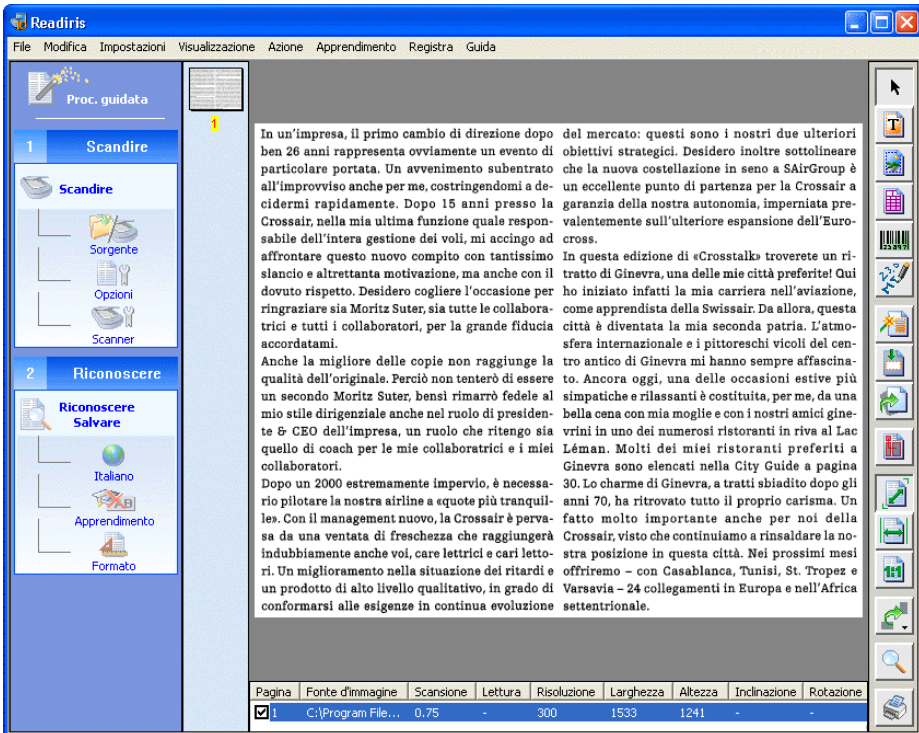
È possibile inoltre selezionare il comando "Apri" dal menu "File" e aprire direttamente un'immagine preacquisita; questa operazione è possibile anche quando l'origine delle immagini corrente è lo scanner.



Il sistema chiederà di selezionare un file immagine. Selezionare il file ITALIAN.JPG nella cartella Readiris. Poiché questo file campione è un'immagine a colori, non viene soltanto letta dal disco: viene infatti creata una versione in bianco e nero "binarizzata" per il processo OCR.



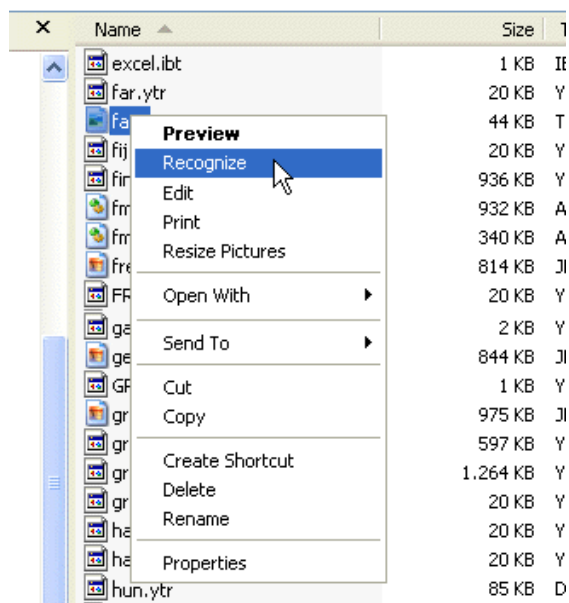
Infine, l'immagine viene visualizzata nella zona immagine. La barra strumenti della pagina e il pannello dei documenti indicano che è caricata una sola pagina in Readiris.



Un terzo modo per aprire immagini preacquisite è costituito dall'utilizzo del **“trascinamento della selezione”**: è possibile trascinare le immagini da Esplora Risorse di Windows alla zona immagine di Readiris o sull'icona Readiris, potendole quindi aprire immediatamente.



Infine è possibile aprire le immagini da Esplora Risorse di Windows. A tale scopo è sufficiente **fare clic con il pulsante destro** del mouse su un file immagine e selezionare il comando "Riconoscimento" dal menu di scelta rapida. (Il comando appare soltanto se il tipo di file è supportato.)



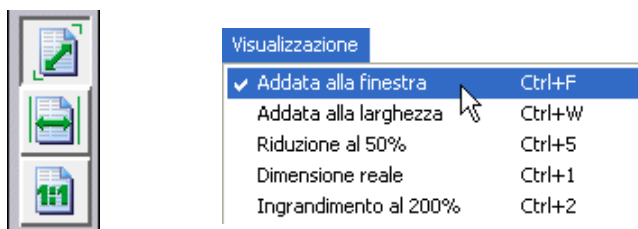
Ciò non significa che l'OCR è stato eseguito immediatamente: per consentire all'utente la piena flessibilità, viene lanciato Readiris e l'immagine viene aperta.

La barra degli strumenti immagine posta a destra della finestra dell'applicazione Readiris contiene tutti i comandi utili per l'anteprima dell'immagine: gli strumenti relativi all'indicazione delle zone di interesse, alla rotazione dell'immagine, all'ingrandimento o riduzione ecc.

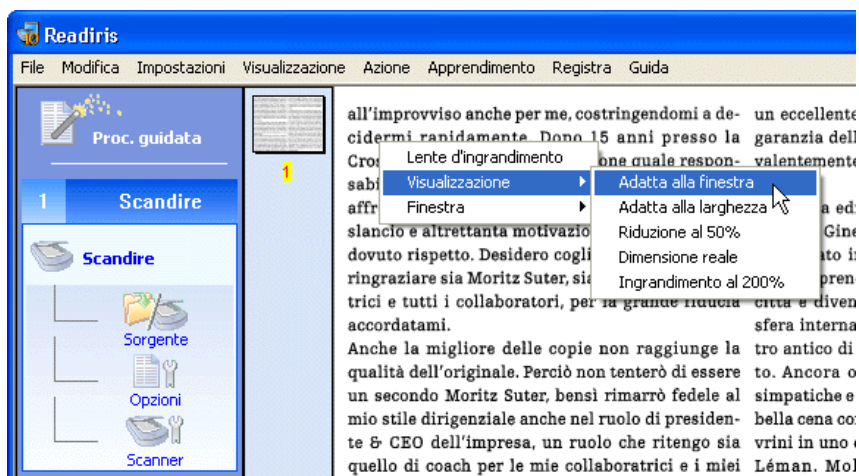
INGRANDIMENTO DELLE IMMAGINI

Readiris è dotato di numerosi comandi che consentono di **ingrandire** l'immagine acquisita, ad esempio per verificare la qualità della scansione.

La barra degli strumenti immagine contiene pulsanti che consentono di ingrandire fino alle dimensioni reali, di adattare l'immagine all'ampiezza della pagina e di inserire l'intera immagine nella finestra di anteprima. Il menu "Visualizzazione" contiene gli stessi comandi ed è dotato di due ulteriori livelli di zoom: è possibile visualizzare l'immagine al 50% e al 200% delle dimensioni reali. In dimensioni reali, un pixel dello schermo corrisponde ad un pixel dell'immagine. (Per tutti i livelli di zoom sono disponibili tasti di scelta rapida.)



I vari livelli di ingrandimento sono disponibili anche nel menu di scelta rapida, che viene visualizzato facendo clic con il pulsante destro del mouse. Da questo menu, selezionare il livello di zoom desiderato.

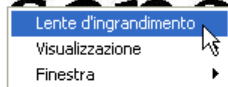


È, inoltre, possibile fare *due volte* clic con il pulsante destro del mouse su una zona immagine acquisita per ingrandire immediatamente alle dimensioni reali. Ripetere l'operazione per ridurre.

Infine, è possibile utilizzare la **lente di ingrandimento** per ingrandire i dettagli di un documento. La lente di ingrandimento è anche disponibile dal menu di scelta rapida facendo clic con il pulsante destro del mouse sull'immagine.



Per saperne

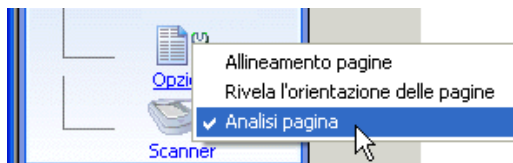




PRIMO, SCOMPOSIZIONE DI UN'IMMAGINE ACQUISITA

Dopo aver acquisito l'immagine, è necessario indicare quali parti si desiderano convertire in testo editabile, tracciando cornici definite "finestre" intorno alle zone di interesse.

In realtà, abilitando l'opzione "Analisi Pagina" disponibile dal pulsante "Opzioni" della barra strumenti principale (o nel menu "Impostazioni"), l'operazione verrà eseguita automaticamente.





La scomposizione automatica della pagina risulta particolarmente utile quando vengono riconosciuti **testi divisi in colonne** e documenti che presentano un complesso layout di pagina, che può comprendere elementi grafici e tabelle.

The screenshot shows the Readiris software interface. The main window displays a scanned document with the title "Give the Brain a Break" and a subtitle "2 device to avoid being sucked under by the information maelstrom". The document text is recognized and displayed in a structured format, with some text highlighted in yellow. The interface includes a menu bar (File, Modifica, Impostazioni, Visualizzazione, Azione, Apprendimento, Registra, Guida) and a toolbar on the right. The left sidebar contains two main sections: "1 Scandire" (Scan) and "2 Riconoscere" (Recognize), each with sub-options like Sorgente, Opzioni, Scanner, Riconoscere Salvare, Inglese, Apprendimento, and Formato. At the bottom, a table shows the recognition results for the scanned page.

Pagina	Fonte d'immagine	Scansione	Lettura	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 1	C:\Program Files\...	7,27	-	300	2074	2602	-	-

La scomposizione della pagina utilizza tre **tipi di finestra**: finestra di testo, finestra grafica e finestra con tabella. Readiris riconosce i blocchi di testo, le

tabelle e le zone grafiche contenenti fotografie, illustrazioni ecc. (Il salvataggio dei grafici e il riconoscimento delle tabelle sarà illustrato più avanti.)

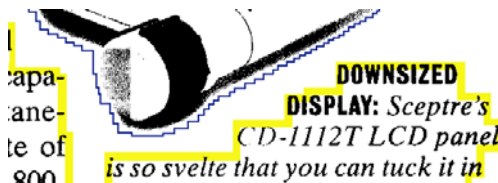
Due tipi di zona devono essere sempre tracciati manualmente: le zone codici a barre e le zone di scrittura in stampatello. (Per ulteriori informazioni sulla lettura dei codici a barre e sul riconoscimento di testo scritto in stampatello, fare riferimento ai rispettivi paragrafi del presente manuale.)

Un **codice colore** indica il tipo di finestra: le zone di testo hanno un bordo arancione, le finestre grafiche hanno un bordo porporino e le tabelle un bordo rosa. Le zone con codici a barre sono verdi e le zone a stampatello blu.

Il numero di finestre è sempre indicato nelle descrizioni degli strumenti.



L'analisi della pagina è veloce, estremamente accurata ed è in grado di riconoscere righe inclinate: traccia forme complesse, "irregolari".



L'analisi della pagina individua zone in cui compare **testo bianco su sfondo nero**. Il riconoscimento di tali inserimenti non costituisce un problema: mentre l'anteprima visualizza il documento correttamente acquisito, quando si presenta la necessità di riconoscere tali blocchi di testo, Readiris "inverte" l'immagine. (Inoltre è possibile creare un'immagine *completamente* capovolta dallo scanner per elaborare pagine con testo bianco su sfondo nero. Vedere di seguito.)



PRIMO E MEZZO, ORDINAMENTO DELLE FINESTRE

Readiris non solo individua i vari blocchi ma li *ordina*. Per impostazione predefinita le zone vengono selezionate capovolte, da sinistra a destra, per adattarsi ai documenti divisi in colonne.

È possibile modificare l'**ordine di selezione**. Fare clic sul tasto "Ordina" della barra strumenti immagine. Il cursore del mouse assume la forma di una mano appena si abilita la modalità di ordinamento.



Fare clic sulle finestre che si desidera includere. Le finestre *non* selezionate vengono semplicemente ignorate ed escluse dal riconoscimento. È facile comprendere quali sono le zone selezionate e quali meno: le finestre selezionate presentano i colori originali, mentre le finestre non selezionate hanno colori più tenui e non sono numerate.

The screenshot shows the Readiris software interface. The main window displays a scanned document page with the following content:

Give the Brain a Break

2 device to avoid being sucked under by the information maelstrom

3 THERE'S ONE PART OF THIS MAGAZINE I TRY HARD NOT to read it's the section we call *Numbers*, near the front. My colleague who prepares it (I dare to write this because she's on vacation) has the knack of collating statistics which some mental noise makes you suck in the proportions of Dutch couples who live in what used to be called sin (32%), the amount paid at auction for an Eric Clapton guitar (\$497,500), the estimated number of condos used in the world annually (9 billion), or how high the pile would be if Bill Gates stacked his wealth in \$1 bills (25,000 km).

The trouble with this fun is that it leaves stal clots in your brain. If you can't do some plumbing to restore thinking's normal flow, they accumulate like cholesterol up there in the gray matter's byways, causing serious mind-blocks. After three glasses of red you declare over dinner that Bill Gates has used nine billion condoms, Eric Clapton had a 25,000-km-long guitar, and 32% of Dutch couples earn \$497,500 a year. The only cure for this condition of factoid overload is a sort of neurotic plague which is said to be the main cause of mental illness in the 21st century—is to learn to sift. Survivors will be those with the capacity to distinguish Internet surfing from info-surflet.

The first step to becoming a sworn suffer is easier than giving up smoking. It only requires a little virtual brain surgery to find a warm spot in one of your lobes and implant there an imaginary device smaller than the delete key on a computer. Its chemical acronym might be DNK, which is not a kinky variant of DNA, but stands for Don't Need to Know.

After undergoing a voluntary DNK implant, I now shiver with pleasure on opening my morning newspapers as factoid after factoid that I don't need on my poor old soft disk is burned from memory's gates. An example: scientists in Texas say each year the Earth and the Moon move 3.82cm further apart, meaning that since Neil Armstrong was taking giant steps there 30 years ago, we are 114cm more distant. Standing at the back door looking up at the Big Cheese, does my soul benefit from this knowledge? No. Slam DNK it!

There are forests of facts which can not be so easily shed, as with some of the far-from-trivial statistics that jolt in our Numbers spot. A recent example is that 29 million people in our world live in some form of slavery. That needs to be stored and pondered. As does, say, a recent finding by the London School of Economics that four million British children live in poverty, relative though that is. This knowledge can't be tossed if we are to remain half human, but precisely because there is such a welter of factoids the importance of big numbers like these dilutes into a milky way of mental numbness.

It's not just numbers, it's also words. Don't-Need-to-Know beginners can, for example, zap any report to do with rats. Not the sewer sort but the laboratory kind, the ones which almost daily lead to variants of, "Tweaking gene 856G and injecting ribonucleodioninase greatly reduces hair loss in male rats over 50. Scientists at Helioblab caution that many years of research will be required before..."

Sport offers superb DNK material. You can whitout all that coaches say before a match and most of what they say after it, along with what the players mutter about why they lost won. It's the event, stupid! If you've watched Lance Armstrong through each stage of the Tour de France, do you really need to learn it wasn't easy on his legs?

Politics is the promised land of DNK. Meaning you blank all promises and pre-election blarney, all state-of-the-nation monologues. One needs to take in only how the votes fall, all details (on the where-there's-smoke basis), and party funding figures. Where politics bleeds into economics, there is also room to turn a deaf ear. DNK all interest rate forecasts, stockpiling and merger talk because it's about as accurate as tarot, and anyway you never hear it in time to benefit from it.

Criticism can be a joy, but keep the box button ready for those experts who can't resist giving the whole plot away, explaining what the painting really means, or who majored in semiotics. It's easy to build such personalized DNK lists. My own long one also includes all advertisements (especially those offering "personalized" things), and anything to do with any royal marriage, anywhere. Also flamed out are all numbers about flying saucers, mobile phone sales, desirable cholesterol levels, winners of the U.S. Masters, and guesstimates of how much Bill Gates has in the bank, or how it might stack up.

The DNK club is free and there are no membership numbers. Our cathartic motto might sound familiar—if you haven't yet DNKed out all slogans - It's No Brain, no Brain.

Pagina	Fonte d'immagine	Scansione	Lettura	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 1	C:\Program Files\...	7.27	-	300	2074	2602	-	-

L'analisi della pagina è abilitata per impostazione predefinita. Se è stata involontariamente disabilitata l'analisi della pagina o sono state cancellate alcune finestre ma si desidera rieseguire un'analisi della pagina, per fare in modo che Readiris scomponga la pagina corrente è sufficiente fare clic sul pulsante "Analizza Pagina" della barra degli strumenti immagine.



Selezionare la lingua del documento *prima* di eseguire l'analisi della pagina quando si elaborano documenti asiatici ed ebraici. Per queste lingue vengono utilizzate routine specifiche: lo spazio tra le righe è normalmente più ampio nei documenti asiatici che nei documenti occidentali, il testo è composto da piccoli elementi grafici ("ideogrammi") che potrebbero facilmente essere interpretati come zone grafiche nei documenti occidentali e il testo si legge dall'alto in basso, da destra a sinistra. Nei documenti ebraici il testo procede da destra a sinistra. Se si è dimenticato di selezionare la lingua desiderata, sarà possibile farlo in un secondo tempo. Readiris riesegue automaticamente l'analisi della pagina!

Alcuni documenti presentano puntini "sparsi" sulla pagina, che possono determinare un bordo nero intorno all'immagine effettiva. Per cancellare tutte le piccole finestre, presumendo che non contengano testo, e riordinare le zone restanti, è possibile fare clic sul comando "Sopprimi le Piccole Finestre" disponibile nel menu "Modifica".

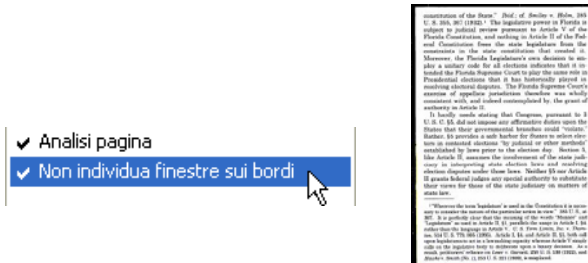
Sopprimi le piccole finestre ... Ctrl+M



December 12, 1999
Dear Friends,
Twenty-five years ago this weekend, thousands of people showed up to the polls in my hometown. It was a special election, called by the local Board of Education.
Imagine a whole town coming together to decide your fate like that. Of course, those weren't the exact words they were. "Should Michael Moore be removed from the Board of Education?" but that was pretty much how it felt. As I sat there all day in the sole polling station in school gymnasium, watching the people stream in to cast their ballot, I wondered why had given up going away to college, let me I could try and try to change some things in the school system. I sat there in those benches overlooking the voting booths and figured I was just nuts.
Two weeks prior to this "judgment day," I had become the youngest elected official in the country - and the first 18 year old ever elected to public office. A constitutional amendment had just passed, lowering the voting age from 21 to 18. Well, I was 18 years old, and I guess the adults figured if we could go get killed (and kill others) for no damn reason, then they figured they should let us vote and drink. Some trade off.
So, as I mentioned to you last year in one of these letters, I was a slightly kicked off teenager (this was before the era of school shootings and Riots). I decided the best way to fight the high school principal was to become his boss - what I was told a student. So, I ran for the school board on the strength of being, um, here today I won. Nine months later, the principal was gone.

È inoltre disponibile una routine automatica simile: il rilevamento delle zone ai bordi delle pagine. Se si disattiva questa routine, l'analisi della pagina ignora tutte le zone che toccano i bordi della pagina. Se lo scanner genera dei bordi neri attorno all'immagine, l'analisi della pagina individua delle zone in cui si trovano solo elementi di disturbo. Le zone grafiche poste ai bordi della pagina non vengono

modificate: spesso le fotografie toccano i bordi pagina e, normalmente, le immagini di sfondo coprono l'intera pagina ecc.



SECONDO, ORGANIZZAZIONE MANUALE IN FINESTRE DELL'IMMAGINE ACQUISITA

L'analisi della pagina consente di organizzare automaticamente in finestre una pagina acquisita. In alternativa, è possibile dividere manualmente in zone un'immagine mediante gli **strumenti di organizzazione in finestre** di Readiris.

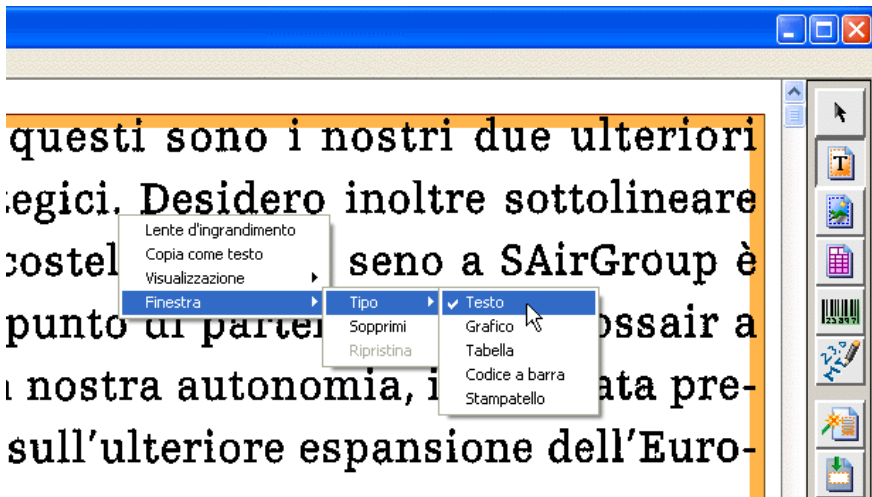




(Come precedentemente detto, le finestre dei codici a barre e della scrittura in stampatello devono essere sempre tracciate manualmente, poiché l'analisi della pagina non è in grado di rilevarle!)

Per **tracciare** un rettangolo intorno a una zona d'interesse, selezionare il relativo strumento nella barra degli strumenti immagine e trascinare il cursore dall'angolo superiore sinistro all'angolo inferiore destro della finestra. (Non sono consentiti lati inferiori a 1 mm. perché non potrebbero contenere neppure un singolo carattere.)

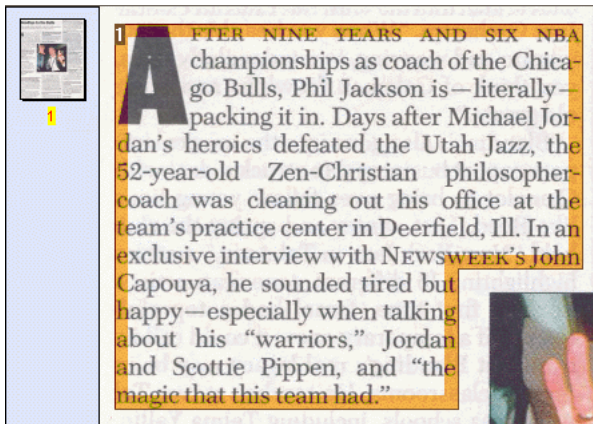
Non importa se è stato selezionato un tipo di zona errato: è possibile modificare rapidamente il tipo facendo clic con il pulsante destro del mouse su una finestra e selezionando il comando "Finestra - Tipo" dal menu di scelta rapida.



Le finestre vengono ordinate automaticamente sulla base dell'ordine di creazione: i numeri indicano la sequenza di ordinamento.

È possibile anche incorniciare i blocchi di testo "irregolari" tracciando intorno ad essi **finestre poligonali**. Unendo zone rettangolari si creano finestre non

rettangolari: quando due rettangoli dello stesso tipo si intersecano, diventano automaticamente un'unica finestra. È come costruire una casa aggiungendo una stanza dopo l'altra... (Non ha senso creare finestre poligonali di tabelle o codici a barre.)



Inoltre, l'organizzazione in finestre manuale può essere associata alla selezione delle finestre: è possibile disegnare nuove finestre anche quando è attivata la modalità di ordinamento. Utilizzare la funzione di ordinamento per includere un numero di finestre rilevate e creare manualmente altre finestre se l'analisi della pagina non ha dato i risultati desiderati. Non appena inizia l'operazione di creazione della finestra nella modalità di ordinamento tutte le finestre non selezionate vengono immediatamente cancellate.

Per modificare, spostare ed eliminare finestre, è necessario innanzitutto **selezionarle**. A questo scopo, scegliere lo strumento "Selezione Finestra" (lo strumento "freccia") nella barra strumenti immagine e fare clic all'interno di una finestra. Compaiono così marcatori rettangolari in ciascun angolo e al centro dei lati della finestra.



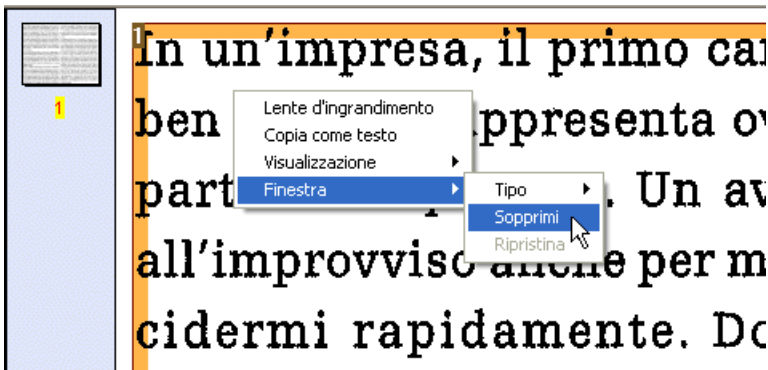
Will he come back? Should he?

Per **delesezionare** finestre, fare clic su qualche altro punto. Per selezionare **ulteriori finestre**, tenere premuto il tasto Maiusc facendo clic sulle finestre desiderate. Per selezionare una finestra e le **finestre incluse** di un tipo diverso, tenere premuto il tasto Ctrl facendo clic sulla finestra principale.

Queste sono le operazioni relative alla selezione delle finestre. Per **modificare** una finestra di testo, selezionarla, posizionare il cursore su un marcatore e trascinare il lato per modificare le dimensioni della finestra.

Per **spostare** una finestra, è sufficiente selezionarla e trascinarla in un'altra posizione.

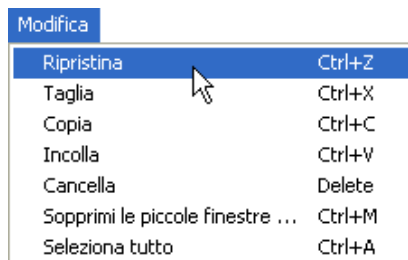
Per **eliminare** finestre, selezionarle, fare clic con il pulsante destro del mouse e scegliere il comando "Finestra - Cancella" dal menu di scelta rapida. In questo modo si cancelleranno tutte le finestre selezionate nonché la finestra posta sotto il cursore.



In alternativa, è possibile selezionare zone e scegliere il comando "Taglia" o "Cancella" dal menu "Modifica". Il comando "Taglia" taglia le finestre trasferendole in una memoria temporanea interna, mentre "Cancella" elimina definitivamente le

finestre. Quando si incollano finestre, queste vengono inserite nella loro posizione originale ed è necessario trascinarle nella nuova posizione.

Infatti, alle finestre si applicano *tutti* i comandi noti del menu "Modifica", che consentono di eliminare, tagliare, copiare e incollare finestre. È possibile utilizzare anche il comando "Ripristina": nel caso in cui alcune finestre siano state involontariamente eliminate, spostate, ridimensionate ecc., mediante il comando "Ripristina" sarà possibile annullare l'ultima operazione.



Inoltre per tutti i comandi sono disponibili tasti di scelta rapida! Segue un esempio: per eliminare tutte le finestre esistenti, è possibile scegliere il comando "Seleziona Tutto" del menu "Modifica" o il relativo tasto di scelta rapida Ctrl+A e fare clic sul comando "Cancella" del menu "Modifica" o sul relativo tasto di scelta rapida Elimina. A questo punto è possibile ricreare il layout necessario. Per ripristinare il layout precedente, è possibile scegliere "Ripristina" o il relativo tasto di scelta rapida Ctrl+Z.

TERZO, SALVATAGGIO DI MODELLI DI ORGANIZZAZIONE IN FINESTRE

Per impieghi successivi, i layout di organizzazione in finestre ottenuti possono essere salvati come **modelli di divisione in zone** mediante il comando "Salva Finestre" del menu "File" e caricati in memoria mediante il comando "Carica Finestre".

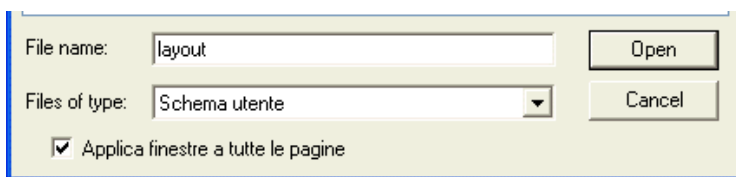


Carica finestre...

Salva finestre...

Qualora si renda necessario riconoscere documenti con un layout simile, ad esempio una relazione di 50 pagine in cui, per ovvie ragioni, sarebbe necessario escludere l'intestazione e il piè di pagina, è possibile utilizzare un unico modello per dividere in zone tutte le 50 pagine.

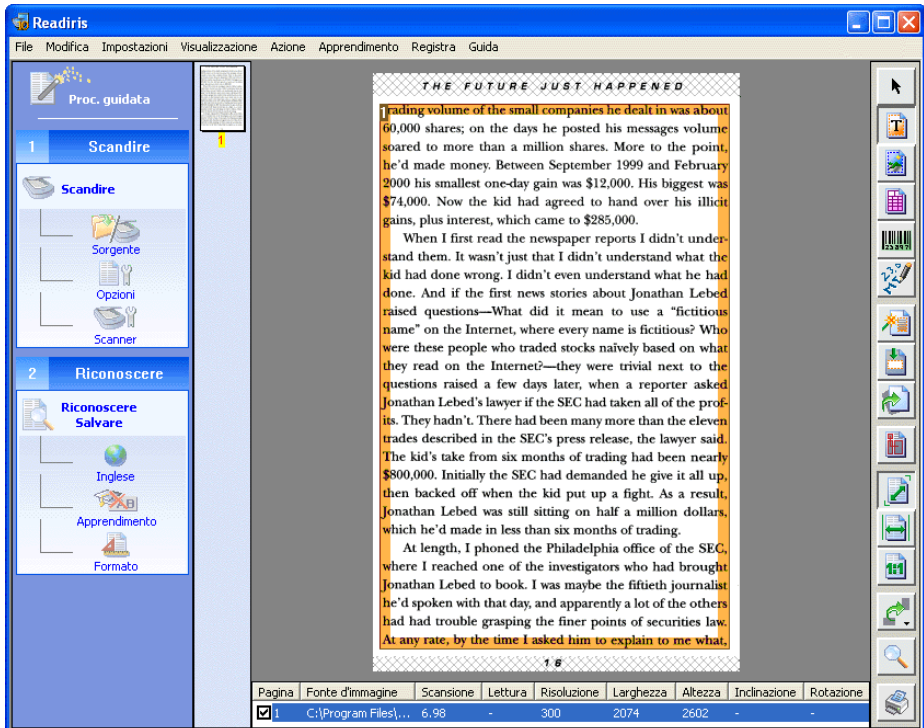
Quando si carica un modello in memoria, l'analisi della pagina viene disabilitata automaticamente e il modello di divisione in zone resta attivo fino a quando dalla barra degli strumenti principale non si riabilita l'analisi della pagina. Durante la memorizzazione di un layout, è possibile attivare l'opzione "Applica Finestre a Tutte le Pagine" per applicare il layout a tutte le pagine del documento corrente.



In realtà, esiste un'alternativa alle configurazioni delle finestre: lo strumento di anteprima "Ignorare Zona Esterna" limita la scomposizione della pagina alla porzione di immagine "ritagliata".



Selezionare questo strumento e la parte di immagine da elaborare. Nel caso di documenti a più pagine, è possibile escludere la stessa zona esterna dall'analisi pagina per ogni pagina. (Rieseguire l'analisi pagina per cancellare il "ritaglio" dell'immagine o modificare manualmente le zone.)



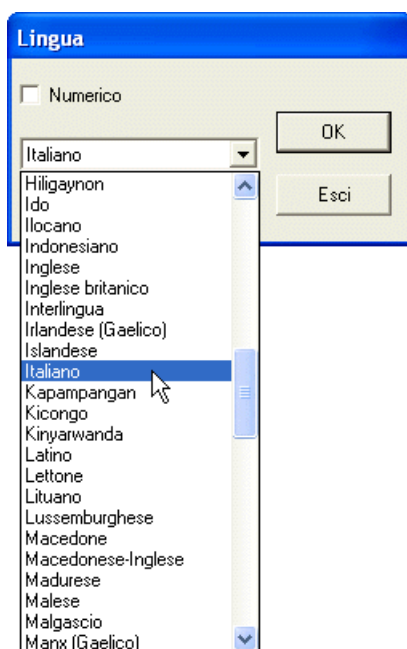
CON READIRIS, VIAGGI IN TUTTO IL MONDO!

Supponendo che le finestre siano state definite correttamente, il testo è quasi pronto per il riconoscimento dei caratteri. “Quasi” perché la lingua e le impostazioni del documento non sono ancora state verificate.

L'impostazione della lingua viene eseguita dalla barra degli strumenti principale.



Per modificare la lingua del documento, fare clic sul pulsante "Lingua".



Per accedere direttamente, premere il tasto di una lettera: se la lingua selezionata è l'italiano e si desidera selezionare, ad esempio, il danese, è possibile premere il "D" della tastiera per passare direttamente alla lingua danese. Nel caso in cui numerose lingue abbiano la stessa iniziale, per accedere alle opzioni premere più volte la lettera. Ad esempio: Readiris legge l'occitano e l'olandese. Premendo "O" una volta viene selezionato l'occitano, premendo "O" una seconda volta viene

selezionato l'olandese e premendo "O" una terza volta, si torna all'occitano. (Per passare a un'altra lettera, ad esempio T, premere il tasto BackSpace prima di digitare il carattere "T".)

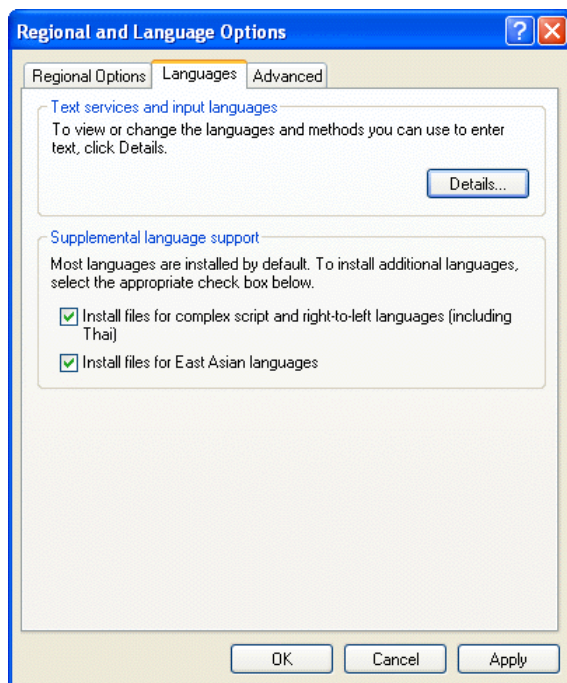
Readiris non si limita all'italiano ma è in grado di gestire fino a 117 lingue. Supporta infatti tutte le lingue americane ed europee, incluse le lingue dell'Europa centrale, le lingue cirilliche ("russo"), le lingue baltiche, il greco e il turco.

Esiste la possibilità di leggere **documenti ebraici** e **asiatici**: il modulo aggiuntivo "Modulo OCR ebraico" consente il riconoscimento di documenti ebraici, mentre l'opzione software "Modulo OCR asiatico" consente la lettura delle lingue giapponese, cinese semplificato, cinese tradizionale e coreano. (Il cinese semplificato è usato nella Cina continentale e a Singapore, mentre il cinese tradizionale è usato a Hong Kong, Taiwan, Macao e dalle comunità cinesi d'oltremare.)

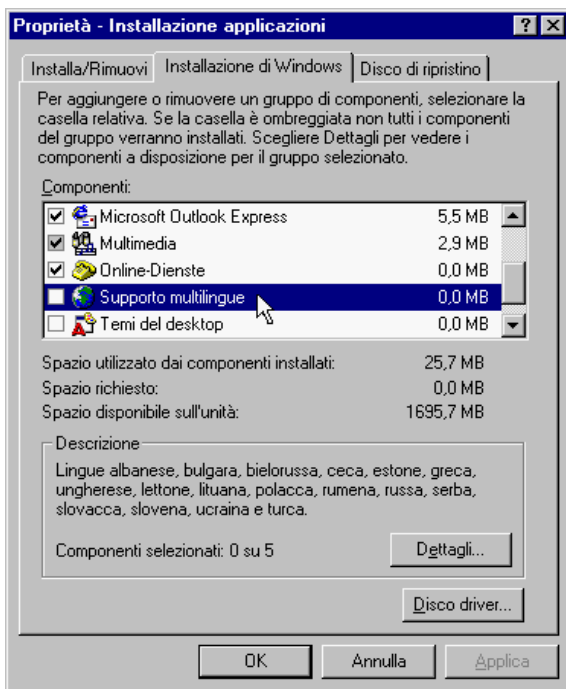
Inoltre l'inglese britannico viene distinto dall'inglese americano - o si dovrebbe dire "internazionale"? Lo stesso accade per lo spagnolo e il messicano.

Per visualizzare i caratteri delle lingue dell'Europa centrale, del greco, del turco, delle lingue cirilliche e baltiche, è necessaria la configurazione Windows appropriata. Per consentire al sistema di Windows di gestire queste lingue, può essere necessario installare il modulo di **Windows "Supporto Multilingue"**.

Su sistemi operativi Windows XP, 2000 e Windows NT 4.0, selezionare l'icona "Impostazioni Internazionali" del "Pannello di Controllo".



Su sistemi operativi Windows ME e 98, per verificare che il modulo di Windows "Supporto Multilingue" sia installato, selezionare l'icona "Installazione Applicazioni" nel "Pannello di Controllo".



Per visualizzare e modificare documenti asiatici ed ebraici, è possibile installare le versioni asiatiche ed ebraiche del sistema operativo Windows o utilizzare Word 2003, Word 2002 o Word 2000 per visualizzare e modificare i seguenti documenti: Office 2003 System, Office XP ed Office 2000 sono stati appositamente studiati per gestire questi documenti in molte lingue diverse. Per ulteriori informazioni sull'argomento, consultare il file **"Leggimi"** di Readiris.

È indispensabile selezionare la lingua del documento appropriata. Basandosi sulla lingua selezionata, il software sa quali **simboli** riconoscere. Il supporto



multilingue assicura il riconoscimento corretto di caratteri “esotici”, quali ad esempio ç, ß, ñ, γ e ø.

Inoltre, per convalidare i risultati ottenuti, il software fa largo uso di **database linguistici**. Si supponga sia necessario leggere la parola "presidente", dove una macchia d’inchiostro rende la "r" simile alla "f". Consultando il vocabolario di italiano, Readiris scoprirà autonomamente di dover leggere la parola "presidente" e che non ha alcun senso riconoscere il simbolo "f". È chiaro che questa **tecnica di “autoapprendimento”** è strettamente legata al contesto linguistico.

I criteri di linguistica offrono un utile aiuto per risolvere **casi ambigui**, come nel caso della lettera "O" che potrebbe essere confusa con il numero "0". Un altro esempio tipico è costituito dalla lettera "l" e dal numero "1" aventi forma identica in molti font; si pensi ai testi prodotti dalle vecchie macchine da scrivere. Il contesto linguistico aiuta a determinare se si tratta di una "l" o di un "1".

La figura sotto riportata mostra varie forme di "l" e "1". Le forme poste sulla prima riga risultano non ambigue, ma le forme poste sulla seconda sono ambigue, tuttavia i criteri di linguistica consentono di distinguerle. Qualora il contesto non sia sufficiente, interviene l’utente.

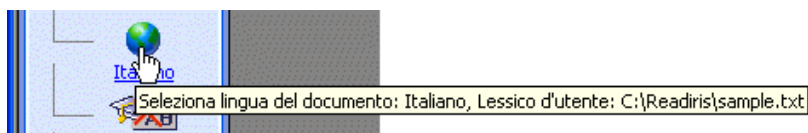
193 1950s. 1hr
Well, Rossellini

LESSICI D’UTENTE PER AMPLIARE LA LINGUISTICA

È possibile migliorare la qualità della linguistica personalizzandola ulteriormente: per quanto potenti possano essere i lessici standard, gli utenti di Readiris Corporate possono “potenziare” la **precisione OCR** caricando i lessici dell’utente con il comando "Lessico d’Utente" nel menu "Impostazioni".

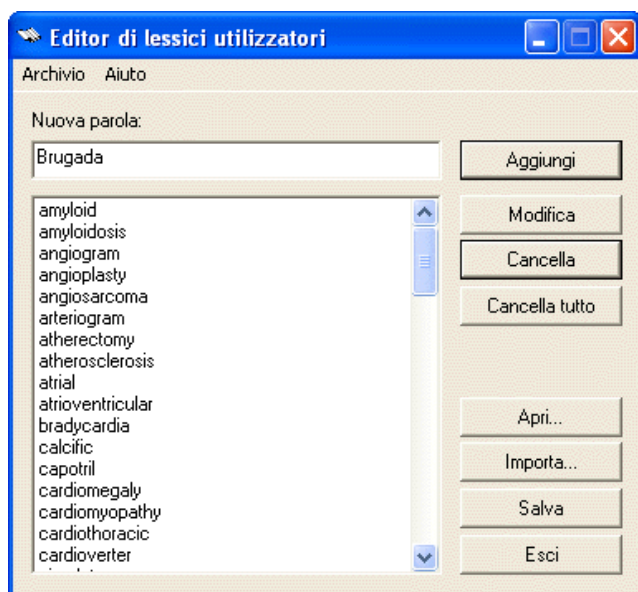


La descrizione del tasto "Lingua" indica il lessico dell'utente corrente.



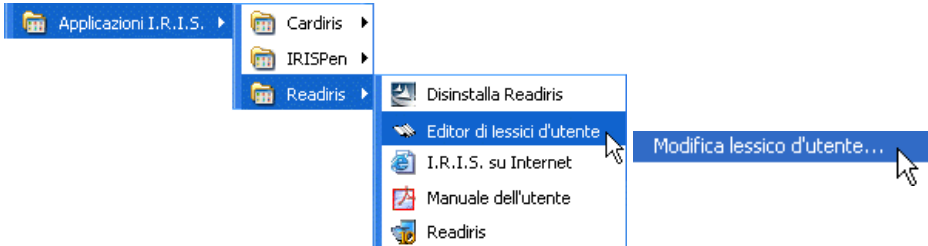
I lessici dell'utente sono elenchi di parole contenenti termini non ricorrenti nei lessici generali, "di base". Basta pensare ai **termini** tecnici, scientifici, legali o alla **terminologia aziendale**.

Readiris è provvisto di elaboratore dei lessici d'utente, un programma di utilità studiato per creare e memorizzare i **lessici dell'utente**. Questo strumento è di facile utilizzo; consultate la guida in linea in caso di dubbi sul suo funzionamento.





È possibile trovare l'elaboratore nel sottomenu "Applicazione I.R.I.S. - Readiris" e sotto il menu "Impostazioni" di Readiris.



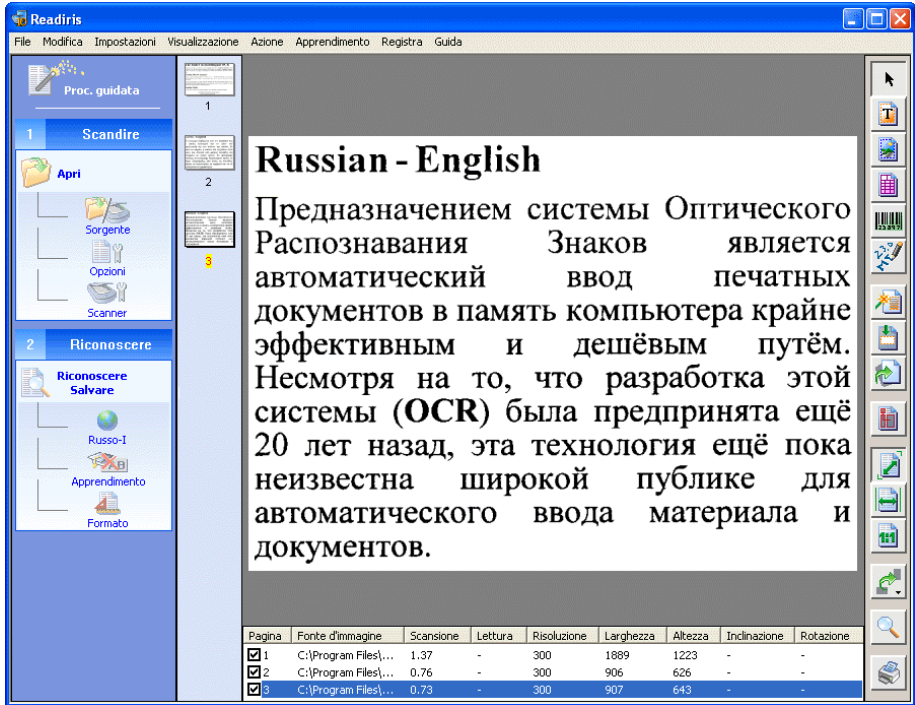
SE NECESSARIO, READIRIS CAMBIA LA LINGUA

Ma Readiris fa molto di più: Readiris può passare da una lingua all'altra all'interno di una frase senza richiedere l'intervento da parte dell'utente. Quando vi sono parole occidentali in documenti in lingua greca, cirillica, ebraica o asiatica - nomi propri non traslitterabili, marchi ecc., essi vengono scritti utilizzando i simboli occidentali conosciuti - infatti Readiris passa automaticamente all'alfabeto corretto. In altre parole, è possibile attivare un **alfabeto misto** di caratteri greci, cirillici, ebraici o asiatici con caratteri occidentali.

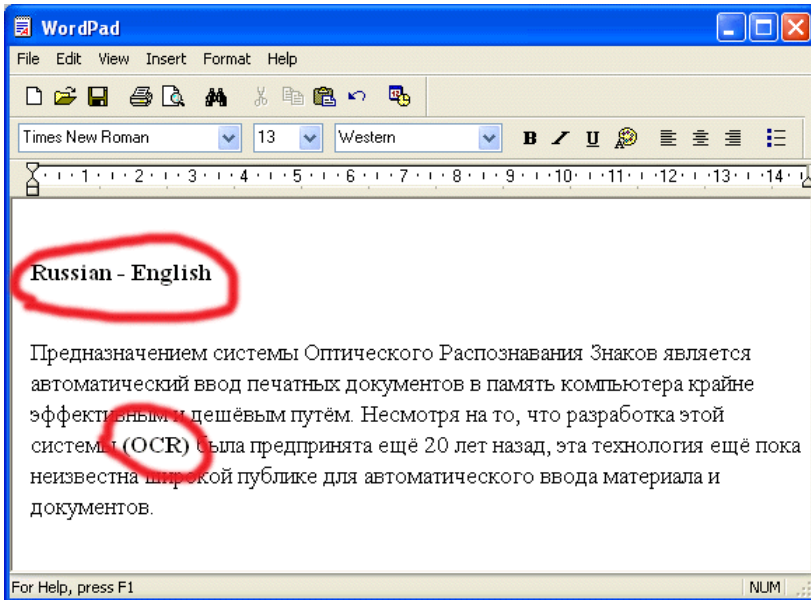
Assicurarsi di selezionare "Greco-Inglese" o la lingua cirillica appropriata, ad esempio "Bielorusso-Inglese". In altre parole, non tentare di selezionare solo "Greco" o "Bielorusso" come lingua del documento sperando che i simboli occidentali risultino chiari!



Di seguito viene fornito l'esempio di un testo russo contenente alcune parole inglesi. Per fare una prova, aprire il file immagine ALPHABETS.TIF ed effettuare il riconoscimento della pagina corrispondente!

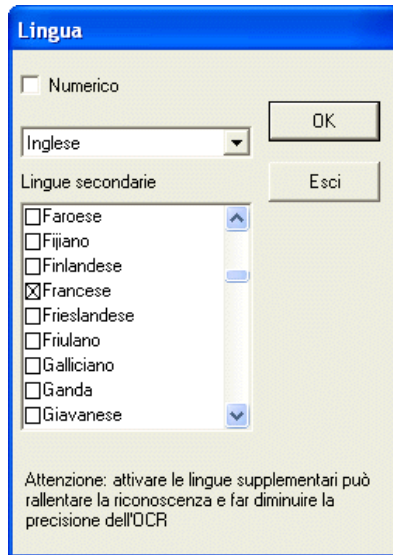


Quando viene aperto in un elaboratore di testi, il risultato finale è come quello presentato. Per visualizzare correttamente il testo russo, può essere necessario selezionare un **font** cirillico.



LETTURA DI DOCUMENTI CON LINGUE MISTE

Readiris Corporate è più potente quando si tratta della lettura di documenti in diverse lingue: con questa versione è possibile selezionare una lingua principale e (fino a) 4 lingue secondarie (dello stesso gruppo linguistico).

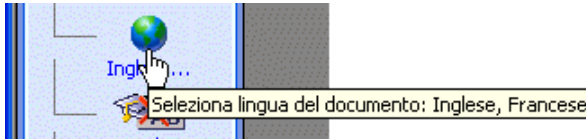


Per leggere un documento nel quale la traduzione in francese si affianca al testo inglese, selezionare l'inglese come lingua principale e il francese come lingua secondaria. Questo non consente solo di estendere il set di caratteri per includere caratteri accentati come ç, è e ù. Entrambi i database linguistici forniscono un “feedback” al processo OCR e questo consente a Readiris di individuare automaticamente la lingua utilizzata in un dato punto del documento.

È possibile selezionare contemporaneamente fino a 5 lingue. Solo le lingue appartenenti allo stesso gruppo linguistico sono attivabili: le lingue incluse in un'altra tabella codici di Windows rispetto alla lingua “principale” *non* possono essere attivate simultaneamente. Innanzitutto, selezionare solo le lingue pertinenti: quanto più ampio è il set di caratteri, tanto più lento è il riconoscimento e quindi maggiore il rischio di errore!



Per conoscere rapidamente le lingue selezionate, utilizzare il tasto "Lingua" per visualizzarne l'elenco...



DEFINIRE LE CARATTERISTICHE DEL DOCUMENTO

Una volta impostata la lingua, si passerà a specificare le altre caratteristiche del documento, ad esempio il tipo di font e la spaziatura dei caratteri, impostando con precisione il riconoscimento del documento stesso. (Questi comandi non si applicano ai documenti asiatici.)

Iniziare con il comando "Tipo di Font" del menu "Impostazioni". Le modalità dei differenziano i documenti "normali" dai documenti stampati **a matrice di punti**. I simboli a matrice di punti "bozza" o "9 aghi" sono composti da punti isolati, separati; per riconoscerli sono necessarie procedure di riconoscimento estremamente specializzate.

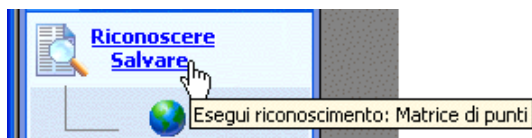
ape--descended life

La stampa a matrice di punti in "modalità lettera", chiamata anche matrice di punti a "25 aghi", richiede l'impostazione normale, proprio come per la **qualità di stampa** di materiale tipografico, dattiloscritto, stampato con stampanti laser e a getto d'inchiostro.

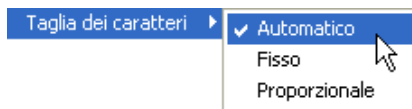
L'impostazione "Automatico" specifica che Readiris individua la modalità del font automaticamente. Lasciare che Readiris individui automaticamente la modalità del font in tutte le situazioni, a meno che non si è certi che soltanto i documenti a matrice di punti siano letti. (Naturalmente, l'opzione "Automatico" è l'impostazione predefinita.)



Il tipo di font è indicato nella descrizione del tasto "Riconoscere-Salvare". Se non si aggiunge alcun messaggio alla descrizione, il rilevamento automatico della qualità di stampa risulta valido, quando il messaggio "Matrice di punti" compare nella descrizione, la lettura per matrici di punti è attivata.



È possibile impostare la **spaziatura dei caratteri** mediante il comando "Taglia dei Caratteri" del menu "Impostazioni".



Con font *fissi* o “a larghezza fissa”, tutti i simboli del font sono della stessa larghezza. In questa frase ad esempio, una "i" e una "w" occupano in senso orizzontale lo stesso spazio sulla riga. Si pensi ai documenti prodotti con una macchina da scrivere, in cui il carrello lascia lo stesso spazio per ogni simbolo dattiloscritto.

Per spaziatura *proporzionale* si intende che la larghezza di un carattere dipende dalla sua forma. Simboli come "m" e "w" sono più larghi, occupano più spazio in senso orizzontale su una riga rispetto a caratteri “stretti” come "l" o "j". In genere, tutti i libri, riviste e giornali sono stampati con spaziatura proporzionale.

La soluzione più semplice consiste nel lasciare ogni volta questa opzione impostata sul valore predefinito "Automatico", per consentire a Readiris di individuare automaticamente la spaziatura dei caratteri.

READIRIS DIVENTA SEMPREPIÙ INTELLIGENTE

Una volta selezionata la lingua del documento e impostate le caratteristiche dello stesso, attivare l'apprendimento interattivo e fare clic sul pulsante "Riconoscere-Salvare".



L'avanzamento del riconoscimento viene indicato sullo schermo. Per interrompere il riconoscimento del testo, fare clic sul pulsante "Stop".



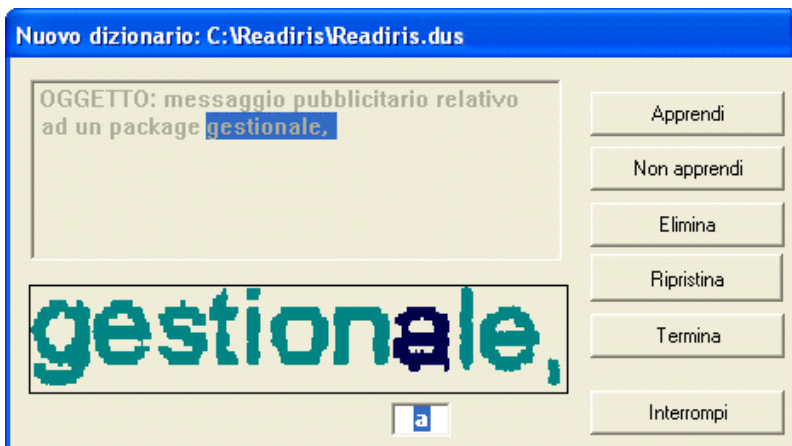
Al termine del riconoscimento, Readiris passa alla fase di apprendimento interattivo in cui l'apprendimento è abilitato dal pulsante "Apprendimento" della barra strumenti principale.

(L'apprendimento interattivo non si applica in documenti asiatici: l'apprendimento non ha senso per queste lingue nelle quali si utilizzano migliaia di simboli differenti ed è necessario immettere "ideogrammi", cosa non facile con le tastiere occidentali!)

L'**apprendimento dei font** può migliorare notevolmente la precisione del sistema di riconoscimento. Quando l'utente prova a leggere caratteri distorti e illeggibili, quali si trovano in documenti reali o forme di font stilizzate che Readiris non riconosce in modo ottimale, l'apprendimento consente di superare questo temporaneo "ostacolo".

L'apprendimento guidato dall'utente viene anche utilizzato per addestrare il sistema a riconoscere **simboli speciali** che Readiris non è in grado di riconoscere, quali i simboli matematici e scientifici e i segni grafici decorativi ("dingbats"). Ad esempio, è possibile addestrare Readiris a riconoscere il simbolo " π " come "pi" o il segno grafico "☎" come "Tel". (Tuttavia, la lista dei simboli riconosciuti non può essere ampliata con i simboli " π " e "☎".)

Il testo riconosciuto viene visualizzato progressivamente e il sistema si arresta su caratteri dubbi o, se si tratta di caratteri in contatto ("legature"), su stringhe di caratteri dubbi. I caratteri dubbi vengono evidenziati e presentati sempre nel loro contesto. I caratteri che non sono stati riconosciuti vengono rappresentati con una tilde (il simbolo "~").



Verificare innanzitutto di aver attivato il dizionario dei font e la modalità di dizionario appropriati, che sono sempre indicati nel titolo della finestra di apprendimento. In caso contrario, fare clic sul pulsante "Interrompi"; viene rivisualizzata l'immagine del documento con la divisione in zone iniziale, abilitare il dizionario di font o la modalità di dizionario appropriati ed eseguire nuovamente l'OCR. (Seguirà una spiegazione sul funzionamento dei dizionari dei font.)



Se necessario, sostituire la forma errata o sconosciuta con un carattere o una stringa di caratteri e fare clic su uno dei seguenti pulsanti.

Apprendi

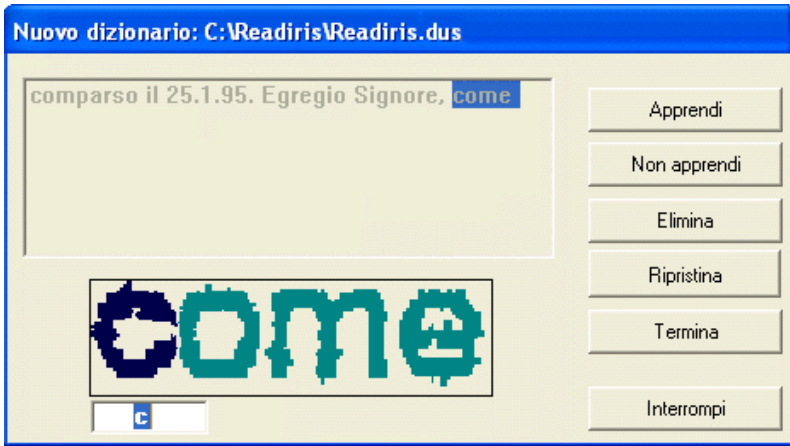
Consente di accettare la soluzione proposta o di correggerla. Il programma salva questo carattere dubbio nel dizionario dei font come “sicuro”, definitivo. Per un riconoscimento successivo non verrà più richiesto l’intervento dell’utente, in quanto la forma viene considerata definitivamente appresa.

Nell’esempio precedente, il sistema si ferma su un carattere pulito e facendo clic su "Apprendi" si accetta una forma che non può essere confusa con altri caratteri.

Non apprendi

Consente di accettare la soluzione proposta o correggerla. La differenza rispetto al pulsante "Apprendi" è data dal fatto che il simbolo appreso viene classificato nel dizionario come “incerto”. Per un riconoscimento successivo il sistema proporrà la soluzione “appresa”, ma richiederà un’ulteriore conferma.

Questo pulsante viene utilizzato per simboli che potrebbero essere confusi con altri: una "e" illeggibile che potrebbe essere scambiata per una "c", una "t" danneggiata che somiglia molto a una "r" ecc.



Nel caso della "c" sopra indicata, danneggiata a tal punto da somigliare al simbolo "e", l'utente dovrebbe fare clic sul pulsante "Non Apprendi", in modo tale da non confonderla con il simbolo "e".

Elimina

La forma visualizzata viene eliminata dall'output. Questo pulsante viene utilizzato per ignorare elementi di disturbo presenti nei documenti, come chiazze, macchie di caffè ecc., che potrebbero essere riconosciute come punti, virgole ecc., nonché per cancellare ogni altro simbolo non desiderato.

Annulla

Questo comando consente di tornare a correggere gli errori, annullando le ultime 32 decisioni.



Termina

Il processo di apprendimento viene interrotto, ma l'OCR prosegue in modalità automatica. Tutte le decisioni prese da questo momento in poi dal sistema vengono accettate senza la convalida dell'utente.

Fare clic su questo pulsante quando il riconoscimento è estremamente preciso e non richiede una revisione dettagliata.

Interrompi

Non confondere il pulsante "Termina" con il pulsante "Interrompi": con "Interrompi" non si ottiene alcun output e si ricomincia tutto daccapo, con "Termina" si crea un testo che però non viene controllato dettagliatamente!

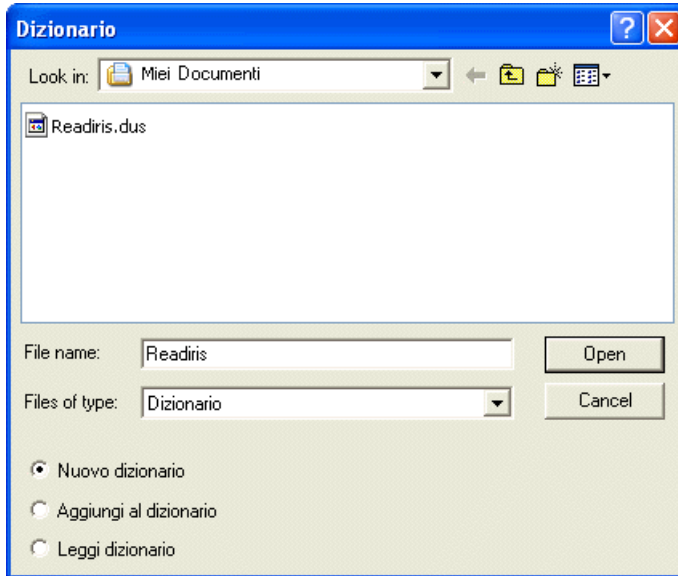
IL RUOLO DEI DIZIONARI DEI FONT

I risultati di ciascuna sessione di addestramento vengono conservati temporaneamente nella memoria del computer ma, per impieghi successivi, è possibile e necessario memorizzarli in file chiamati "dizionari dei font".

Non confondere i dizionari dei font con i lessici! I dizionari dei font contengono forme carattere apprese durante la fase OCR interattiva, i lessici sono database linguistici di ausilio nel riconoscimento.

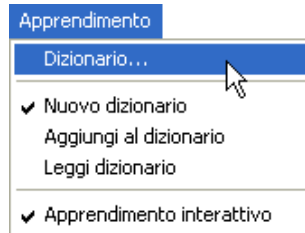
Quando si desidera riconoscere documenti simili, è necessario caricare in memoria questi dizionari dei font per utilizzare l'intelligenza supplementare che contengono. In questo modo Readiris tiene conto dell'intelligenza memorizzata in queste biblioteche dei font. Si potrebbe affermare che ad ogni impiego Readiris diventa più intelligente!

Come è possibile questo? Il funzionamento dei dizionari dei font è controllato dal menu "Apprendimento": selezionare un dizionario con il comando "Dizionario" e determinarne la modalità di funzionamento.



I **dizionari** dei font sono limitati a 500 forme, pertanto si consiglia di creare dizionari separati per applicazioni specifiche, ad esempio per tipo di documento. I dizionari hanno per impostazione predefinita l'estensione *.dus. Quando il dizionario è pieno, l'addestramento non ha più alcun effetto: i risultati dell'apprendimento non vengono più conservati in memoria o scritti in un dizionario.

È possibile impostare la modalità del dizionario all'interno del comando "Dizionario" o direttamente dal menu "Apprendimento". Sono disponibili tre modalità di dizionario: nuovo, aggiungi e leggi.



Selezionando "Nuovo Dizionario", i risultati del processo di apprendimento verranno salvati in un *nuovo* dizionario. Qualora si selezioni un dizionario esistente, i suoi contenuti verranno cancellati.

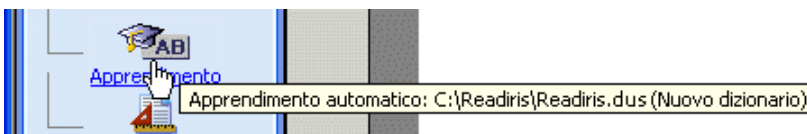
Con la modalità "aggiungi" i risultati del processo di apprendimento verranno salvati in un dizionario *esistente*: il riconoscimento si avvale di un'intelligenza supplementare già presente nel dizionario, a cui vengono aggiunte nuove forme di font. In altre parole, questa opzione consente di creare man mano un dizionario dei font.

(Quando si assegna un nome di file ad un nuovo dizionario e si attiva la modalità "aggiungi", viene creato un dizionario dei font vuoto che deve essere completato.)

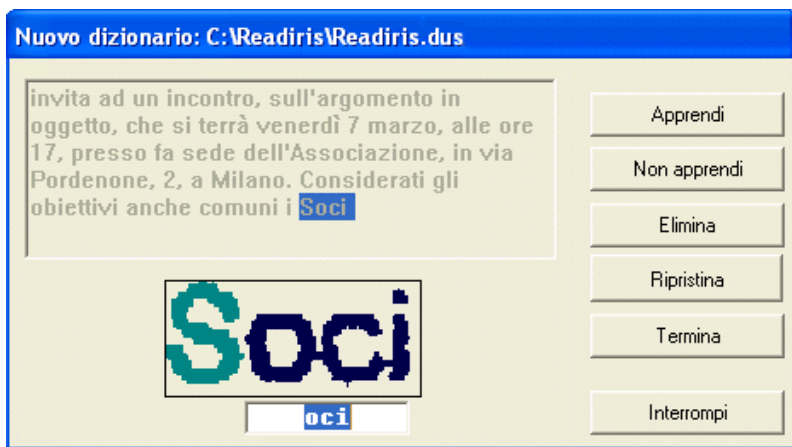
Con l'ultima opzione, "Leggi Dizionario", il dizionario funziona nella modalità di sola lettura: si utilizza il dizionario *senza* aggiungervi nuove forme di font.

Selezionare la modalità "nuovo" quando si esegue il riconoscimento di una sola pagina. Per riconoscere numerose pagine dello stesso tipo, come pagine con gli stessi font e la stessa qualità di stampa, selezionare la modalità "nuovo" per la prima pagina, la modalità "aggiungi" per alcune pagine successive e la modalità "leggi" per la parte restante del documento.

La descrizione del pulsante "Apprendimento" indica sempre quale dizionario dei font è attualmente attivo e in quale modalità.



Quando si passa all'apprendimento interattivo, nel titolo della finestra vengono indicati il dizionario e la relativa modalità di funzionamento; qualora essi siano errati, fare clic sul pulsante "Interrompi" e ricominciare.



INVIO DEL RISULTATO ALL'APPLICAZIONE

L'addestramento interattivo pone termine al riconoscimento dei caratteri. Poiché Microsoft Word agisce per impostazione predefinita come applicazione destinataria, l'elaboratore di testi viene avviato automaticamente al termine del riconoscimento, se necessario. Il testo riconosciuto viene inserito.

Durante la formattazione del documento riconosciuto, sullo schermo può comparire una barra di avanzamento. (La comparsa o meno di questa barra dipende dalle dimensioni del documento e dalla complessità della formattazione.)

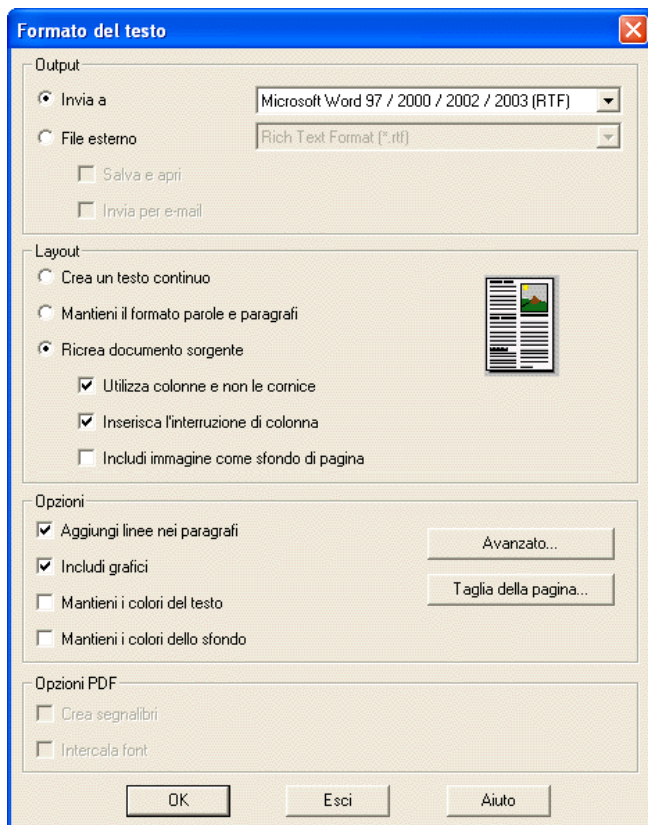


L'immagine acquisita viene nuovamente visualizzata con la divisione in zone creata, per essere disponibile per ulteriori elaborazioni, restando visualizzata fino a quando non si esegue la scansione di un'altra pagina.

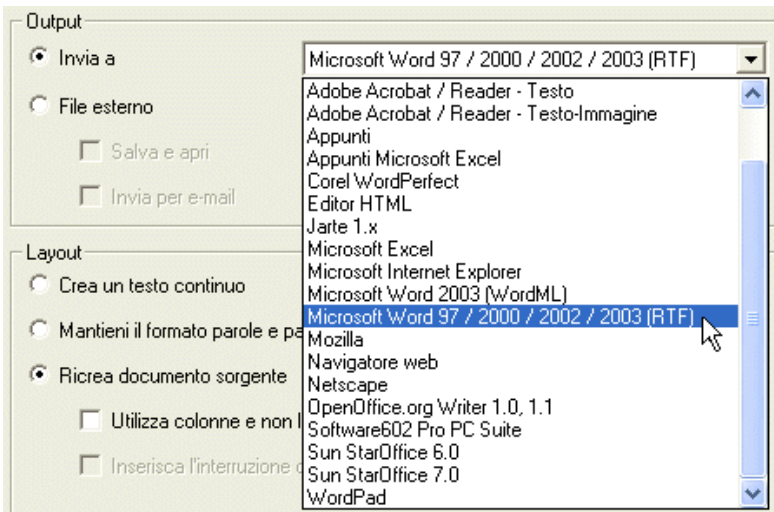
In questo modo un documento cartaceo viene effettivamente convertito in un file editabile, impiegando tempi fino a 40 volte inferiori rispetto a quelli necessari per la ridigitazione manuale. Procedere e confrontare il testo con l'immagine posta all'interno della finestra di Readiris.

Readiris dispone di tre metodi diversi per salvare il risultato OCR: inviare direttamente il documento riconosciuto a una applicazione scelta, salvare il risultato in un file esterno e copiare il risultato negli appunti di Windows.

Utilizzando il tasto "Formato" sulla barra degli strumenti principale (o il comando "Formato del Testo" nel menu "Impostazioni") viene selezionata la **destinazione**.



La funzione "Invia a" consente un collegamento OCR diretto fra lo scanner e le applicazioni Windows: sarà possibile **inviare** i documenti acquisiti direttamente all'elaboratore di testi, al foglio elettronico, al browser web o ad Adobe Reader ecc.



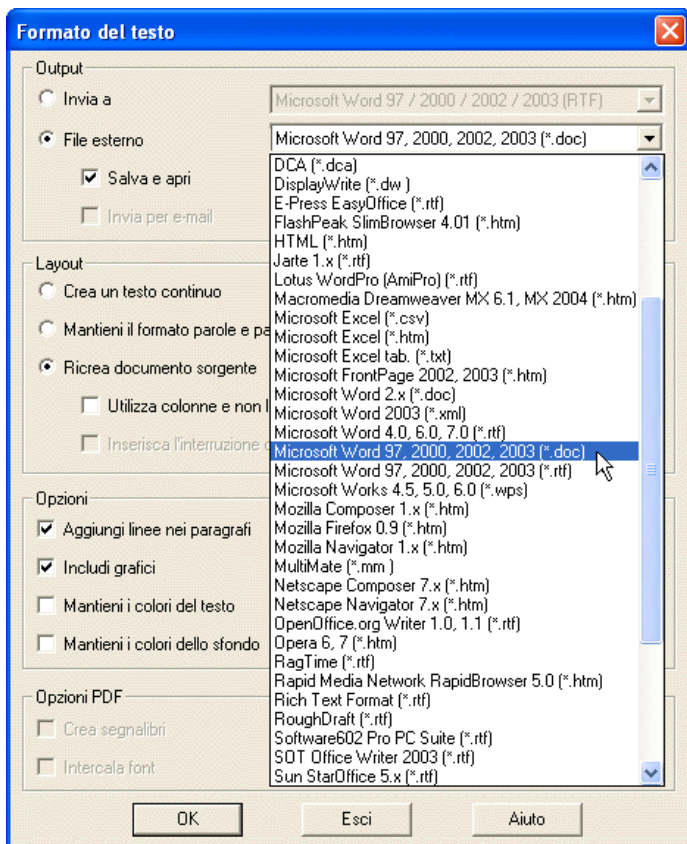
Al termine del riconoscimento, viene avviata l'applicazione scelta e aperto il documento riconosciuto all'interno di un nuovo file di testo o foglio di lavoro.



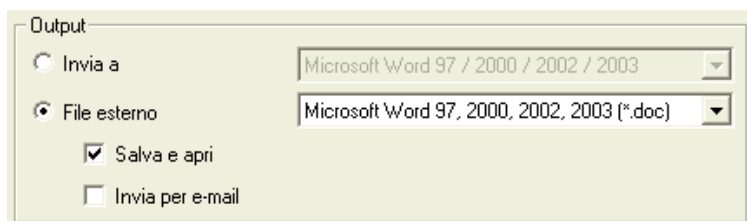
Non dimenticare che l'opzione "Invia a" consente anche di copiare il testo riconosciuto negli **appunti** di Windows. Non vi è quindi alcuna necessità di esportare il risultato... o di salvarlo in un file esterno!

SALVATAGGIO DEI RISULTATI IN UN FILE DI TESTO

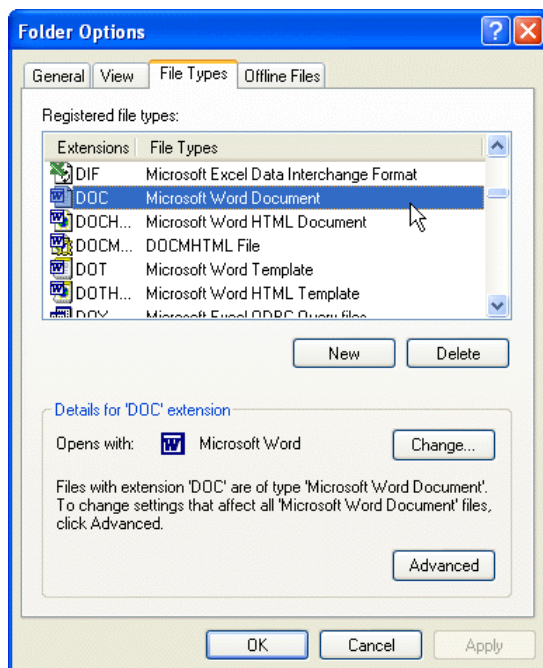
È possibile scrivere il risultato OCR in un file "esterno". Come detto, Readiris supporta un'ampia gamma di formati file comprendenti tutti i sistemi di elaborazione testi, fogli elettronici, applicazioni web ecc. (Readiris supporta anche il formato WordML, il nuovo formato di testo di Microsoft Office 2003!)



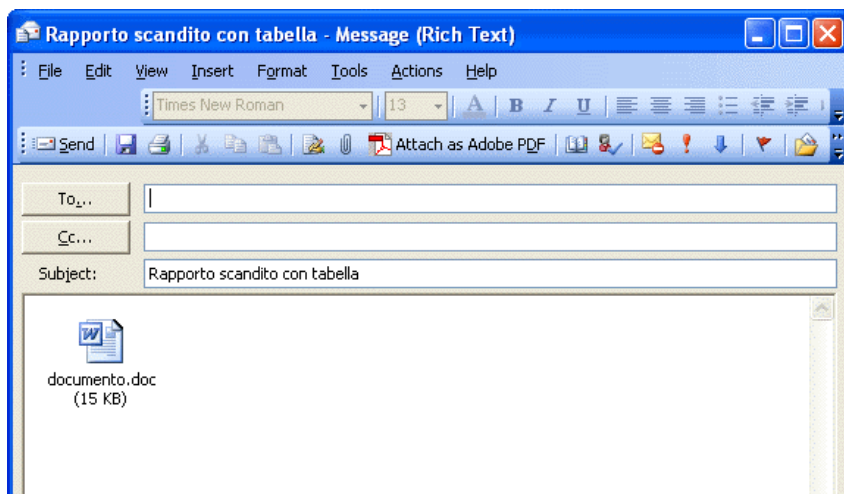
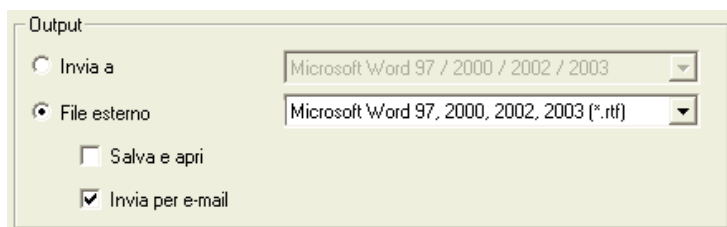
L'opzione "Salva e Apri" è simile alla funzione "Invia a": si apre il documento riconosciuto una volta salvato.



Tuttavia il metodo utilizzato per inviare alle applicazioni scelte è diverso. In questo caso i **tipi file di Windows** definiscono le applicazioni da avviare. È come se si facesse doppio clic sul file di output in Esplora Risorse di Windows... Con l'opzione "Invia a", Readiris invia direttamente alle applicazioni scelte.

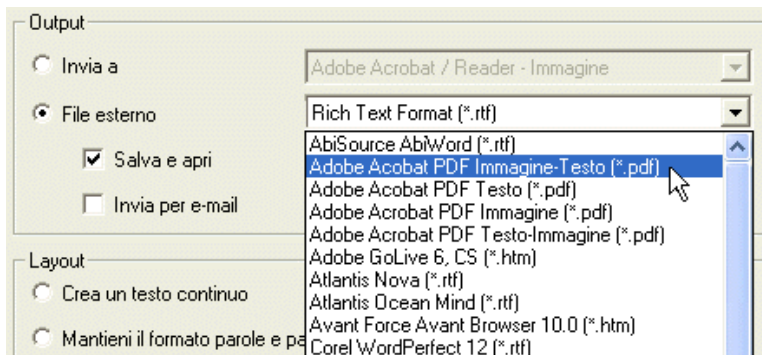
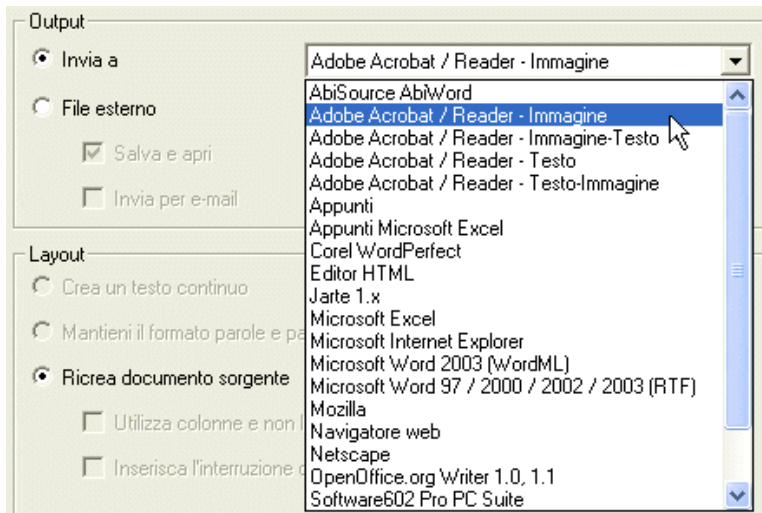


L'opzione "Invia per e-mail" crea un nuovo messaggio **e-mail** ed inserisce il documento riconosciuto come allegato. Esiste un modo più veloce per distribuire un documento cartaceo tanto rapidamente...?



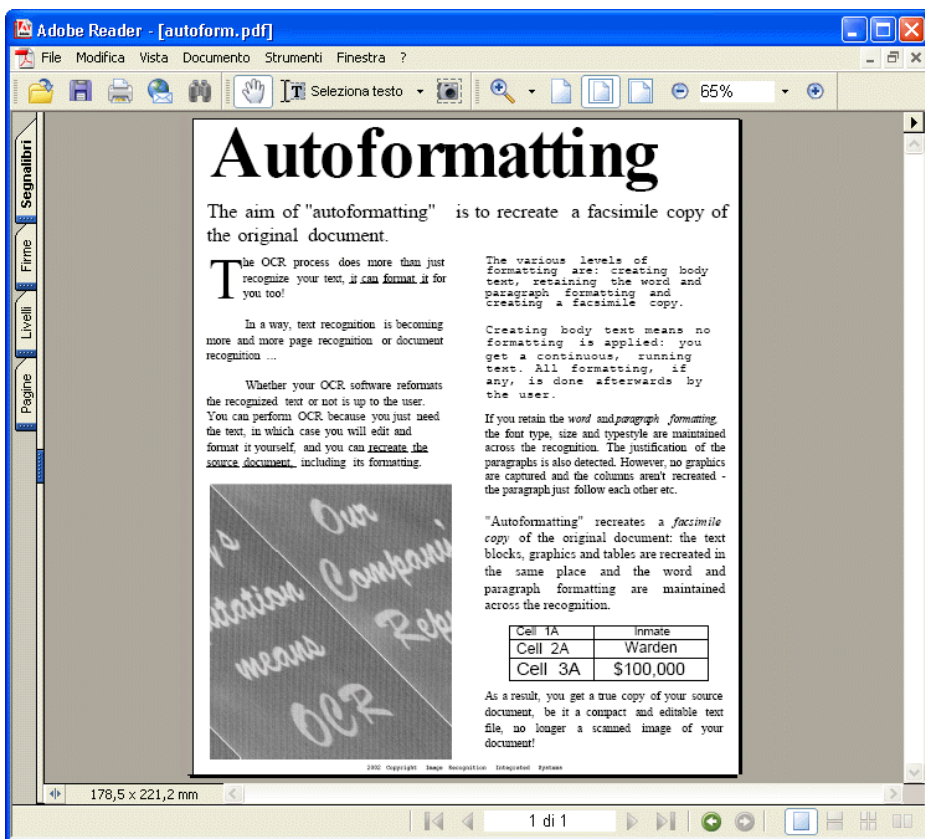
CREAZIONE DI UN DOCUMENTO ESPORTABILE

Verrà esaminato più nel dettaglio un formato: **Adobe Acrobat PDF**. Readiris consente di creare documenti PDF basati su testo e immagine.



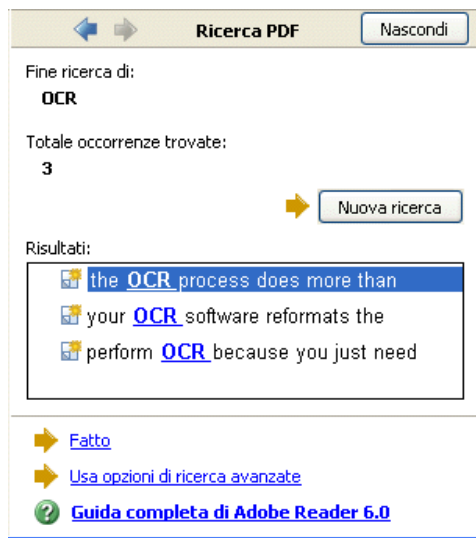
Qual è la differenza tra questi formati? Quando si seleziona il formato "PDF Testo", Readiris crea un file PDF contenente il risultato del testo. (Il risultato può contenere elementi grafici, ma solo quando nella pagina sono presenti zone grafiche)

come fotografie, disegni ecc.) In altre parole, l'immagine della pagina *non* è contenuta nel file PDF a un solo livello! Anche il formato "PDF immagine" è del tipo a un livello, ma naturalmente contiene solo l'immagine acquisita, non i risultati dell'OCR.

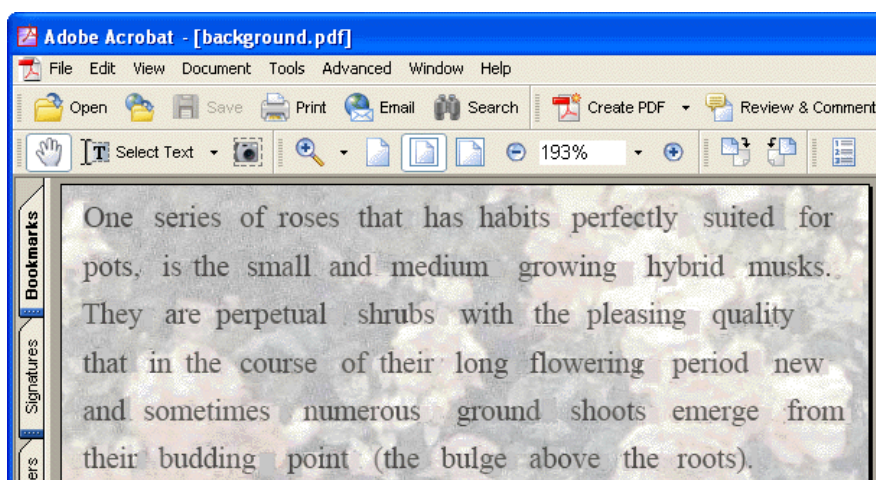
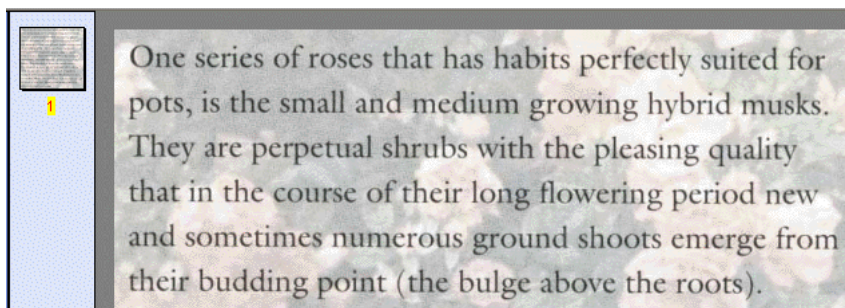




I formati "PDF Testo-Immagine" e "PDF Immagine-Testo" producono risultati diversi: Readiris crea un file ricercabile PDF contenente il testo riconosciuto e l'immagine della pagina. Nei file PDF "Testo-Immagine" l'immagine della pagina viene posizionata *sotto* il testo, generando un file PDF a due livelli, mentre nei file PDF "Immagine-Testo" l'immagine della pagina viene posizionata *sopra* il testo. Utilizzare la funzione di ricerca di Adobe Reader per semplificare l'operazione.



I file PDF del tipo "testo-immagine" sono abbastanza particolari: i pixel del testo riconosciuto vengono cancellati per creare un documento leggibile! Ad esempio, se si visualizza il testo riconosciuto in nero sopra bitmap di caratteri neri si ottiene un testo con una pesante ombreggiatura. Per fare una prova, l'utente può utilizzare l'immagine contenuta nel file BACKGROUND.JPG. (In Readiris Corporate sono disponibili funzioni simili per altri formati di testo...)



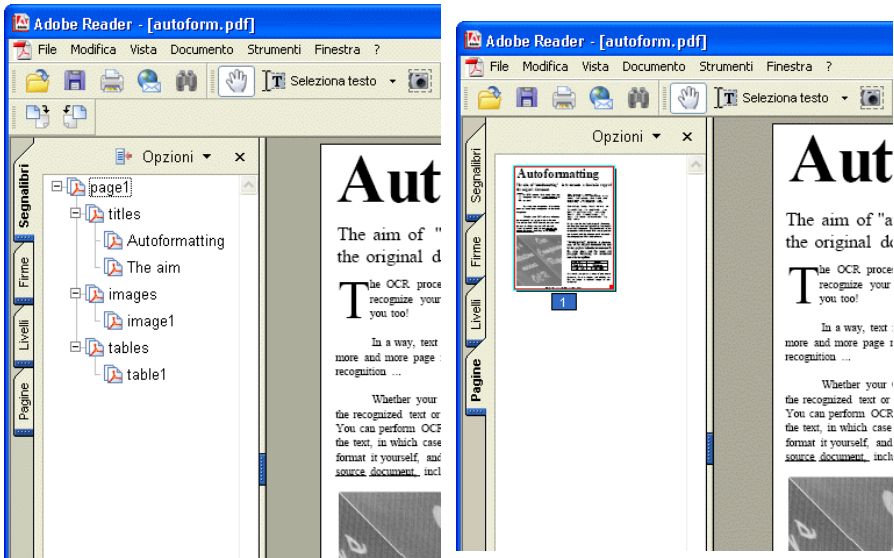
I file PDF basati sul testo includono gli URL dei siti web come collegamenti visibili, quindi è sufficiente fare clic su di loro per passare al corrispondente sito web!



Fare clic sul pulsante "Formato" per scoprire alcune opzioni relative al formato Acrobat PDF: "Crea Segnalibri" e "Incorpora Font".

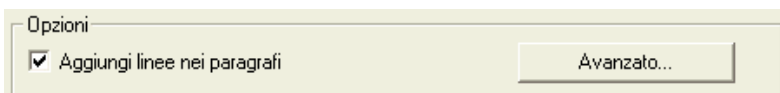


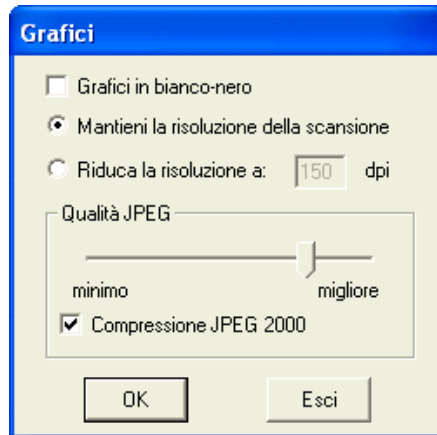
L'opzione "Crea Segnalibri" consente di creare **segnalibro** ("bookmarks") per ogni elemento del documento - per gli elementi grafici come per i blocchi di testo e le tabelle. Per le zone di testo, Readiris applica un algoritmo intelligente per esprimere un titolo, una sorta di "riepilogo" per ciascuna zona; le tabelle e gli elementi grafici vengono semplicemente numerati. (È possibile creare dinamicamente con il software Adobe Reader un altro elemento per lo spostamento nei documenti PDF, le **miniature**.)



L'opzione "Intercala Font" inserisce i font nei file PDF. Incorporando i font si impedisce la sostituzione dei font quando l'utente visualizza e stampa il documento riconosciuto. Garantisce che gli utenti, indipendentemente dalla configurazione del computer, vedano il testo con i font originali. Tuttavia, incorporando i font, si aumenta (leggermente) la dimensione del file dei documenti riconosciuti!

Un'altra opzione dei file PDF è "nascosta" tra le opzioni grafiche avanzate della finestra di dialogo "Formato". Quando si creano file PDF, è possibile selezionare il metodo di compressione delle immagini al loro interno: JPEG o JPEG 2000. (JPEG 2000 è la nuova e più compatta versione dello standard JPEG.)





... O LEGGERLI

Considerando un punto di vista opposto, dal momento che Readiris offre un supporto totale del formato PDF Adobe Acrobat, non si vogliono solo creare file PDF... ma anche *leggerli!*

Convertire documenti PDF può essere una delle più utili applicazioni di Readiris. Diversi sono i motivi. Innanzitutto, si tratta di un modo di convertire immagini in testo: aprire i documenti PDF con le immagini, eseguire il riconoscimento e salvare il risultato OCR in un documento di testo, in qualsiasi formato di testo supportato. I file di testo sono editabili, i file immagini non lo sono.

Secondo caso: è possibile convertire file PDF con immagini in documenti PDF di testo. Si esegue in tal caso il riconoscimento sui file PDF “solo immagini” e si salvano i risultati OCR... come documenti PDF di testo. I file PDF di testo sono ricercabili e modificabili, i file PDF “solo immagini” non lo sono.

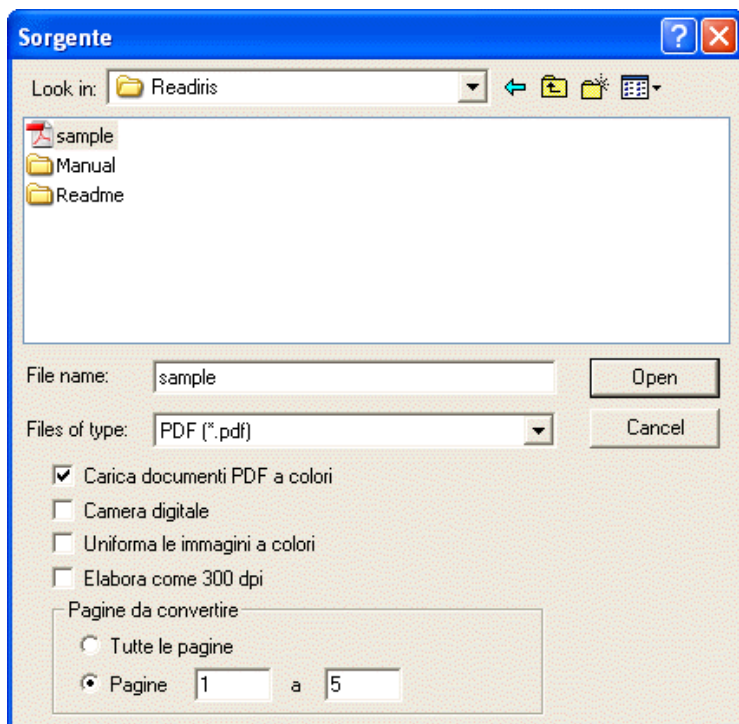
Inoltre convertire file PDF è un modo per “sbloccare” il contenuto PDF. È possibile riconoscere documenti PDF di “sola lettura” il cui testo è normalmente

inaccessibile. Con file PDF non protetti, il contenuto può essere recuperato (copiato e salvato in un file di testo), mentre con i file di "sola lettura" il contenuto non può essere estratto. Questi documenti possono essere solo visualizzati e stampati!

Vale la pena sottolineare due aspetti importanti: Readiris non apre documenti PDF protetti da password, ma riesce a superare tutte le altre barriere di sicurezza. Nello specifico, Readiris è in grado di accedere ai documenti PDF protetti da "password per le autorizzazioni", ma non a quelli protetti da "password di apertura" che definiscono gli utenti autorizzati ad accedere al documento. Inoltre, Readiris non converte documenti PDF contenenti immagini compresse JPEG 2000.

Procedere come a solito: memorizzare i file PDF aprendo le immagini preacquisite, come fax, istantanee scattare con fotocamera digitale ecc. Fare clic sul pulsante "Stop" o premere il tasto Esc per interrompere il processo di caricamento fra due pagine.

È disponibile un'opzione specifica per i file PDF. È possibile aprirli come documenti a colori e in bianco e nero. Viene offerta questa opzione perché la scansione di documenti a colori è molto più lenta...



In secondo luogo, è possibile indicare le pagine specifiche che si desidera convertire. Se si ha la necessità di acquisire solo un capitolo di una lunga pubblicazione in formato PDF, non ha senso caricare tutto il libro... Si risparmierà molto tempo indicando solo l'**intervallo di pagine** desiderato. (Lo stesso vale per le immagini TIFF a più pagine.)

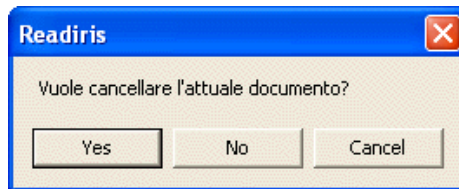
RICONOSCIMENTO DI PAGINE MULTIPLE

Dopo l'OCR l'immagine acquisita viene rivisualizzata con la suddivisione in zone creata per la successiva elaborazione.

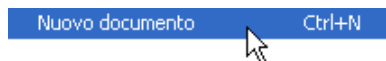
A questo punto è possibile aprire il testo riconosciuto con il programma di videoscrittura o editor di testi, importarlo nel software di desktop publishing o in qualsiasi applicazione basata su testo. Proseguire e confrontare il testo con l'immagine presente nella finestra Readiris.

Ma come salvare il testo di pagine successive o, in altri termini, come elaborare documenti costituiti da più pagine? In realtà è molto semplice: proseguire con il riconoscimento delle pagine e salvare i risultati nello stesso file. Verificare che il file non sia aperto, poiché in tal caso non sarebbe possibile scrivervi. In secondo luogo ricordarsi di impostare il dizionario dei font in modalità "aggiungi", così da poter aggiungere il testo e continuare comodamente l'apprendimento dei font.

Non appena verranno acquisite altre pagine o aperti file immagine all'interno di un documento, si dovrà decidere se iniziare un nuovo documento o completare il documento corrente.



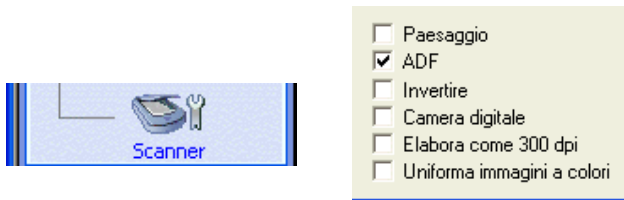
Rispondere "No" per aggiungere le pagine al documento corrente, rispondere "Sì" per creare un nuovo documento. Questa risposta ha lo stesso effetto del comando "Nuovo documento" del menu "File".



Ma è disponibile un metodo più efficiente per riconoscere diverse pagine rispetto alla scansione e all'OCR pagina dopo pagina: l'elaborazione diretta di **documenti a più pagine**.



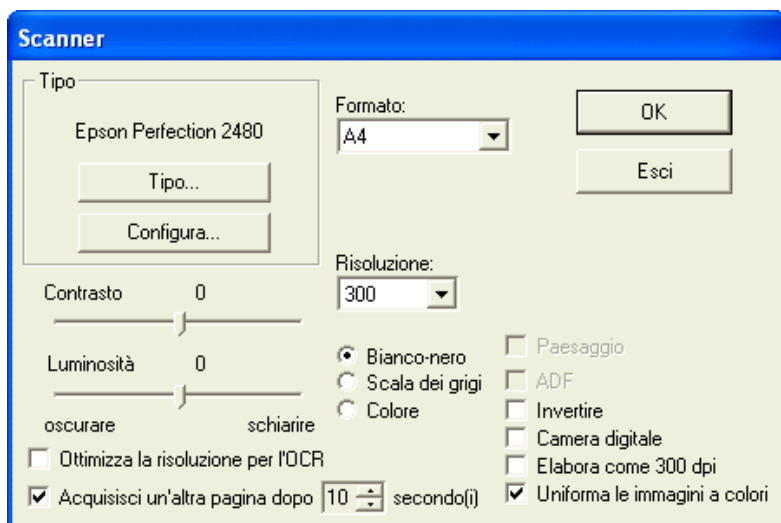
Per acquisire un documento costituito da diverse pagine in un'unica operazione, attivare l'alimentatore di fogli dello scanner con l'opzione "ADF" disponibile facendo clic sul tasto "Scanner".



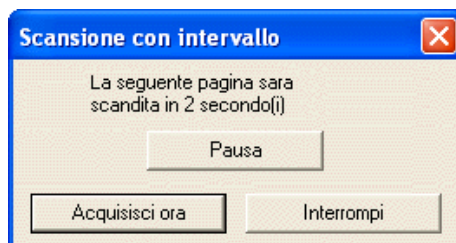
Inserire le pagine del documento nell'alimentatore fogli automatico ed iniziare la scansione: tutte le pagine vengono acquisite fin quando l'alimentatore fogli risulta vuoto.

Se lo scanner piano utilizzato *non* è dotato di alimentatore di fogli, la **scansione a intervalli** consente di acquisire facilmente documenti composti da più pagine. Lo scanner acquisisce automaticamente la pagina successiva dopo un intervallo in secondi impostato dall'utente, durante il quale sarà possibile sostituire la pagina.

È possibile indicare l'intervallo desiderato accedendo alla pagina delle impostazioni dello scanner: fare clic sul pulsante "Scanner" e indicare il valore desiderato nell'opzione "Acquisisci Un'Altra Pagina dopo x Secondo(i)".



A questo punto, fare clic sul pulsante "Scandire" per avviare il processo. Per terminare la scansione automatica, fare clic su "Interrompi" nella finestra di dialogo della scansione a intervalli.

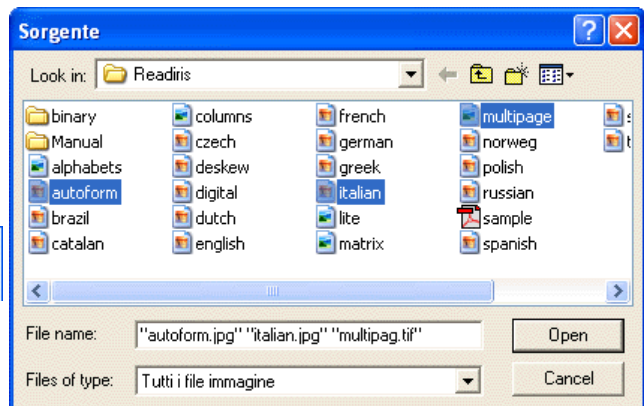


Readiris offre la massima flessibilità: è possibile evitare l'intervallo di scansione per eseguire una scansione immediata, interrompendo momentaneamente il processo, per poter, ad esempio, rispondere al telefono. Fare clic su "Pausa" nella

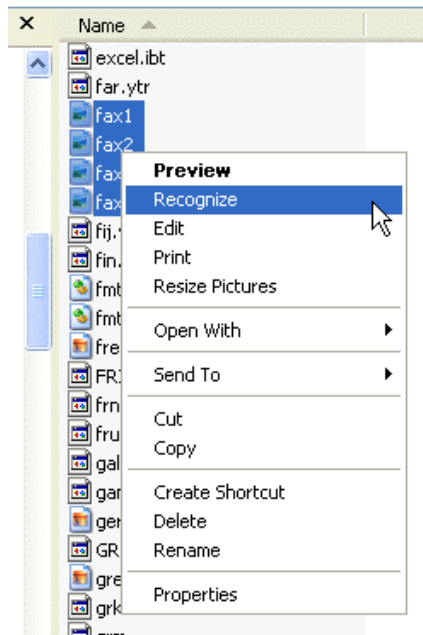


finestra di dialogo della scansione automatica per interrompere momentaneamente la scansione a intervalli. Fare clic su "Riprendi" per riprendere l'operazione. In alternativa, fare clic su "Acquisisci Ora" nella finestra di dialogo della scansione a intervalli per avviare immediatamente la scansione: l'intervallo viene abbreviato!

È anche possibile *aprire* più immagini preacquisite. Per caricare diverse immagini, selezionare la prima immagine e tenere premuto il tasto Ctrl durante la selezione delle altre immagini. Per caricare intervallo continuo di immagini, selezionare la prima immagine e tenere premuto il tasto Maiusc mentre si seleziona l'ultima immagine.



Lo stesso effetto può essere ottenuto facilmente da Esplora Risorse di Windows: selezionare vari file di immagine, fare clic con il pulsante destro del mouse e selezionare il comando "Riconoscimento" dal menu di scelta rapida. È possibile ripetere questa operazione: tutte le immagini inviate a Readiris vengono aggiunte al documento corrente fino a quando non si farà clic sul comando "Nuovo Documento".



È possibile anche *trascinare* diverse immagini preacquisite da Esplora Risorse di Windows alla finestra di Readiris. Vale lo procedimento: tutte le immagini trascinate sulla finestra Readiris vengono aggiunte al documento corrente fino a quando non si farà clic sul comando "Nuovo Documento".

Readiris ordina le immagini automaticamente: l'immagine 001.tif precede 002.tif che precede 003.tif ecc.

Il **pannello dei documenti** mostra le **statistiche** e le informazioni relative alle pagine acquisite come origine e risoluzione dell'immagine, tempo necessario alla scansione e al riconoscimento ecc. È possibile nascondere e visualizzare il pannello dei documenti utilizzando l'opzione corrispondente del menu "Visualizzazione".



✓ Pannello dei documenti

Inoltre, si ottengono informazioni su quale immagine è stata memorizzata. Se è stata aperta un'immagine a più pagine, ovviamente esiste un solo file per tutte le immagini. Durante la *scansione* di documenti, il pannello dei documenti indica solo il modello dello scanner.

Pagina	Fonte d'immagine	Scansione	Letture	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\multipage.tif	1.11	-	300	2000	2388	-	-
<input checked="" type="checkbox"/> 2	C:\Program Files\Readiris\multipage.tif	0.82	-	300	2000	1888	-	-
<input checked="" type="checkbox"/> 3	C:\Program Files\Readiris\multipage.tif	0.77	-	300	1912	2004	-	-

Trascinare il cursore di ridimensionamento colonne per modificare le dimensioni di una colonna. (Non è possibile modificare l'*ordine* delle colonne.) Se la colonna è troppo stretta per visualizzare i dati, posizionare su di essa il cursore del mouse: i relativi dati verranno visualizzati nella descrizione corrispondente.

Fonte d'immagine	Scansione
C:\Program Files\...	2.81

Pagina	Fonte d'immagine	Scansione
<input checked="" type="checkbox"/> 1	C:\Program Files\Readiris\italian.tif	

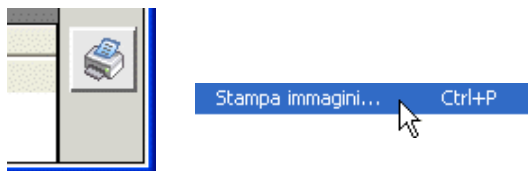
È possibile visualizzare le stesse informazioni per tutte le pagine utilizzando il comando "Informazioni" del menu "File", ma le informazioni relative alle singole pagine possono anche essere visualizzate posizionando il cursore sulla miniatura di una delle pagine nella **barra strumenti della pagina** a sinistra dello schermo. Questa barra degli strumenti viene visualizzata appena si inizia l'elaborazione di una pagina e consente di accedere ai comandi della pagina (utilizzando il pulsante destro del mouse).



La pagina corrente è evidenziata nel pannello dei documenti e nella barra strumenti della pagina, nonché menzionata nella barra del titolo di Readiris.

Per cancellare una pagina, selezionarla nel pannello dei documenti e premere il tasto Canc. (In alternativa, selezionarla nella barra strumenti della pagina, fare clic con il pulsante destro del mouse e selezionare il comando "Cancella la Pagina" dal menu di scelta rapida.)

Per avere una panoramica del documento, è possibile **stampare** rapidamente le **immagini** acquisite facendo clic sul pulsante "Stampa" nella barra degli strumenti immagine (o selezionando il comando "Stampa le Immagini" del menu "File").

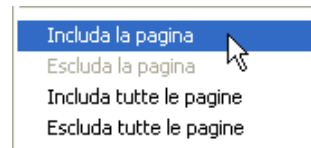


È possibile stampare la pagina corrente o tutte le pagine. Selezionare il numero di pagine o miniature che si desidera stampare su una pagina.

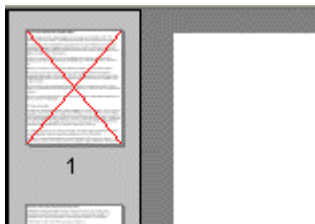


Ma non è necessario stampare tutte le pagine: è infatti possibile escludere (provvisoriamente) determinate pagine utilizzando il pannello dei documenti, nonché i comandi corrispondenti del menu "Modifica" e i comandi del menu di scelta rapida. È sufficiente fare clic sul numero di pagina nel pannello dei documenti per escluderla dalla stampa (e dal processo di riconoscimento). Per reinserirla, fare di nuovo clic. Per maggiore flessibilità, il menu "Modifica" contiene comandi equivalenti che si applicano a *tutte* le pagine.

Pagina	Fonte d'immagine	Scansione
<input type="checkbox"/> 1	C:\Program Files\Readiris\...	2.45
<input checked="" type="checkbox"/> 2	C:\Program Files\Readiris\...	2.25
<input checked="" type="checkbox"/> 3	C:\Program Files\Readiris\...	2.55



Le miniature delle pagine escluse vengono sbarrate. Ricordare che è sempre possibile stampare la pagina corrente, anche se momentaneamente “disattivata”!



Caricare l'immagine campione MULTIPAGE.TIF e iniziare il riconoscimento. Le varie pagine vengono visualizzate l'una dopo l'altra; la barra del titolo di Readiris indica il numero di pagina.



The screenshot shows the Readiris software interface. The main window displays a document with several lines of text highlighted in orange. A small dialog box titled "Riconoscimento in corso (4/5)" is overlaid on the document, showing a progress bar at 47% and a "Stop" button. The interface includes a menu bar (File, Modifica, Impostazioni, Visualizzazione, Azione, Apprendimento, Registra, Guida), a left sidebar with navigation options (Scandire, Riconoscere, Salvare), and a bottom table of document pages.

Pagina	Fonte d'immagine	Scansione	Lettura	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 3	C:\Program Files\...	1.56	1.53	300	1912	2004	-	-
<input checked="" type="checkbox"/> 4	C:\Program Files\...	1.60	2.54	300	2004	2000	-	-
<input checked="" type="checkbox"/> 5	C:\Program Files\...	1.66	1.27	300	1984	2628	-	-

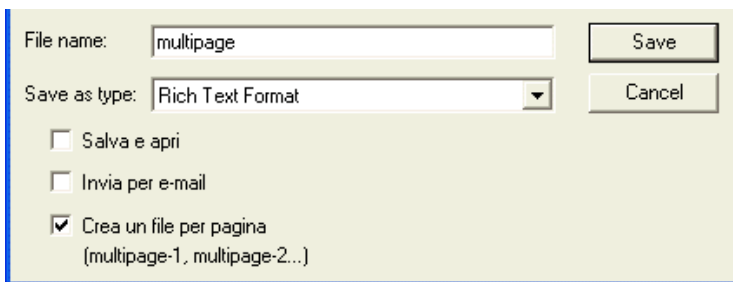
Se l'apprendimento interattivo è attivato, le fasi di riconoscimento ed apprendimento vengono eseguite pagina per pagina. La modalità dizionario "nuovo" viene utilizzata per la prima pagina mentre la modalità "aggiungi" per le pagine successive.

Premendo il tasto "Termina", tutte le decisioni del sistema vengono accettate senza necessità di convalida da parte dell'utente. In altre parole, l'apprendimento

interattivo viene annullato per *tutte* le pagine; l'OCR di questo documento continua in modo automatico.

Il risultato del riconoscimento di documenti a più pagine viene salvato in un unico file di output. (Quando il risultato del riconoscimento viene inviato ad un'applicazione destinataria, vengono create pagine multiple all'interno di un unico documento.)

Ciò accade quando l'opzione "Crea Un File per Pagina" è disattivata mentre si salva il documento riconosciuto. Questa opzione prevede che ogni pagina di un documento a più pagine sia salvata in un file separato. Se si assegna al file il nome testo.doc, i file verranno denominati testo-1.doc, testo-2.doc ecc. (Questa opzione non è disponibile quando si inviano i risultati dell'OCR a un'applicazione scelta, ma solo quando si crea un file esterno.)



File name:

Save as type:

Salva e apri

Invia per e-mail

Crea un file per pagina
(multipage-1, multipage-2...)

MODIFICA DI DOCUMENTI A PIÙ PAGINE

L'utente può modificare documenti a più pagine soprattutto per correggere errori di scansione: può cancellare una pagina dal documento o spostare pagine in altri punti dello stesso.

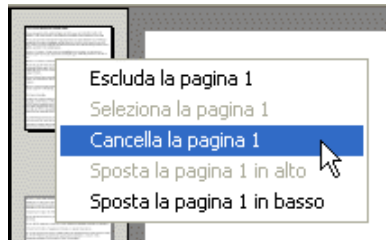
Si inizia con lo spostamento. Per *passare a una pagina* specifica, fare clic su di essa nel pannello dei documenti o nella barra strumenti della pagina. (È anche possibile posizionare il cursore sulla relativa miniatura nella barra strumenti della pagina, fare clic con il pulsante destro del mouse per aprire il menu di scelta



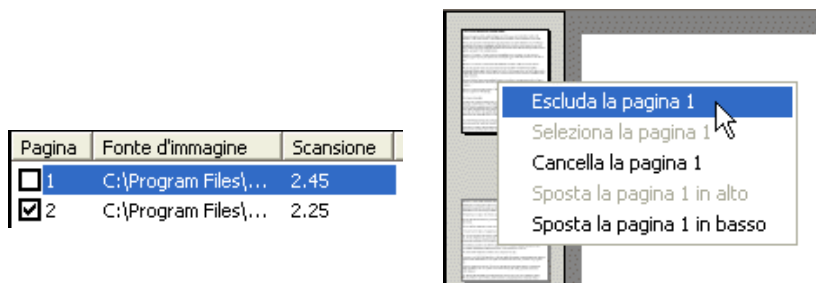
rapida e selezionare il comando "Seleziona la Pagina".) Per andare alla pagina precedente, premere PgSu, per andare alla pagina successiva premere PgGiù. Premere Home per andare alla prima pagina e premere Fine per andare all'ultima pagina. In alternativa, utilizzare i comandi corrispondenti del menu "Visualizzazione".

Prima pagina	Home
Pagina precedente	PageUp
Pagina successiva	PageDown
Ultima pagina	End

Modifichiamo ora il documento. Per *cancellare una pagina*, selezionarla nel pannello dei documenti e premere il tasto Canc. In alternativa, è possibile posizionare il cursore del mouse sulla relativa miniatura, fare clic con il pulsante destro del mouse per visualizzare il menu di scelta rapida e selezionare il comando "Cancella la Pagina".

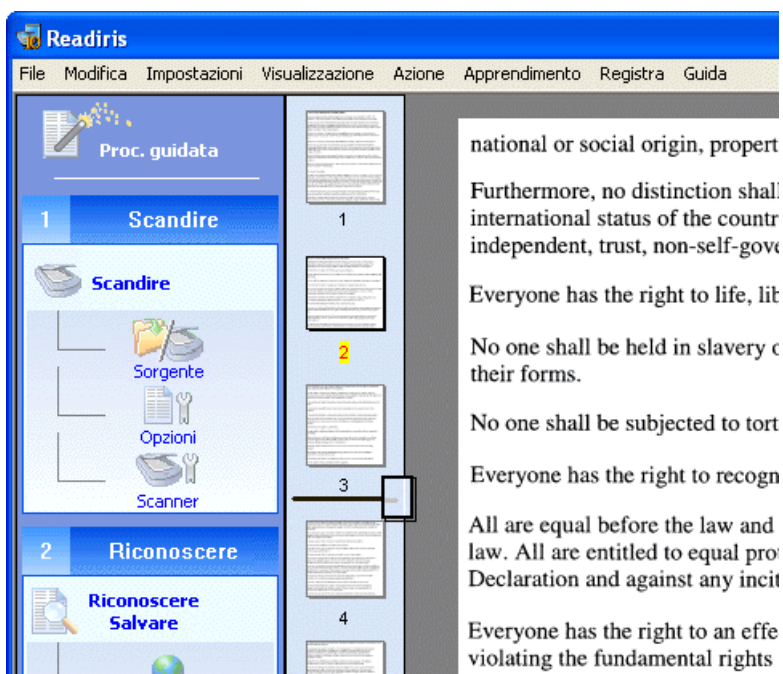


Si ricorda che è possibile escludere temporaneamente le pagine, non di eliminarle, dal processo di riconoscimento (e di stampa dell'immagine): il pannello dei documenti, la barra strumenti della pagina e il menu "Modifica" contengono i comandi necessari.



Per *spostarsi alla pagina precedente*, utilizzare il comando "Sposta la Pagina in Alto" e per *spostarsi alla pagina successiva* utilizzare il comando "Sposta la Pagina in Basso".

Per *spostare una pagina* in un punto del documento completamente diverso, trascinatene l'icona nel nuovo punto.



CREAZIONE DI UN NUOVO DOCUMENTO

È possibile utilizzare il comando "Nuovo Documento" del menu "File" per chiudere il documento corrente.

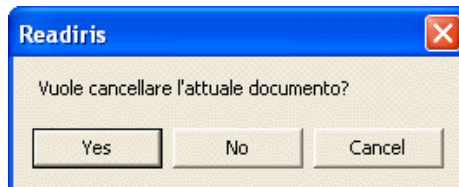
Nuovo documento Ctrl+N



Questo comando “cancella la lavagna”. Qualsiasi documento caricato in memoria, contenente una pagina singola o più pagine, viene cancellato. A questo punto è possibile creare un nuovo documento.

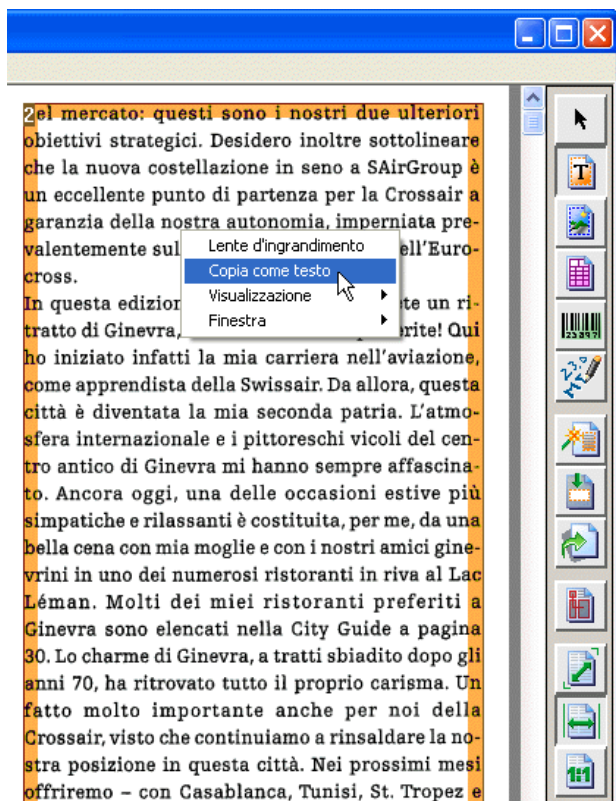
È possibile creare un nuovo documento anche dall'interno del documento corrente. Se l'OCR non è stato eseguito, il sistema deduce che si desidera aggiungere altre pagine al documento corrente. Ad esempio, è possibile acquisire tutte le pagine contenute nell'alimentatore automatico dello scanner, riempire nuovamente l'alimentatore e ripartire. Tutte le pagine acquisite comporranno un singolo documento. In alternativa è possibile acquisire una serie di pagine e aggiungere alcuni file immagine, ad esempio fax. Queste pagine formano a loro volta un unico documento. Tutto quello che occorre fare è cambiare l'origine dell'immagine utilizzando il tasto "Origine".

Se, una volta *ultimato* l'OCR, viene riavviata la scansione o eseguito il caricamento delle immagini, il sistema chiede di creare un nuovo documento o di completare il documento corrente.



RICONOSCIMENTO DI ZONE DI TESTO

Si è visto fino ad oracome riconoscere le pagine e come elaborare un documento a più pagine. Ma è possibile riconoscere con la stessa facilità meno di una pagina? Sì, è possibile! Fare clic con il pulsante destro del mouse e selezionare il comando "Copia come Testo" dal menu di scelta rapida: la finestra di testo sotto il mouse viene riconosciuta e inviata agli appunti.



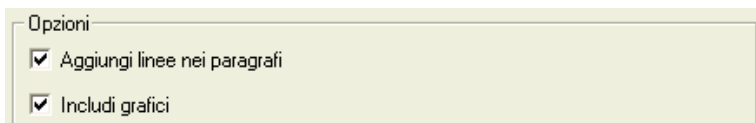
Si applicano le impostazioni di sistema correnti, come lingua, tipo di font ecc. Il risultato dell'OCR viene collocato negli appunti come "testo continuo", non formattato.

ORGANIZZAZIONE DEL DOCUMENTO DI TESTO

Salvare o esportare il testo significa molto più che selezionare una modalità di output o assegnare un nome al file di output. È possibile anche selezionare un formato file e determinare l'aspetto del testo riconosciuto. In breve, prima di lanciare l'esecuzione è necessario decidere come gestire il testo.

Alcune opzioni del tasto "Formato" consentono di modificare l'aspetto del testo riconosciuto.

Il **flusso di testo** del documento di output è direttamente influenzato dall'opzione "Aggiungi Linee nei Paragrafi".



Per fare in modo che Readiris individui i paragrafi, mantenere questa opzione abilitata: Readiris applicherà così il normale **ritorno a capo automatico** tipico degli elaboratori di testo, altrimenti dopo ciascuna riga verrebbe aggiunto un ritorno a capo e le parole separate dal trattino rimarrebbero divise. L'individuazione dei paragrafi è abilitata per impostazione predefinita.

Segue un esempio. Quando le prime tre righe di una colonna sono "Il nuovo presi-", "dente salutava dal balcone." e "Sua moglie lo aveva raggiunto.", l'individuazione del paragrafo fornisce il seguente risultato: "Il nuovo **presidente** salutava dal balcone. **Sua** moglie lo aveva raggiunto." Le parti separate dal trattino della parola "presidente" sono state "reincollate" ed è stato aggiunto uno spazio alla fine della frase, creando un testo continuo.

Se l'individuazione del paragrafo *non* fosse stata abilitata, sarebbe stato mantenuto il layout originale, con un ritorno a capo aggiunto alla fine di ogni riga.

L'opzione *non* è valida quando il formato PDF è selezionato: i file Adobe Acrobat PDF memorizzano sempre il testo riga per riga!

(Il pulsante "Formato" contiene alcune opzioni di formattazione non ancora trattate: l'argomento verrà trattato nelle prossime pagine.)



CONFIGURAZIONE DELLO SCANNER

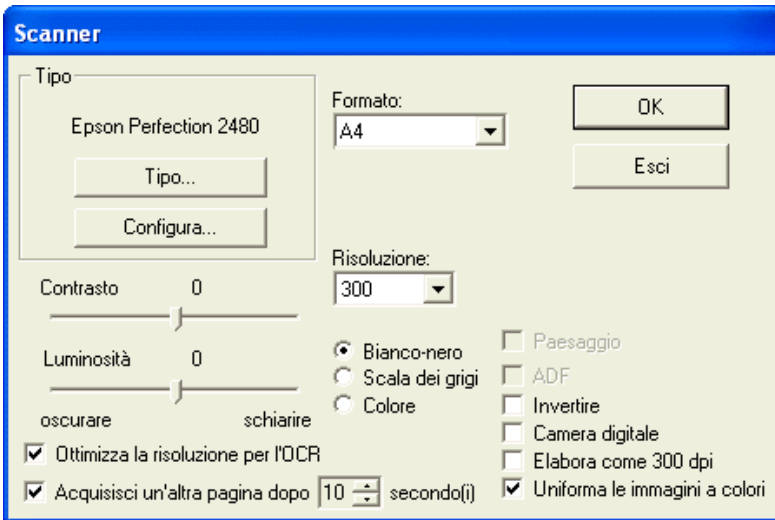
A questo punto è necessario configurare lo scanner. Si presume che l'hardware dello scanner e i piloti siano stati installati correttamente.

Nel caso in cui la licenza del software Readiris sia collegata a un modello di scanner, questa operazione può non essere necessaria in quanto lo scanner potrebbe già essere installato sotto Readiris.

Fare clic sul pulsante "Scanner" della barra degli strumenti principale.

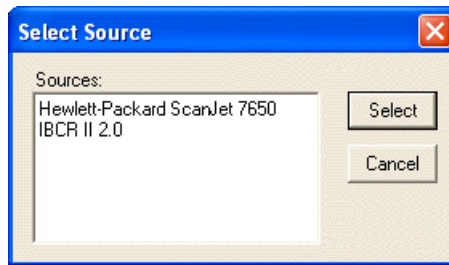


Per determinare il **modello dello scanner** in dotazione, fare clic sul pulsante "Tipo".



Quando si seleziona "<Image>" (immagine) come "scanner", le immagini preacquisite funzionano sempre come origine delle immagini; non sarò necessario selezionare il disco come origine delle immagini utilizzando il pulsante "Sorgente" della barra degli strumenti principale.

Il pulsante "Configura" è disponibile solo quando lo scanner ne permette l'utilizzo. Tale pulsante consente l'accesso a parametri di scansione avanzati; con gli scanner Twain, facendo clic sul pulsante "Configura" è possibile selezionare l'origine Twain. (È possibile utilizzare anche il comando "Seleziona Origine" dal menu "File".)



Una volta selezionato lo scanner, la stessa finestra consente di impostare la risoluzione di scansione, il formato e l'orientamento della pagina, la luminosità ed il contrasto, consentendo inoltre di indicare se utilizzare il sistema di lettura automatica di documenti dello scanner o meno. Con scanner Twain compatibili, tutti i parametri di scansione vengono spesso impostati all'interno dell'interfaccia utente Twain.

Impostare la **luminosità** e, se disponibile, il **contrasto**.

Abilitando l'opzione "Paesaggio", la pagina selezionata sarò orientata in senso orizzontale, anzichè verticale. In realtà, l'orientamento della pagina viene applicato ai formati di pagina ridotti: con uno scanner piano A4 è possibile, ad esempio, eseguire la scansione di pagine A5 (le cui dimensioni sono dimezzate) in formato verticale o orizzontale, ma ovviamente è possibile eseguire la scansione di una superficie completa A4 solo in una direzione!



Infine, l'opzione "Invertire" permette di generare delle **immagini "invertite"** nella modalità di scansione in bianco e nero; potete attivare questa opzione per elaborare pagine intere con testo bianco su sfondo nero.

La scansione a intervalli è stata già illustrata. Si tratta di una modalità molto efficace per la scansione di documenti a più pagine quando lo scanner piano utilizzato non è dotato di alimentatore di fogli.

COME COLORARE LE SCANSIONI

Readiris supporta immagini in bianco e nero, in scala di grigi e a colori, pertanto è possibile scegliere la **modalità colore** che meglio risponde alle proprie esigenze. Per includere immagini "linear" nei documenti riconosciuti, eseguire la scansione in bianco e nero; per includere fotografie in bianco e nero eseguire la scansione in scala di grigi e per includere immagini a colori eseguire la scansione a colori.

Ma perché ridurre la profondità di bit dell'immagine durante la scansione? È superfluo precisare che le immagini in scala di grigi e a colori sono più lente da acquisire e richiedono una memoria RAM più ampia rispetto alle immagini a "due colori".

La scansione, in scala di grigi e a colori, non è utile soltanto per salvare gli elementi grafici con una qualità sufficiente, ma in alcuni casi è utile nonché necessaria per ottenere buoni risultati OCR! Quando il testo è stampato su uno sfondo a colori, la scansione a colori può creare differenze di tonalità che mancano nelle immagini in bianco e nero. Quando è presente soltanto un contrasto limitato fra il testo e lo sfondo, lo sfondo può creare elementi di "disturbo" che rendono difficile o addirittura impossibile il riconoscimento.

Si pensi, ad esempio, ad un testo nero stampato su uno sfondo scuro: durante la scansione di un documento di questo tipo in bianco e nero, potrebbe non essere possibile cancellare il colore dello sfondo senza perdere le informazioni di testo, malgrado tutto quello che è possibile fare per regolare la luminosità dello scanner...



Readiris crea una versione in bianco e nero per ogni immagine in scala di grigi e a colori. Grazie alle routine intelligenti è possibile risolvere anche i casi più difficili: ecco come viene binarizzata una immagine “difficile”!

MASAYOSHI SON, 42, president and CEO, is the master Net empire builder. His conglomerate holds stakes in 300 Internet companies in the U.S., Japan, Europe, and other Asian countries. Today, Softbank manages about \$4 billion in venture capital funds for global investments.

YASUMITSU SHIGETA, 35, has invested in more than 70 Web or mobile Net-based ventures in Japan and the U.S., including Tumbleweed Communications and Phone.com. Shigeta is also developing new businesses that take advantage of the growth of the Internet and mobile communications.



Per visualizzare un'immagine acquisita in bianco e nero, deselezionare l'opzione "Visualizza Documento a Colori" del menu "Visualizzazione".

✓ Visualizza documento a colori Ctrl+O



In realtà, non verranno visualizzate immagini in bianco e nero, anche quando si acquisiscono immagini a due colori! Ciò accade perché Readiris ottimizza le immagini per renderle perfettamente leggibili a schermo. La tecnologia I.R.I.S. di visualizzazione ad alta risoluzione consente di convertire immagini in bianco e nero in immagini in scala di grigi.

Reading dot matrix documents

You can read dot matrix document without changing the font mode. The software detects whether "normal" text or dot matrix printouts are being read.

Far out in the uncharted backwaters of the unfashionable end of the Western Spiral arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-two million miles is an utterly insignificant little blue green

Reading dot matrix documents

You can read dot matrix document without changing the font mode. The software detects whether "normal" text or dot matrix printouts are being read.

Far out in the uncharted backwaters of the unfashionable end of the Western Spiral arm of the Galaxy lies a small unregarded yellow sun. Orbiting this at a distance of roughly ninety-two million miles is an utterly insignificant little blue green

Le immagini in scala di grigi e a colori sono attenuate, sfumate.

A word about OCR

The aim of OCR is to automatically enter printed text documents in a very effective and low cost way. Although the first research and development on Optical Character Recognition (OCR) began more than 30 years ago, this technology is still unknown by most of the people who could use it for their document entry applications.

A word about OCR

The aim of OCR is to automatically enter printed text documents in a very effective and low cost way. Although the first research and development on Optical Character Recognition (OCR) began more than 30 years ago, this technology is still unknown by most of the people who could use it for their document entry applications.

Di conseguenza, non è necessario ingrandire, anche nei computer portatili con schermo LCD o computer desktop con schermo a 15" di bassa qualità.

Ingrandire e ridurre i particolari per visualizzare l'immagine "grezza" così come è stata acquisita!

APPARECCHI DIVERSI, DIVERSE RISOLUZIONI

Indipendentemente dalla modalità di scansione, selezionare una **risoluzione** di 300 dpi per applicazioni normali. Utilizzare una risoluzione di 400 dpi per stampe di dimensioni ridotte (inferiori a 10 punti) e quando la qualità del documento è estremamente bassa.

Readiris legge **dimensioni in punti** compresi fra 6 e 72 punti (da 0,21 a 2,54 cm.).

6 punti

72 punti

Readiris riconosce inoltre "**capilettiera**", ossia grosse maiuscole che coprono più righe. (Naturalmente queste non possono essere di dimensioni superiori a 72 punti.) Vengono riconosciuti anche i capilettiera a tonalità invertite...



Readiris legge letterine ("drop letters") che coprono diverse linee e le indirizza alla prima riga.

Le Festival de Wallonie est dans sa phase ultime et nale avec l'ouverture de ses dernières sections.
Cette semaine, c'est la branche

Un OCR ottimale richiede una risoluzione fra 300 e 400 dpi; Readiris avverte quando si elaborano immagini con una risoluzione inferiore ai 200 dpi o maggiore di 800 dpi. Inoltre, la risoluzione di queste immagini è contrassegnata in rosso all'interno del pannello dei documenti.

Pagina	Fonte d'immagine	Scansione	Lettura	Risoluzione	Larghezza	Altezza
<input checked="" type="checkbox"/> 1	C:\Program Files\...	1.70	-	96	1533	1241

Readiris è in grado di correggere scansioni con troppi dettagli! Attivare a tale scopo l'opzione "Ottimizza la Risoluzione per l'OCR" nelle impostazioni di scansione. Se la risoluzione dell'immagine delle scansioni supera i 600 dpi, la risoluzione viene ridotta per il processo OCR.

Ottimizza la risoluzione per l'OCR
 Acquisisci un'altra pagina dopo secondo(i)

Sono disponibili altri due modi per evitare questo avvertimento: quando si leggono fax con una risoluzione di 100 o 200 dpi, quando si creano immagini con una fotocamera digitale (dove la risoluzione non è nota) e quando si aprono immagini in cui l'intestazione del file contiene una risoluzione errata. Per elaborare queste immagini senza problemi, utilizzare l'opzione "Elabora come 300 dpi". Questa impostazione è adatta sia per la scansione diretta sia per l'apertura di immagini già acquisite.

Invertire
 Camera digitale
 Elabora come 300 dpi
 Uniforma immagini a colori

Carica documenti PDF a colori
 Camera digitale
 Uniforma le immagini a colori
 Elabora come 300 dpi

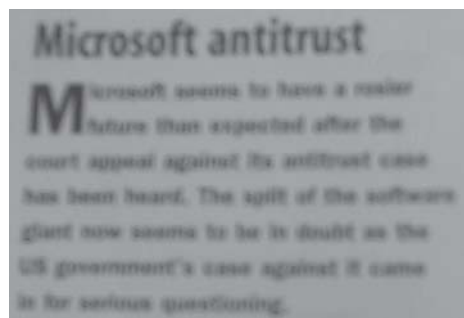
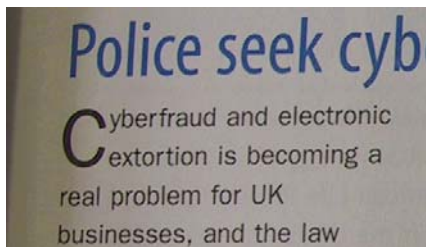
Se le immagini sono state acquisite con una **fotocamera digitale** anziché mediante uno scanner, è obbligatorio attivare un'opzione speciale che risulta valida anche per immagini acquisite direttamente o preacquisite.

<input type="checkbox"/> Invertire	<input type="checkbox"/> Carica documenti PDF a colori
<input checked="" type="checkbox"/> Camera digitale	<input checked="" type="checkbox"/> Camera digitale
<input checked="" type="checkbox"/> Elabora come 300 dpi	<input type="checkbox"/> Uniforma le immagini a colori
<input type="checkbox"/> Uniforma immagini a colori	<input checked="" type="checkbox"/> Elabora come 300 dpi

Operando in questo modo, si migliora l'immagine prima di riconoscerla. Esistono specifici aspetti di cui tenere conto se si tratta di fotocamere digitali: esse producono immagini a bassa risoluzione, anche si tiene la telecamera molto vicina ai documenti, e la risoluzione dell'immagine è in ogni caso sconosciuta.

Esistono ulteriori aspetti più delicati da prendere in considerazione per riconoscere adeguatamente le immagini riprese con una fotocamera digitale.

Innanzitutto, selezionare la risoluzione dell'immagine più alta possibile. Creare ad esempio immagini da 2.048 x 1.536 in cui siano anche supportate immagini da 1.042 x 768 e da 640 x 480. Successivamente, passare alla modalità "macro" per riprendere primi piani, che è sempre consigliabile quando si fotografano documenti. (Questa modalità è stata studiata specificatamente per scattare fotografie di fiori, insetti ecc.) Non utilizzando questa modalità, si ottengono immagini non nette e illeggibili.

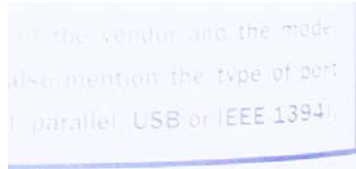




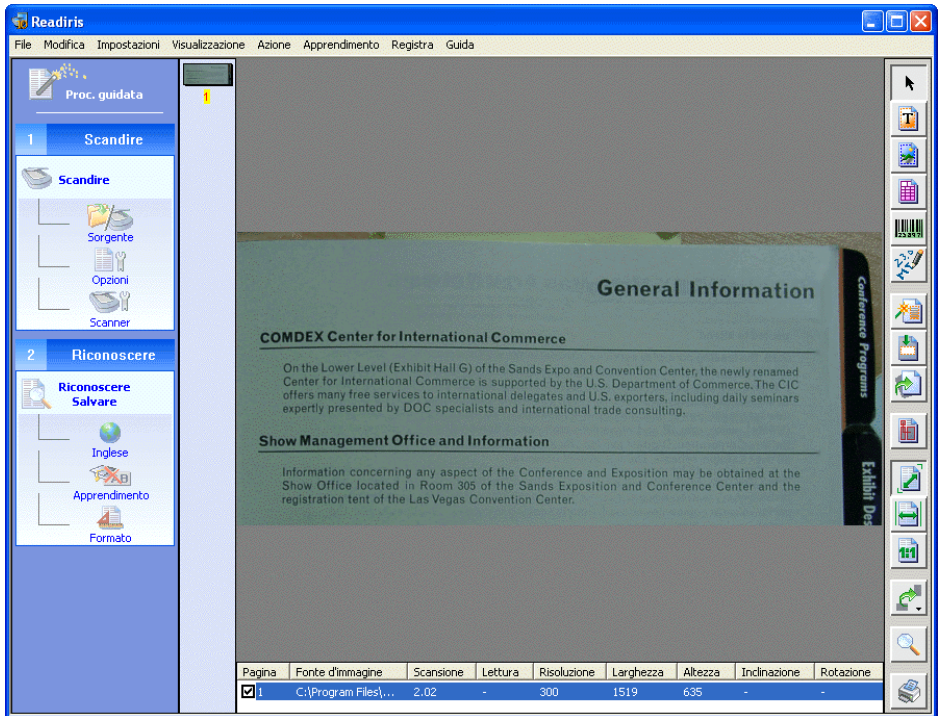
Attenzione alla compressione, che dovrà essere limitata o nulla. Una compressione troppo alta riduce la nitidezza del testo acquisito. Ingrandire manualmente per acquisire il documento; alcune fotocamere sono dotate di software per il rilevamento di immagini assemblate (“photo stitching”), che però non è utile per la lettura di documenti.

Tenere la fotocamera direttamente sul documento per evitare scatti angolati. Tuttavia, evitare che vi siano ombre sul documento determinate dalla propria mano o dall’apparecchio stesso. Scattare immagini stabili. Se necessario, montare un treppiedi.

Non fare uso del flash per scattare immagini su carta lucida, evitando così i riflessi. In generale, adattare la luminosità e il contrasto in base alla situazione: luce del giorno, luce da lampada, neon ecc. (Alcune fotocamere possono essere regolate per scattare foto di documenti bianchi.)



Per fare una prova, aprire l’immagine DIGITAL.JPG nella cartella Readiris e applicare la procedura illustrata.



SALVATAGGIO DELLE IMPOSTAZIONI PREDEFINITE

Impostare correttamente tutti i parametri di scansione e fare clic sul comando "Salva Impostazioni Standard" del menu "File" per salvare le impostazioni correnti come impostazioni predefinite per impieghi successivi.

Salva impostazioni standard

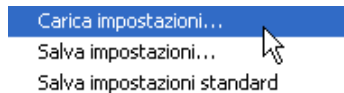




I file di impostazione non solo contengono le **impostazioni** dello scanner: consentono di determinare anche l'eventuale impiego dell'apprendimento interattivo, la lingua, la modalità di output (ad esempio il testo inviato a WordPad) ecc. In breve, *tutte* le impostazioni di funzionamento di Readiris vengono memorizzate nei file di impostazione.

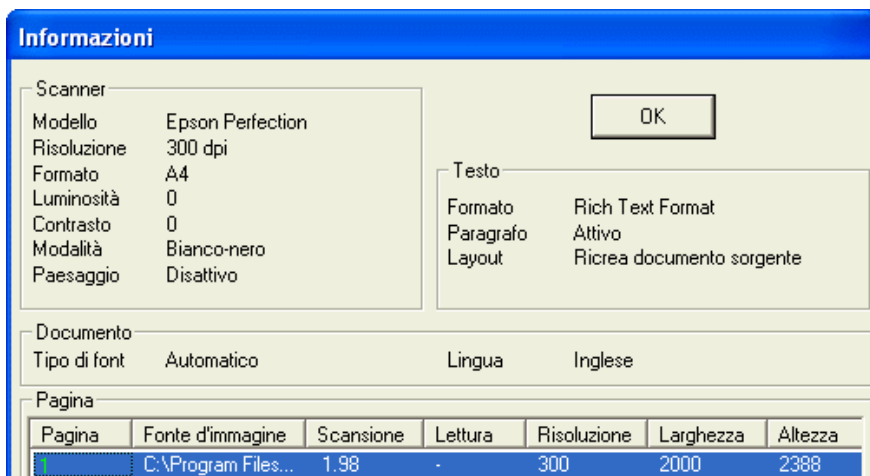
SALVATAGGIO DI IMPOSTAZIONI SPECIFICHE

Ad ogni avvio si utilizzeranno ovviamente le impostazioni predefinite, ma è possibile anche salvare impostazioni specifiche per non dover ridefinire i parametri di funzionamento. I comandi "Salva Impostazioni" e "Carica Impostazioni" del menu "File" svolgono tale funzione.



Esempio: se è spesso necessario eseguire il riconoscimento di documenti in lingua italiana che presentano un layout specifico, si consiglia di creare un file di impostazioni per questo tipo di documento. Selezionare quindi "Italiano" come lingua del documento, caricare un modello di divisione in zone specifico per non dover rieseguire ogni volta la stessa divisione dello schermo, disabilitare l'apprendimento ma attivare un dizionario dei font nella modalità "leggi", in quanto si utilizzano sistematicamente gli stessi tipi di caratteri ecc.

In caso di dubbi sulle impostazioni correnti, per scoprire quali esse siano non è necessario consultare in ogni menu e comando. Per una panoramica, è possibile utilizzare il comando "Informazioni" del menu "File".



Questo comando visualizza inoltre le informazioni disponibili nel pannello dei documenti per tutte le pagine.

SCANSIONE DI DOCUMENTI

Una volta impostato lo scanner è possibile cominciare ad acquisire i documenti. Ecco alcuni elementi da ricordare.

In primo luogo, prestare attenzione al disallineamento. Sebbene l'analisi ed il riconoscimento tollerino un certo disallineamento, può risultare difficile analizzare correttamente e riconoscere una pagina in presenza di un eccessivo disallineamento. Un disallineamento limitato (inferiore a 0,5°) può essere ignorato in quanto non compromette la precisione dell'OCR.

L'opzione "Allineamento Pagine" del pulsante "Opzioni" e del menu "Impostazioni" definisce se le pagine acquisite in modo disallineato verranno **raddrizzate** automaticamente; un disallineamento limitato viene ignorato. Questa opzione è disattivata per impostazione predefinita.

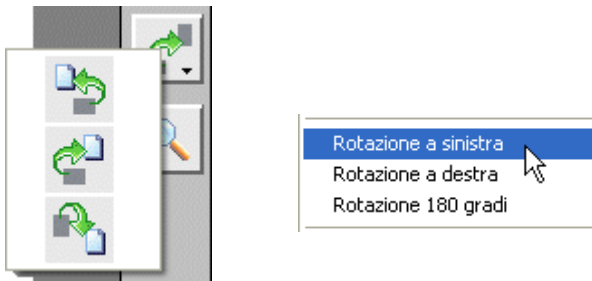


Se si dimentica di attivare questa opzione, utilizzare il tasto "Allinea Pagina" sulla barra degli strumenti immagine o del menu "Azione" per "raddrizzare" le pagine acquisite in modo distorto.



Il raddrizzamento richiede alcuni secondi: l'immagine viene analizzata per rilevare l'angolo di disallineamento e, se necessario, l'immagine a colori o in scala di grigi e la sua versione in bianco e nero vengono raddrizzate e l'analisi della pagina viene ripetuta.

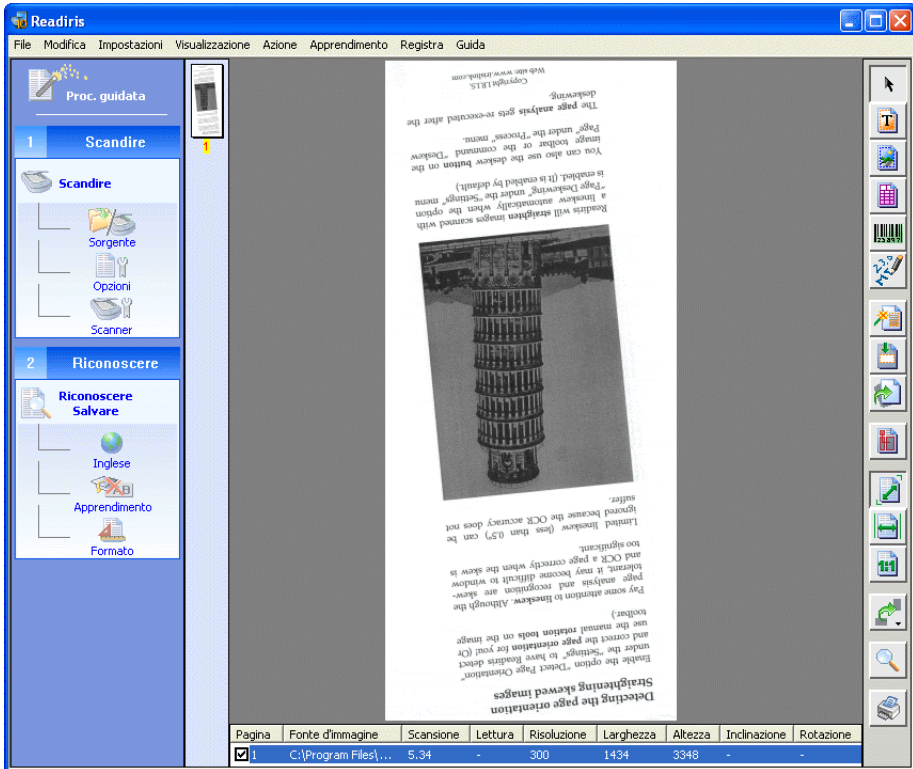
Potrebbe essere necessario regolare l'orientamento della pagina. A tale scopo utilizzare gli strumenti di **rotazione** della barra degli strumenti immagine. (I comandi corrispondenti si trovano nel menu "Visualizzazione".) Sono disponibili tre sensi di rotazione: verso sinistra, verso destra e capovolta. La rotazione richiede alcuni secondi poichè è l'immagine stessa ad essere aggiornata e non solo la visualizzazione sullo schermo.



Tuttavia, Readiris può correggere le pagine orientate male. Attivare l'opzione "Rivela l'Orientamento delle Pagine" del pulsante "Opzioni" (o sotto il menu "Impostazioni") e Readiris correggerà l'orientamento della pagina ove necessario.



A scelta, è possibile utilizzare l'immagine DESKEW.JPG della cartella Readiris. Attivare le opzioni "Allineamento Pagine" e "Rivela l'Orientazione delle Pagine" prima di aprire l'immagine e lasciare che Readiris raddrizzi la Torre di Pisa... come desiderato.

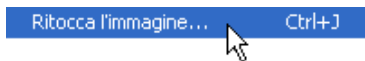


Si sottolinea che nel pannello dei documenti sono indicati l'angolo di disallineamento corretto e la rotazione eseguita!

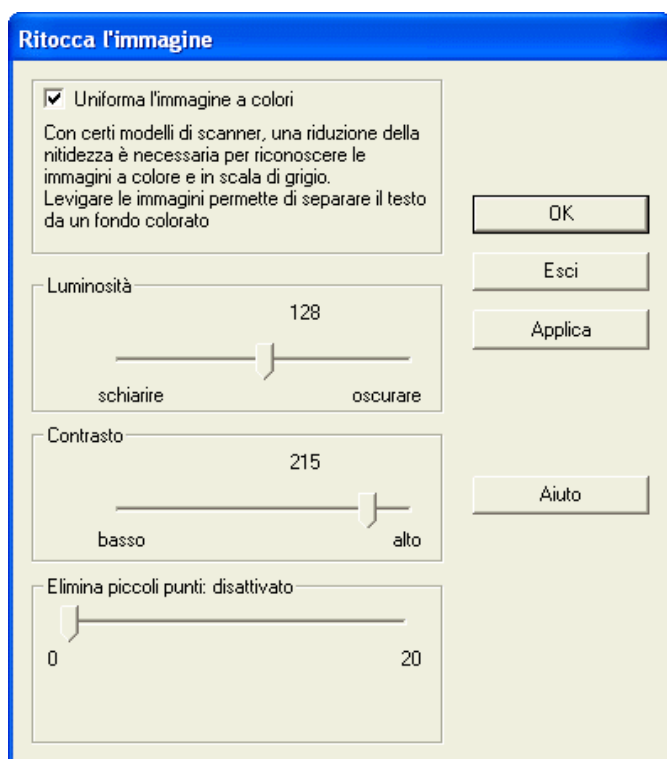
Larghezza	Altezza	Inclinazione	Rotazione
1776	3506	5.69°	ruotata a 180 gradi

REGOLAZIONE DELL'IMMAGINE ACQUISITA

Come già indicato, una serie di routine efficaci e intelligenti converte automaticamente le immagini a colori e in scala di grigi in immagini in bianco e nero. All'occorrenza l'utente può ottimizzare ulteriormente l'immagine per il successivo processo OCR. A tale scopo selezionare il comando "Ritocca l'Immagine" nel menu "Azione".



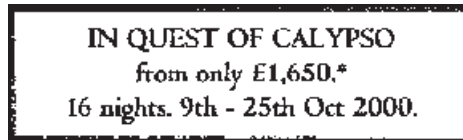
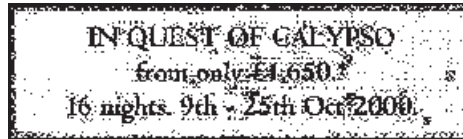
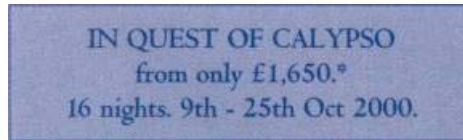
Quando si accede a questo comando, la versione in bianco e nero viene automaticamente visualizzata, come se fosse stata disabilitata l'opzione "Visualizza Documento a Colori". A questo punto, è necessario spiegare in dettaglio alcuni complicati, ma importanti, concetti.



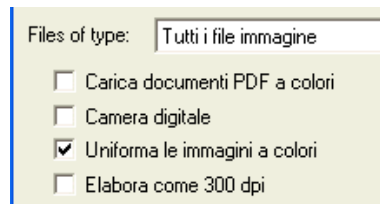
L'opzione "Uniforma l'Immagine a Colori" rende le immagini in scala di grigi e a colori più omogenee "appiattendole", uniformando le differenze di intensità relative. Come risultato si ottiene un contrasto più netto tra il primo piano - il testo - e lo sfondo - un colore, un'immagine ecc.

Questa funzione di "pre-elaborazione" può sembrare molto tecnica e di difficile comprensione ma sicuramente ha il suo scopo: con alcuni modelli di scanner, la riduzione dell'intensità è necessaria per riconoscere le immagini a colori e

in scala di grigi. L'uniformazione è talvolta il solo modo per separare il testo dallo sfondo colorato. Di seguito viene presentata un'immagine campione illeggibile senza l'uniformazione dell'immagine.



Questa opzione può essere attivata quando si memorizzano immagini preacquisite.



Passiamo ora alla **luminosità**. Questa impostazione determina la luminosità complessiva dell'immagine. La diminuzione o l'aumento della luminosità dell'immagine si applica a tutti i pixel. Ciò consente di eliminare lo sfondo della pagina. Di seguito vengono forniti due esempi. Nel primo esempio, tutte le zone dell'immagine sono scure. Pertanto, si dovrà schiarire l'immagine per eliminare lo sfondo della pagina. L'elemento in primo piano, ossia il testo, rimane



sufficientemente scuro da poter essere rilevato durante la binarizzazione. Nel secondo esempio, l'immagine è così chiara che neppure il testo in primo piano viene visualizzato nell'immagine binarizzata! Pertanto, è necessario scurire l'immagine per rendere leggibile il testo.

**Verenigde Staten,
een antwoord te vi
maar met name or**

*wyjścia każdego
brawia, że nasze*



*wyjścia każdego
brawia, że nasze*

**Verenigde Staten,
een antwoord te vi
maar met name or**

*wyjścia każdego
brawia, że nasze*

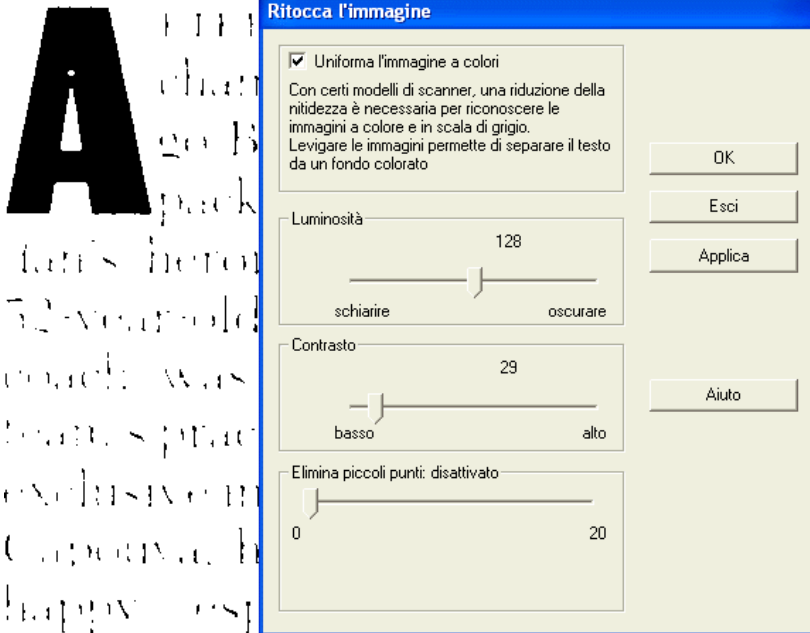
L'impostazione del **contrasto** definisce il contrasto locale tra le zone più scure e quelle più chiare dell'immagine. (Normalmente il testo è più scuro dello sfondo, mentre la situazione risulta opposta quando si lavora su testo invertito.) L'obiettivo è quello di far risaltare le forme dei caratteri in modo netto rispetto allo sfondo (colorato). Di seguito viene presentato un esempio nel quale è necessario intervenire poiché le impostazioni predefinite producono la frammentazione dei caratteri.

**A Look at International
Planning the Future**

**A Look at International
Planning the Future**

**A Look at International
Planning the Future**

Si noti che non viene attuata alcuna regolazione dell'immagine finché non si fa clic sul pulsante "Applica". Se si fa clic su "OK" viene eseguita la regolazione e si chiude la finestra. Di seguito viene fornito un esempio in cui l'immagine in bianco e nero è stata fortemente schiarita... senza tenere conto della precisione del processo OCR!



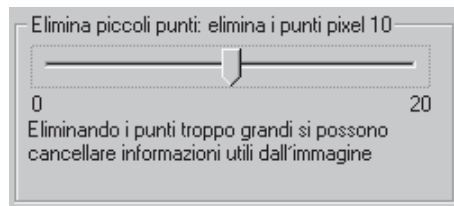
Queste opzioni riguardano le immagini a colori e in scala di grigi, mentre l'ultima opzione, "Elimina Piccoli Punti", riguarda esclusivamente le immagini in bianco e nero. L'eliminazione dei piccoli punti fa in modo che i pixel "parassiti" vengano rimossi dalle immagini in bianco e nero.



The Olympic Games

The Olympic Games

Assicurarsi di non cancellare “punti” troppo grandi, in quanto si rischierebbe di cancellare anche i puntini sulle “i”, porzioni di lettere a matrice di punto ecc.!

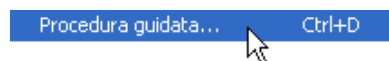
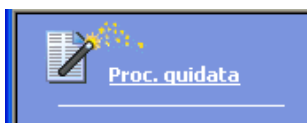


Il modo migliore per ottimizzare le immagini per il processo OCR è il seguente: posizionare la finestra di regolazione in modo tale da non impedire di valutare la regolazione eseguita sull'immagine. Adattare i parametri facendo clic ogni volta su "Applica" finché l'immagine risulterà nitida e chiara.

LA PROCEDURA GUIDATA, UN AIUTO INFALLIBILE!

A questo punto, è possibile iniziare la scansione dei documenti. Anziché scorrere tutti i parametri, utilizzare la **procedura guidata OCR**, un modo decisamente agevole per il riconoscimento delle pagine.

Fare clic sul comando "Procedura Guidata" nella barra degli strumenti principale o selezionare l'opzione corrispondente dal menu "Azione".



La procedura guida agevolmente l'utente lungo il rilevamento OCR: è sufficiente rispondere a qualche domanda per ottenere facilmente risultati rapidi con Readiris.



Si noti che la procedura guidata OCR viene lanciata a ogni avvio di Readiris; è possibile disattivarla disabilitando l'opzione "Inizia Procedura Guidata alla Partenza" nella prima schermata visualizzata (o con l'opzione equivalente del menu "Impostazioni").

READIRIS RICREA IL LAYOUT DEL DOCUMENTO

La procedura guidata OCR rende automatico il processo di riconoscimento, ma l'OCR "automatico" *non* deve essere confuso con l'autoformattazione. "Autoformattazione" significa che Readiris ricrea una **copia facsimile** del documento acquisito, applicando la formattazione di parola, paragrafo e pagina del documento originale.

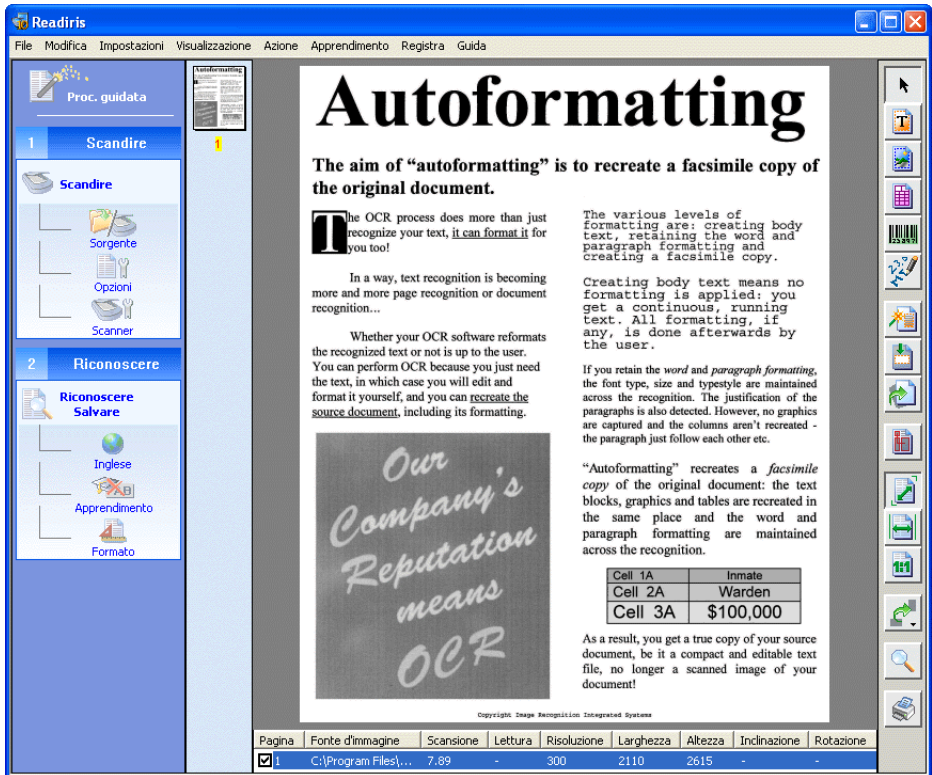


Durante il riconoscimento vengono utilizzati tipi di carattere (serif e sans serif, proporzionale e fisso, normale e condensato) simili a quelli del documento di partenza e vengono mantenuti le dimensioni in punti e gli stili dei caratteri (grassetto, corsivo, sottolineato, apice e pedici). Si ricreano le tabulazioni e l'allineamento (a sinistra, centrato, a destra e giustificato) di ciascun blocco di testo. Lo stesso avviene per gli elenchi puntati e numerati. Gli indirizzi e-mail e gli URL delle pagine web vengono rilevati e ricreati come collegamenti ipertestuali nel documento creato. Il posizionamento delle colonne, dei blocchi di testo e degli elementi grafici è conforme al documento originale.

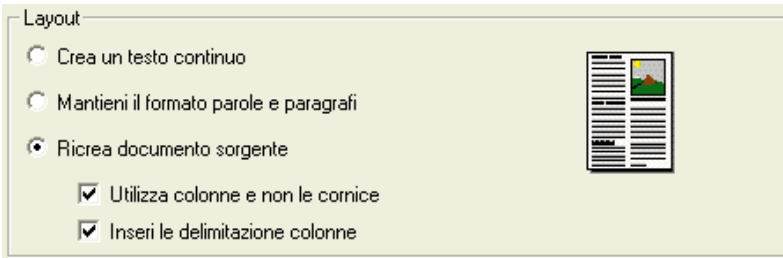
In altre parole, Readiris permette di archiviare una copia conforme di documenti, pur trattandosi ora di file di testo compatto ed editabile e non di un'immagine acquisita!

Pertanto, quando si utilizza "l'autoformattazione", viene *parzialmente* applicato l'ordinamento di finestre: è possibile includere ed escludere zone, ma qualsiasi riordinamento delle zone viene semplicemente ignorato...

Segue un esempio. Per comprendere meglio questa caratteristica, aprire l'immagine AUTOFORMAT.JPG nella cartella Readiris.

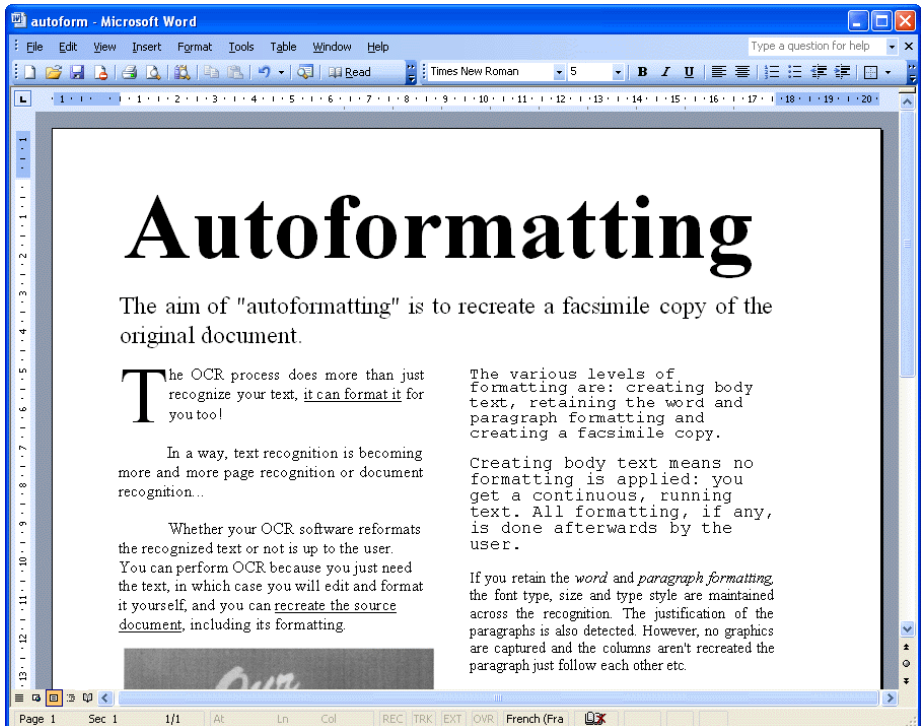


Fare clic sul pulsante "Formato" nella barra degli strumenti principale e inviare il risultato OCR a Microsoft Word o selezionare il formato RTF ("Rich Text Format") o Word (*.doc). Selezionare "Ricrea Documento Sorgente" come opzione di layout. (Per impostazione predefinita, viene attivata l'opzione "Aggiungi Linee nei Paragrafi" per attivare il ritorno a capo all'interno dei paragrafi.)

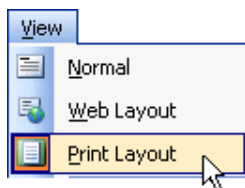


La ricostruzione del layout è disponibile a seconda della modalità di output selezionata. Alcuni formati “semplici” che producono testi “normali” quali Testo (ANSI), Testo formato MS-DOS (ASCII) ecc., *non* supportano codici di formattazione avanzati e quindi non sono in grado di offrire l’autoformattazione. Il formato PDF Adobe Acrobat è stato studiato per copiare l’aspetto dei documenti: i documenti PDF implicano per loro natura l’autoformattazione!

Quando il testo riconosciuto viene aperto utilizzando un elaboratore di testi, il testo appare in questo modo senza *alcun* intervento da parte dell’utente.



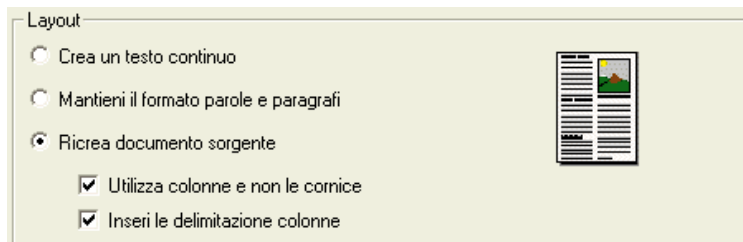
Per visualizzare correttamente l'effetto, è necessario abilitare la modalità "WYSIWIG" dell'elaboratore di testi, generalmente chiamata modalità "layout di pagina". Tuttavia, inviando il documento riconosciuto direttamente a Microsoft Word, il layout di pagina o il layout di stampa viene attivato automaticamente.



In breve, Readiris non solo riconosce i testi ma è anche in grado di formattarli per l'utente. L'OCR non si limita al riconoscimento di testi, ma è diventato uno strumento anche per il **riconoscimento di documenti!**

COLONNE, NON CORNICE

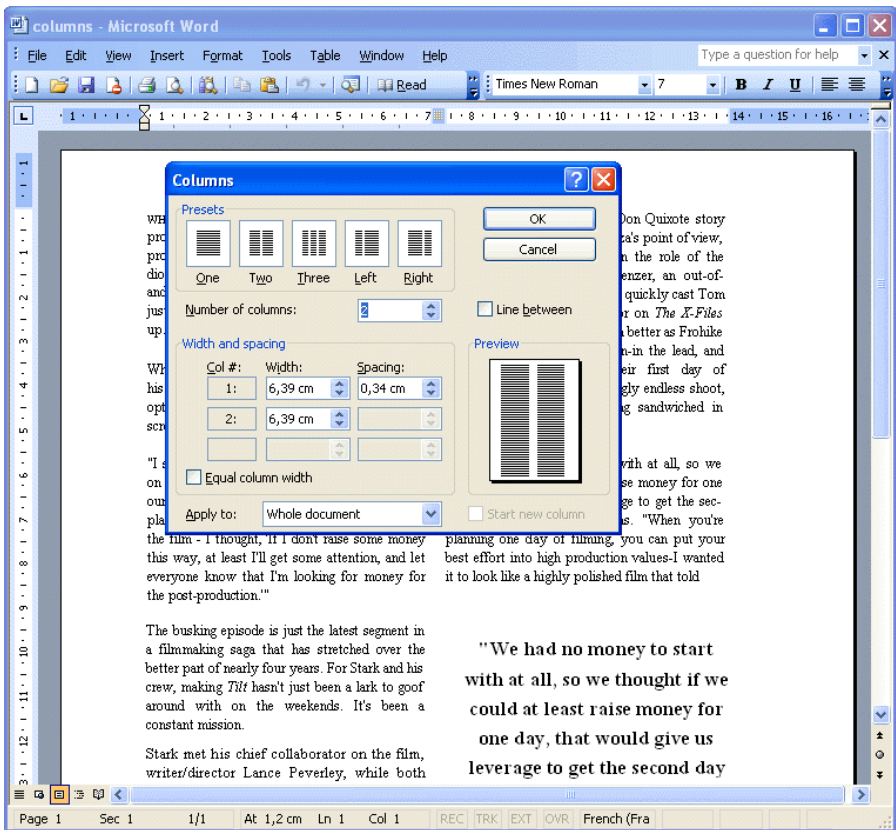
L'opzione di formattazione "Utilizza Colonne e Non le Cornice" stabilisce *come* viene eseguita l'autoformattazione: i blocchi di testo, le tabelle e gli elementi grafici possono essere memorizzati in riquadri o in **colonne** modificabili.



I riquadri ("frames") sono contenitori separati di testo utilizzati per posizionare diversi blocchi di testo, immagini e tabelle su una pagina. Con le colonne, il testo fluisce naturalmente da una colonna a quella successiva. Inoltre, è più facile modificare i testi incolonnati.

Si supponga vi siano colonne effettive nel documento acquisito: se il sistema non è in grado di individuare colonne nel documento di origine, questa modalità di formattazione utilizza i cornice come posizione "alternativa".

Per fare una prova, è possibile utilizzare l'immagine COLUMNS.TIF nella cartella Readiris.



L'opzione "Inserisci l'Interruzione di Colonna" rifinisce la ricreazione di colonne: stabilisce se inserire o meno interruzioni di colonna "fisse" alla fine di ogni colonna.



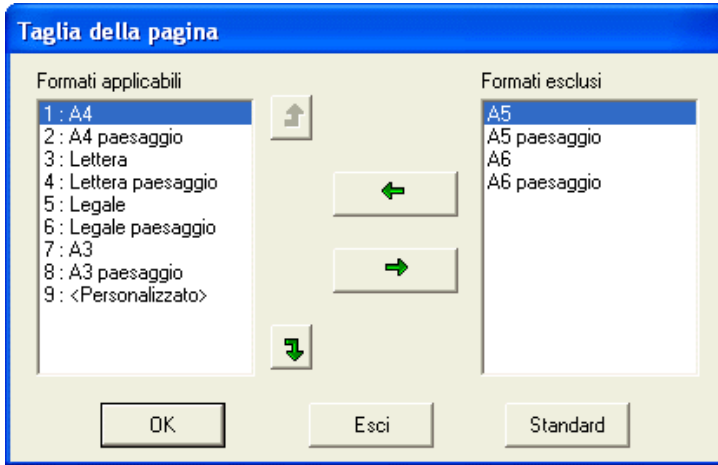
Con le interruzioni di colonna, il testo che da modificare, aggiungere o spostare rimane all'interno della colonna. Il testo non si sposta automaticamente con un'interruzione di colonna. Tutto il testo dopo un'interruzione di colonna viene spostato nella parte superiore della colonna successiva!

Attivare questa opzione se si desidera mantenere le interruzioni di colonna dove queste sono state rilevate nel documento riconosciuto, indipendentemente dalle modifiche del testo eseguita dopo l'OCR. Nei quotidiani e nelle riviste, le diverse colonne di una pagina corrispondono spesso a diverse "porzioni" di un articolo. Se il testo si spostasse da una colonna a quella successiva, se ne perderebbe il controllo.

Disattivare questa opzione quando si dispone di un testo in colonna: in questo modo si garantirà lo spostamento naturale del testo da una colonna a quella successiva.

Esiste un elemento che può convincere l'utente a scegliere di *non* ricreare il documento di origine: il formato pagina dei documenti elaborati. Cosa si intende? Gli esempi seguenti saranno utili alla comprensione. Si sceglie di eseguire la scansione di pagine Lettera ma si crea un output A4 perché è un formato più facile da stampare, mentre l'output Lettera richiede l'alimentazione manuale. Un avvocato, ad esempio, preferisce acquisire documenti in formato Lettera ma di salvarli in formato Legale.

Per questa ragione Readiris consente di indicare il formato carta desiderato per i documenti elaborati. Fare clic sul pulsante "Taglia della Pagina" nella finestra di dialogo "Formato".



Selezionare i formati carta da usare e quelli da escludere. È possibile utilizzare i formati carta desiderati per formattare i documenti riconosciuti, escludendo i formati indesiderati. Per ordinare i formati carta preferiti, Readiris scorre tutti i formati di carta nell'ordine indicato e utilizza il primo formato che sia in grado di contenere il documento acquisito. Il pulsante "Standard" consente di riapplicare le impostazioni predefinite. (Vengono utilizzate le impostazioni di Windows!)

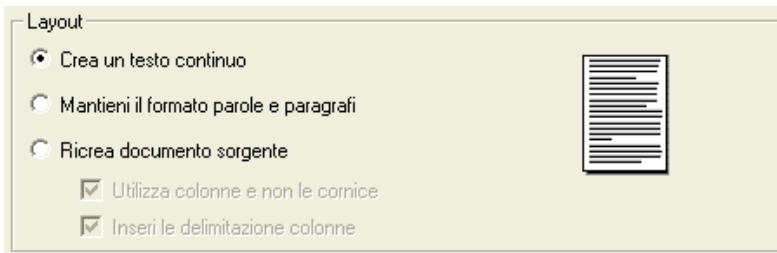
Questa opzione non è valida per i file HTML, un formato di testo concepito per l'uso in Internet che non presenta formati di pagina. Non è valida neppure per i file PDF, nei quali il documento di origine viene ricreato con precisione attraverso un processo di adattamento.

FORMATTAZIONE DI TESTO, PARTE 2

Le altre opzioni di layout sono "Crea un Testo Continuo" e "Mantieni il Formato Parole e Paragrafi".



Come illustra l'icona a destra, creare un **testo continuo** significa produrre un testo non formattato. Il testo verrà acquisito, ma ignorando completamente la sua formattazione. Utilizzare questa opzione quando si desidera riacquisire solo un testo, non il suo layout.



Il testo continuo è quello che si ottiene quando si esegue rapidamente il riconoscimento di una zona di testo facendo clic con il pulsante destro del mouse e selezionando il comando "Copia come Testo". Una volta eseguito il riconoscimento, si incollerà il testo continuo nell'applicazione di elaborazione testi.

L'opzione "Mantieni il Formato Parole e Paragrafi" rappresenta una via intermedia: durante il riconoscimento vengono mantenute la **formattazione della parola**, come tipo di font, dimensione in punti e stile del carattere, e la **formattazione del paragrafo**, come tabulazioni e allineamento.

Non confondere questa opzione di formattazione con l'autoformattazione "completa": questa opzione pone solo un paragrafo di seguito all'altro, non ricrea colonne e non copia la relativa posizione di varie zone di testo.

ESPORTAZIONE DI TESTO RIPETUTA

È possibile esportare i risultati dell'OCR diverse volte senza dover ripetere il riconoscimento. Modificare il formato di testo e le opzioni di formattazione facendo clic sul pulsante "Formato" e ancora clic sul pulsante "Riconoscere-Salvare". Non viene eseguito l'OCR a meno che non siano state definite nuove finestre o

modificate quelle esistenti, altrimenti Readiris riformatta semplicemente i risultati dell'OCR salvandoli nel nuovo formato di testo o inviandoli all'applicazione scelta.

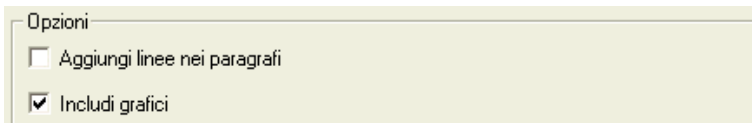


Lo stesso vale per qualsiasi altro elemento che modificate: quando aggiungete una pagina al lavoro OCR, solo quella pagina sarà riconosciuta. Quando create una nuova zona testo su una pagina, solo quella zona sarà riconosciuta prima dell'esportazione dei risultati.

Potete ad esempio riconoscere un documento di 10 pagine e salvarlo in un file Word, scandire il compendio ed inviarlo per e-mail ad un collega impaziente, scandire l'appendice - una tabella - e salvare tutti i risultati in un file HTML da mettere nel sito web della società.

SALVATAGGIO SEPARATO DEGLI ELEMENTI GRAFICI

In questo esempio, l'elemento grafico è stato inserito nel documento riconosciuto; questo dipende dall'opzione di formattazione "Includi Grafici". Il salvataggio degli elementi grafici all'interno del testo dipende ancora una volta dalla modalità di output. I formati di testo "semplici" come Testo (ANSI) ecc. non memorizzano elementi grafici!



Con Readiris è possibile salvare gli elementi grafici senza eseguire il riconoscimento del testo. Dal momento che Readiris crea immagini in bianco e nero, in scala di grigi e a colori, è possibile acquisire senza problemi immagini "lineart" e fotografie.

In che modo? Tracciare una zona grafica intorno alle illustrazioni, ai disegni ecc. La creazione manuale di finestre grafiche avviene nello stesso modo del



disegno di finestre di testo e tabelle: è sufficiente selezionare lo strumento "Finestra Grafica" sulla barra degli strumenti immagine.

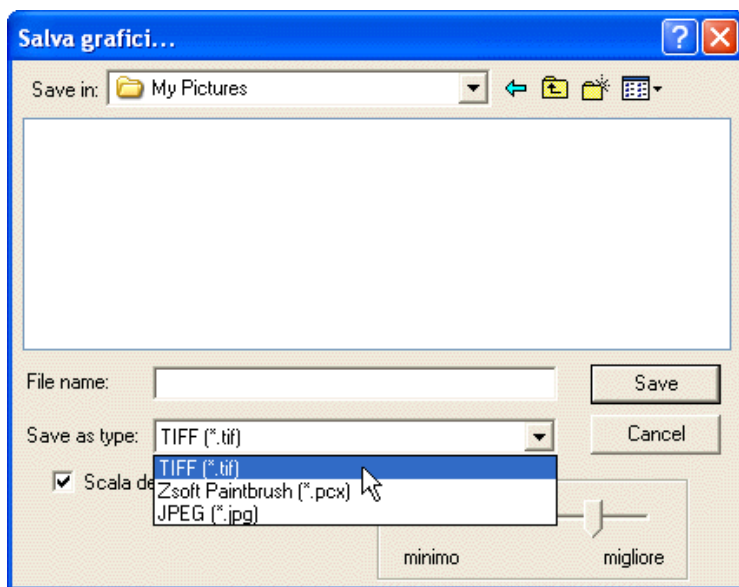


Scegliere quindi il comando "Salva Grafici" nel menu "File".

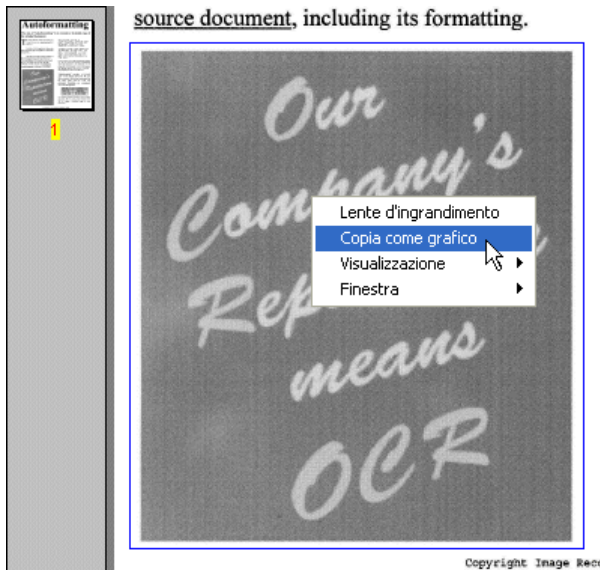


Il sistema chiederà di specificare un nome per il file. Scegliere il formato grafico che si intende utilizzare. Selezionare un formato supportato dal software di fotoritocco. Sono disponibili i formati JPEG, TIFF e ZSoft Paintbrush (*.pcx). Readiris Corporate supporta inoltre il formato compatto JPEG 2000 (*.j2c)!

Abilitare l'opzione "Scala di Grigi/Colori" per salvare l'elemento grafico come immagine a colori o in scala di grigi. Per salvare elementi grafici in bianco e nero in formato TIFF viene utilizzata la compressione Gruppo 4. Per salvare elementi grafici in scala di grigi e a colori in formato TIFF, viene utilizzata la compressione JPEG ed è possibile definire la qualità JPEG.



Per inviare un elemento grafico agli appunti anziché salvare un file immagine, fare clic con il pulsante destro del mouse sulla finestra e selezionate il comando "Copia come Grafico": la zona grafica sotto il puntatore può ora essere incollata.



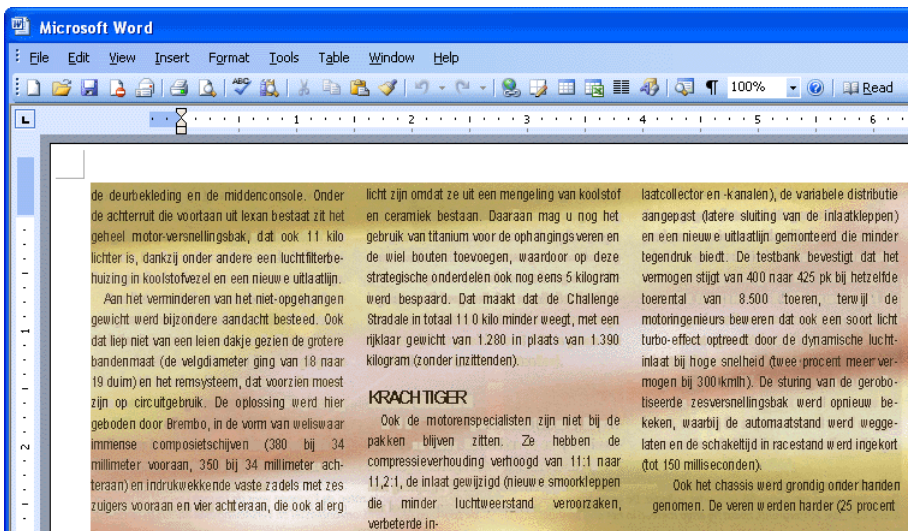
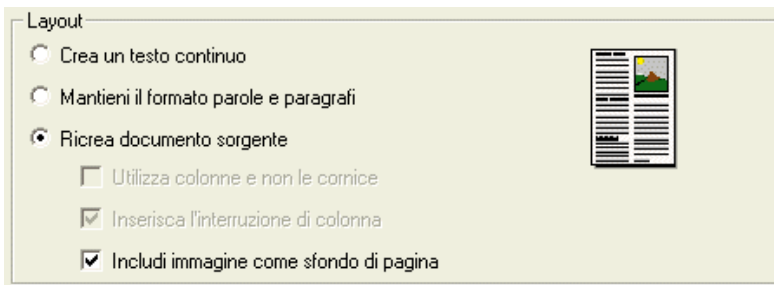
source document, including its formatting.

SALVATAGGIO DI SFONDI COLORATI

Readiris Corporate consente di salvare lo sfondo all'interno dei documenti riconosciuti. L'opzione "Includi Immagine come Sfondo di Pagina" posiziona l'immagine acquisita come sfondo di pagina sotto il testo riconosciuto; tale funzione è supportata per i formati HTML, RTF, Word (*.doc) e WordML (*.xml). (Quando si generano file PDF, è possibile creare lo stesso effetto selezionando il formato "PDF Testo-Immagine".)

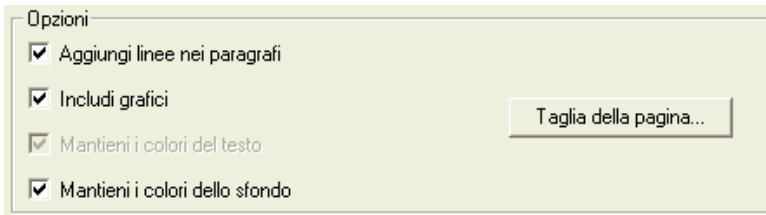
Naturalmente, questa opzione fa aumentare considerevolmente le dimensioni dei documenti riconosciuti poiché aggiunge allo sfondo l'immagine acquisita con tutti i particolari. *Tutti* i particolari? Non esattamente. Poiché vale anche per l'elaborazione PDF "testo-immagine", i pixel del testo riconosciuto vengono

cancellati per creare un documento leggibile. Ad esempio, se si visualizza il testo riconosciuto in nero sopra bitmap di caratteri neri si ottiene un testo con una pesante ombreggiatura. (L'immagine di esempio BACKGROUND.JPG ne mostra il funzionamento.)

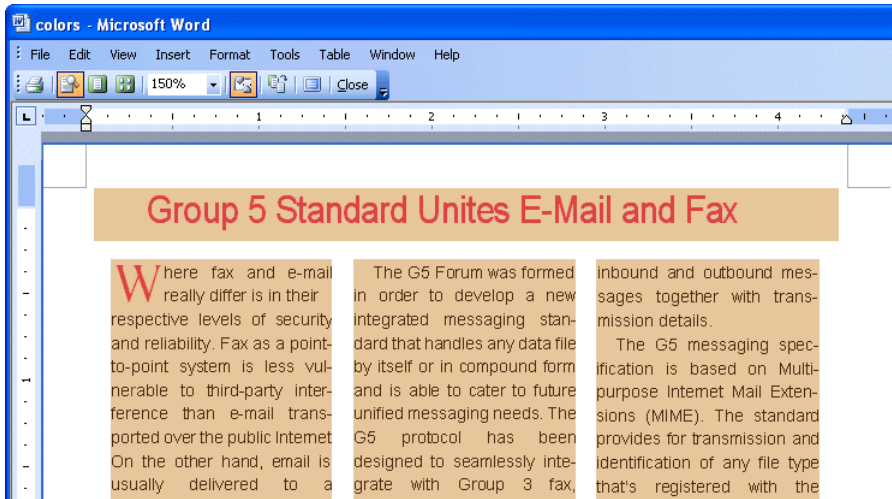




Readiris offre inoltre un'alternativa meno forte e più compatta: l'opzione "Mantieni i Colori dello Sfondo" consente di mantenere i colori campione della pagina durante il riconoscimento. (Ne consegue che i colori del testo vengono mantenuti.)



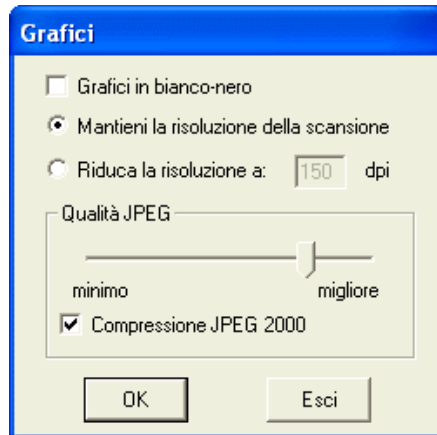
Si otterrà un colore di sfondo uniforme, se presente nel documento di origine, per i singoli paragrafi. Anche questa opzione è valida per i documenti HTML, Word e WordML. In questo caso, i dettagli di una fotografia a pagina intera *non* vengono mantenuti, mentre lo è il colore campione di una cornice di testo. (Per fare una prova, utilizzare l'immagine di esempio COLORS.JPG.)



INSERIMENTO COMPLETO DI ELEMENTI GRAFICI

Readiris Corporate offre altre opzioni avanzate per la gestione degli elementi grafici. Queste opzioni sono disponibili dal pulsante "Avanzato" del comando "Formato Testo".

Queste impostazioni si applicano anche a *tutti* gli elementi grafici: le zone grafiche comprese nei documenti riconosciuti e l'immagine di pagina collocata sopra il testo in un file PDF Adobe Acrobat "immagine-testo".



Stabilire la modalità colore: salvare gli elementi grafici come immagini a colori-scala di grigi o in bianco e nero. Selezionare la risoluzione dell'immagine nei documenti riconosciuti: mantenere la risoluzione di scansione o ridurla. (Non è possibile *umentare* la risoluzione in questo modo.) Quando si inviano i documenti riconosciuti a un sito web come file HTML, è possibile ridurne le dimensioni alla risoluzione dello schermo. Per impostazione predefinita, le immagini dei file HTML vengono ridotte a 72 dpi. (Tuttavia, è possibile indicare manualmente un altro valore.) Infine, è possibile selezionare la qualità JPEG. (Le immagini JPEG sono utilizzate per memorizzare grafici a colori e in scala di grigi in documenti PDF, Word, RTF ecc.)

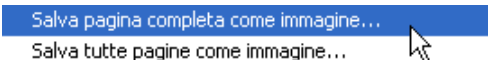
(L'opzione "Compressione JPEG 2000" è stata già discussa. La compressione JPEG 2000 viene applicata a tutti gli elementi grafici presenti nei documenti PDF.)

Queste opzioni consentono di influenzare la **dimensione del file** dei documenti di output. Un esempio: quando si generano file PDF "immagine-testo" con elementi grafici a due colori, si memorizzano le immagini in file TIFF compressi Gruppo 4. Memorizzare le stesse scansioni come immagini a colori e, per impostazione predefinita, si memorizzano file JPEG con un'elevata qualità (0,8).

LETTURA DI FAX E RICONOSCIMENTO DIFFERITO

Salvare immagini come file immagine offre un'altra possibilità: è possibile salvare una pagina *intera* e successivamente eseguire l'**OCR differito**, come si è fatto con le immagini preacquisite delle esercitazioni.

Eseguire la scansione del documento. Selezionare il comando "Salva Pagina Completa come Immagine" nel menu "File" per salvare una pagina singola. Si riceverà la richiesta di salvare l'intera pagina come file PDF, JPEG, JPEG 2000 (*.j2c), TIFF o ZSoft Paintbrush (*.pcx).



Salva pagina completa come immagine...
Salva tutte le pagine come immagine...

Selezionare il comando "Salva Tutte Pagine come Immagine" per salvare un documento a più pagine. In questo caso, sono disponibili due formati file: PDF e TIFF a più pagine.

A questo punto è possibile selezionare il disco come origine delle immagini ed aprire il file immagine con il pulsante "Aprire" (o con il comando corrispondente nel menu "Azione"). (Se si utilizza il comando "Apri" del menu "File", non è necessario aggiornare l'origine delle immagini.)

Poiché vengono supportate le immagini a colori, in scala di grigi e in bianco e nero alla stessa maniera, Readiris consente di aprire documenti PDF Adobe Acrobat, immagini fax DCX (una versione a più pagine del formato Paintbrush), immagini DjVu (*.djv, *.djvu), immagini JPEG, immagini JPEG 2000 (*.j2c, *.jp2), immagini PNG, immagini TIFF (non compresse, compresse di LZW, PackBits, Gruppo 3, Gruppo 4 e JPEG), immagini TIFF a più pagine, bitmap Windows (*.bmp) e immagini ZSoft Paintbrush (*.pcx).

Questa proprietà è particolarmente utile per convertire i **fax** in file di testo modificabili. Per leggere i fax con Readiris, il software individua la tipica risoluzione fax, ossia 100 x 200 dpi ("qualità normale"), 200 x 200 dpi ("qualità fine") e 200 x 400 dpi ("qualità superfine") e "rielabora" queste immagini automaticamente per garantire risultati OCR ottimali.



Ciononostante, è sempre una buona idea chiedere di inviare fax con la qualità “fine”. Questi fax daranno risultati OCR migliori.

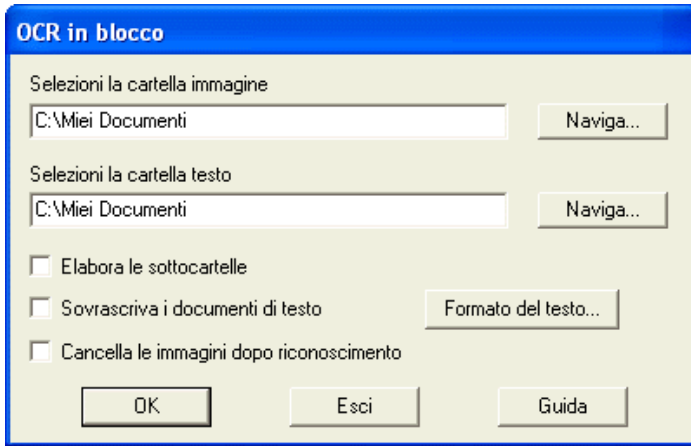
Per aprire immagini, non dimenticare che è possibile fare clic con il pulsante destro del mouse sulle immagini in Esplora Risorse di Windows e selezionare il comando "Riconoscimento" dal menu di scelta rapida per aprire le immagini. In alternativa, utilizzare la funzione di “trascinamento della selezione”: dopo aver trascinato file di immagini da Esplora Risorse di Windows alla zona immagine o sull'icona di Readiris, è possibile aprire direttamente i file.

RICONOSCIMENTO DI BLOCCHI DI DOCUMENTI

Readiris Corporate è molto più potente quando deve riconoscere immagini preacquisite: è possibile riconoscere automaticamente interi blocchi di documenti ed è possibile creare una cartella controllata. Eccone i dettagli.

L'**OCR in blocco** esegue il riconoscimento di tutte le immagini preacquisite in una cartella specifica. È possibile, ad esempio, eseguire la scansione di giorno e leggere i risultati di notte. Acquisire tutti i documenti da riconoscere. Una volta fatto, eseguire l'OCR in blocco e lasciar lavorare il computer. Il giorno successivo, tutti i documenti saranno stati riconosciuti.

I documenti riconosciuti hanno lo stesso nome dei file immagine. L'estensione del file dipende ovviamente dal formato di output selezionato: un file immagine 001.TIF viene convertito in 001.DOC se è stato selezionato Word come formato di output.



Selezionare la cartella immagine e la cartella di testo. La cartella di testo può essere diversa dalla cartella immagine, ma ciò non è obbligatorio. Quando la cartella immagine è identica alla cartella di output, i documenti testo e le scansioni saranno insieme, a meno non si attivi l'opzione "Cancella le Immagini dopo il Riconoscimento".

L'opzione "Elabora le Sottocartelle" stabilisce se saranno elaborate anche le sottocartelle della cartella immagine. Questo ha senso quando i documenti preacquisiti vengono collocati in sottocartelle specifiche. È possibile avere, ad esempio, una cartella denominata "25-2-2005" a indicare il giorno e le sottocartelle "1", "2" ecc. che contengono i documenti effettivi.

Se attivata, questa opzione avvia l'elaborazione di tutte le sottocartelle; non è possibile limitare l'OCR a una o più sottocartelle specifiche. Quando la cartella di testo è diversa dalla cartella immagine, vengono create delle sottocartelle, ossia ne viene creato un "mirror" nella cartella di uscita!

Fare clic su "OK" per avviare l'esecuzione. Non è necessario altro lavoro per convertire i documenti: il processo di riconoscimento è completamente automatico. (In questo caso, non si applica l'apprendimento interattivo.)



Inoltre, si è limitati ai formati di testo “esterni”. I risultati dell’OCR vengono salvati su disco. Inviare i risultati della lettura direttamente a un’applicazione scelta, inviarli per e-mail o aprirli automaticamente dopo il riconoscimento non ha senso se l’OCR viene eseguito su un computer “non utilizzato”. Tutte le altre opzioni OCR, come lingua e tipo di font del documento corrente ecc., rimangono valide: assicurarsi di attivare le opzioni corrette prima di avviare il riconoscimento.

L’OCR in blocco è dotato di alcune funzioni avanzate. L’opzione "Sovrascriva i Documenti di Testo" stabilisce se il processo OCR può sostituire i risultati del riconoscimento precedente. Disattivare questa opzione quando si aggiungono nuovi file immagine a una cartella contenente immagini precedentemente elaborate. (In caso contrario, i file immagine sarebbero nuovamente riconosciuti elaborando la cartella una seconda volta...)

L’opzione "Cancella le Immagini dopo il Riconoscimento" stabilisce se i file immagine saranno cancellati dopo il riconoscimento. In altri termini, Readiris può “ripulire” la cartella immagini in modo automatico!

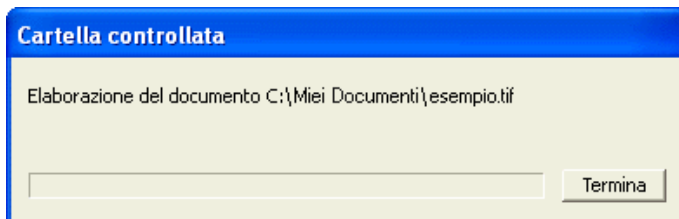
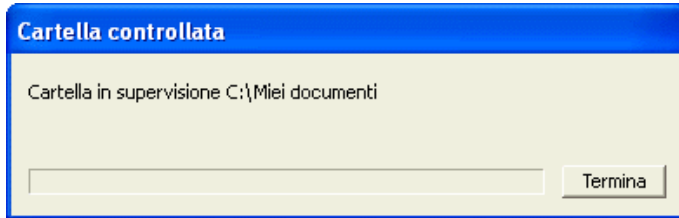
Attivare questa opzione quando si memorizzano-archiviano esclusivamente i documenti riconosciuti, eliminando i file immagine “temporanei”. Disattivare questa opzione quando si salvano sia le scansioni che i documenti riconosciuti. (Vale la pena ripetere che Readiris genera documenti PDF del tipo “immagine-testo”: questo formato di output salva le immagini acquisite e il documento riconosciuto in un unico file!)

CREAZIONE DI UNA CARTELLA CONTROLLATA

L’uso di una “cartella controllata” è molto simile alla procedura di OCR in blocco. La principale differenza è che, in questo caso, l’utente non avvia il riconoscimento. Al contrario, Readiris esegue sistematicamente il riconoscimento su qualsiasi file immagine che viene collocato in una cartella specifica. È possibile lasciar lavorare il software OCR giorno dopo giorno... È sufficiente acquisire nuovi documenti e questi verranno immediatamente riconosciuti.



Fare clic su "OK" per attivare la “supervisione” della **cartella controllata** e Readiris elabora le immagini progressivamente quando queste vengono create o collocate nella cartella. Fare clic su "Termina" per interrompere la supervisione.



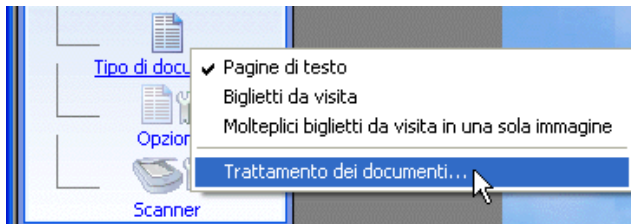
Come avviene con l'OCR in blocco, Readiris elabora le immagini di tutti i formati file supportati. Non è possibile limitare l'OCR a un formato grafico

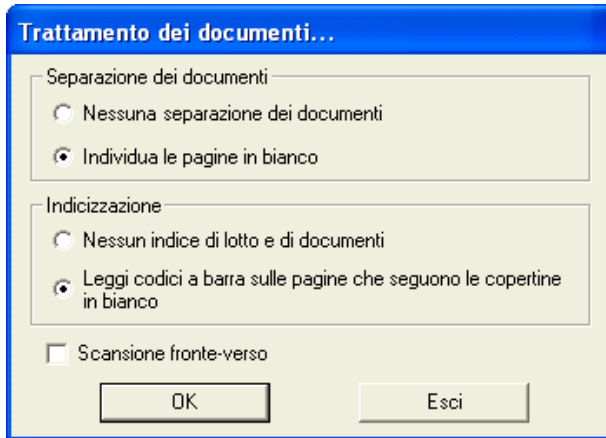


specifico: Readiris elabora le immagini di *tutti* i formati supportati. (I file con un formato file diverso vengono semplicemente ignorati.)

ORGANIZZAZIONE IN BLOCCHI

L'organizzazione in blocchi non significa che le pagine acquisite vengono elaborate direttamente senza consentire altre operazioni. Al contrario, Readiris Corporate prevede routine particolari che consentono di elaborare in modo intelligente i blocchi scanditi. Per comprenderne il funzionamento, selezionare il comando "Trattamento di Documenti" dal pulsante "Tipo di Documento" (o dal menu "Impostazioni").



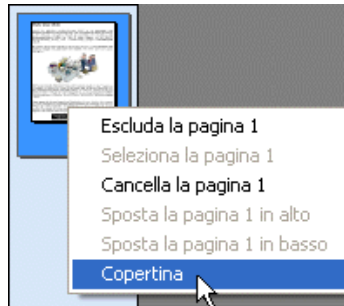


Per **separare i documenti** è possibile inserire pagine di copertina vuote tra le pagine. Per pagina “vuota” si intende una pagina in cui i pixel neri sono quasi del tutto assenti - fatta eccezione per i bordi. La barra strumenti della pagina evidenzia la copertina e il pannello dei documenti la contrassegna.



Pagina	Copertina	Codice a barra	Fonte d'immagine	Scansione
<input checked="" type="checkbox"/> 1	si		C:\Program Files...	2.48
<input checked="" type="checkbox"/> 2	-	TEST21	C:\Program Files...	2.38
<input checked="" type="checkbox"/> 3	-		C:\Program Files...	1.52

È inoltre possibile segmentare manualmente una serie di immagini acquisite. Selezionare una pagina, vuota o meno, nella barra strumenti della pagina, fare clic con il pulsante destro del mouse e selezionare il comando "Copertina" dal menu di scelta rapida. (Si noti che il contenuto della pagina di copertina viene perduto: le copertine servono da separatori, ma *non* vengono lette.)

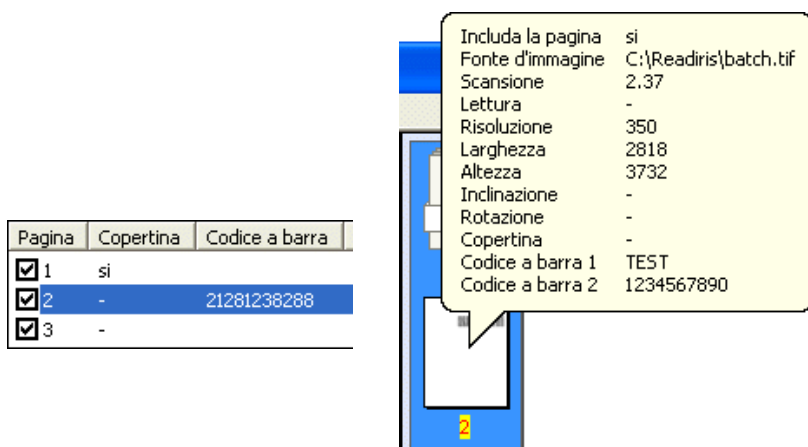


Si inizia separando in documenti diversi la serie di immagini acquisite. Successivamente, occorre farne un'accurata **indicizzazione**. A questo scopo, viene applicata la lettura dei codici a barre. Tutti i codici a barre presenti in una pagina che segue una copertina vuota vengono letti automaticamente. (Per fare una prova, utilizzare l'immagine di esempio BATCH.TIF.)



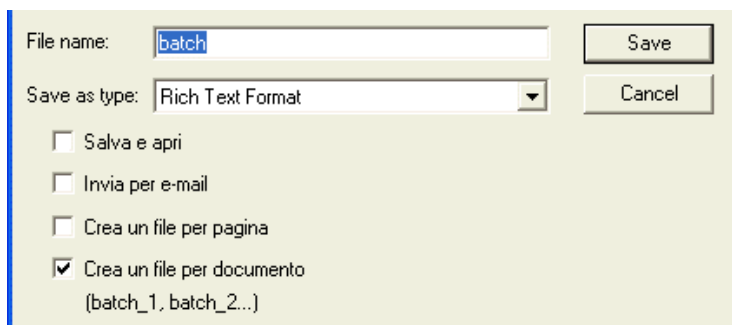
L'opzione "Scansione Fronte-Retro" indica che si sta eseguendo la scansione di entrambi i lati dei documenti. Ovviamente il retro viene ignorato per rilevare un codice a barre nella pagina successiva: tale pagina diventa l'immagine sul *fronte* successiva, non l'immagine successiva!

L'interfaccia di Readiris presenta altre funzioni utili per l'elaborazione di blocchi di documenti: il pannello dei documenti e le descrizioni delle icone della pagina mostrano il valore dei codici a barre riconosciuti.



I risultati della lettura dei codici a barre vengono salvati nell'indice XML, non nel documento riconosciuto.

Se si elaborano i blocchi in modo manuale, l'opzione "Crea Un File per Documento" della finestra di dialogo "Salva" consente di salvare documenti separati, non blocchi interi. (Ciò si verifica solo quando si crea file esterni, non quando si inviano i risultati a un'applicazione scelta.)





RICONOSCIMENTO DI TABELLE

Fino ad ora sono stati spiegati il riconoscimento di testi e di fax e il salvataggio degli elementi grafici. A questo punto è possibile illustrare l'elaborazione di una tabella. Prendere una tabella numerica ed eseguirne la scansione oppure aprire l'immagine campione TABLES.JPG nella cartella Readiris.

L'immagine TABLES.JPG contiene in realtà due tabelle per un motivo ben preciso. L'analisi della pagina divide in zone le due tabelle e Readiris è in grado di ricostruirle, ricreando le tabelle cella dopo cella nel foglio elettronico o inserendo un oggetto tabella all'interno del file dell'applicazione di elaborazione testi.

Verranno analizzate le diverse soluzioni iniziando dalla tabella "con griglia" o "incorniciata", che è caratterizzata da bordi intorno alle celle.

Reading Tables

Readiris recognizes tabular data and recreates them cell by cell in worksheets or as table objects inside wordprocessor files.

To insert tables as table objects, you must retain the word and paragraph formatting or recreate the source document; see the "Format" button on the main toolbar.

The page analysis detects "gridded" and "ungridded" tables. "Gridded" or "framed" tables have borders around the cells - as does the example below. The borders of the table cells get recreated.

Performance test optical media				
CD-ROM Digital Versatile Disk	Average access time (msec)	CPU utilization (%)	Video clip playbacks (frames dropped)	Sequential read 16 KB (K bps)
CD-ROM 4x speed	442	4.2	10 8	612
CD-ROM 12x speed	137	20.9	5 4	1,586
CD-ROM 24x speed	80	58.2	3 2	2,258
CD-ROM 32x speed	60	72.1	- -	2,987
DVD	58	78.9	- -	3,143

Tested on 333 MHz Pentium II PC with 64 MB RAM and 4 GB SCSI HD

"Ungridded" tables don't have any borders around the cells. When the columns of ungridded tables are too widely spaced, the page analysis may not detect a table window to avoid confusion with columnized text blocks.

When your tables exclusively contain **numeric characters**, enable the numeric reading mode with the "Language" button on the main toolbar for increased accuracy.

Pagina	Fonte d'immagine	Scansione	Letture	Risoluzione	Larghezza	Altezza	Inclinazione	Rotazione
<input checked="" type="checkbox"/> 1	C:\Program Files\...	7,22	-	300	2092	2517	-	-

Eeguire il riconoscimento dopo aver abilitato l'opzione di layout, "Mantieni il Formato Parole e Paragrafi" o "Ricrea Documento Sorgente" e le tabelle vengono ricreate. Aprire l'elaboratore di tsti per esaminare il risultato: le celle e i bordi sono stati ricreati da Readiris uno per uno. (È anche possibile che siano stati inclusi paragrafi di testo nel file di testo.)



table - Microsoft Word

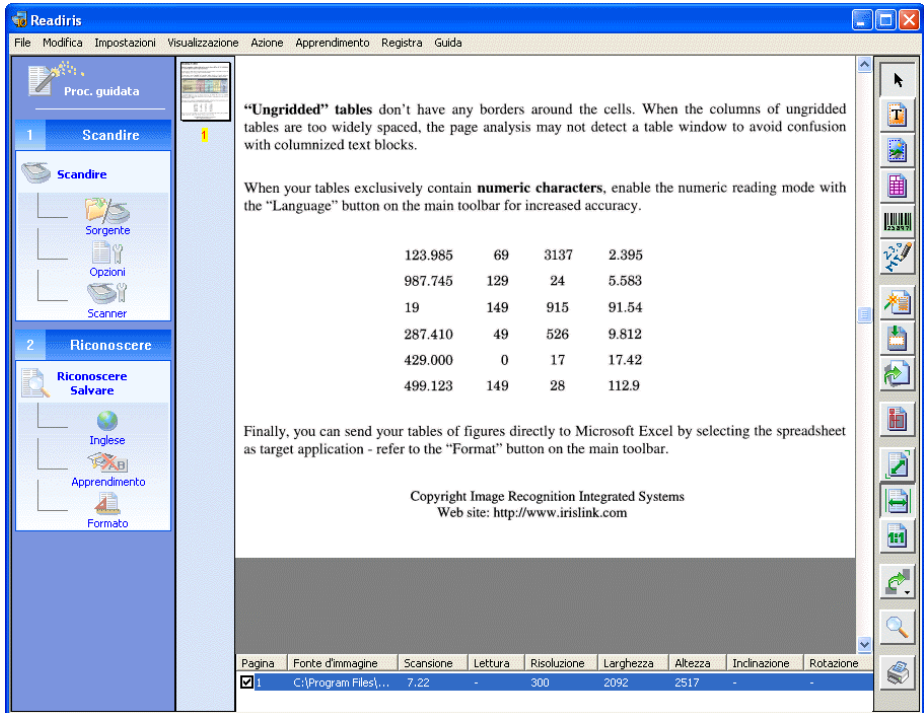
File Edit View Insert Format Tools Table Window Help

Type a question for help

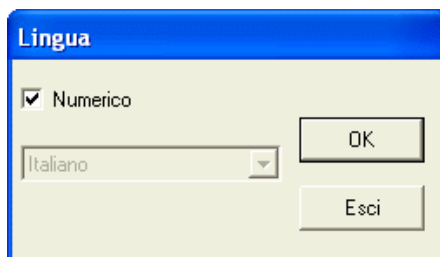
Times New Roman 12 B I U

Performance test optical media					
CD-ROM Digital Versatile Disk	Average access time (msec)	CPU utilization (%)	Video clip playbacks (frames dropped)		Sequential read 16 KB (K bps)
CD-ROM 4x speed	442	4.2	10	8	612
CD-ROM 12x speed	137	20.9	5	4	1,586
CD-ROM 24x speed	80	58.2	3	2	2,258
CD-ROM 32x speed	60	72.1	-	-	2,987
DVD	58	78.9	-	-	3,143
Tested on 333 MHz Pentium II PC with 64 MB RAM and 4 GB SCSI HD					

Questo è l'esempio "senza griglia". Non vi sono bordi intorno alle celle. L'analisi della pagina individua comunque la tabella!

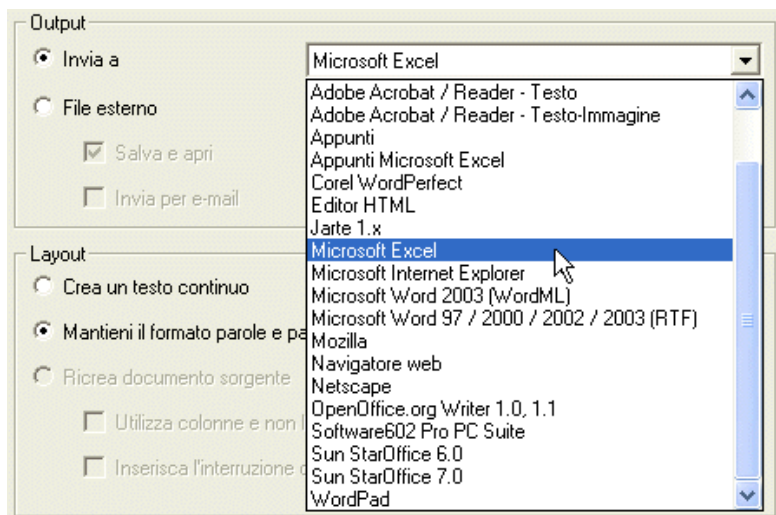


Per una precisione ottimale di OCR, limitare il riconoscimento ai **simboli numerici** mediante il pulsante "Lingua". (La modalità numerica non riguarda strettamente i numeri, ma include i simboli da “0” a “9”, “+”, “*”, “/”, “%”, “,” (virgola), “.” (punto), “(”, “)”, “-”, “=”, “\$”, “£”, “¥” e il simbolo “€”).



Poiché questo procedimento è possibile solo quando la tabella non contiene simboli alfabetici; in caso contrario non sarebbe possibile riconoscere correttamente le porzioni di testo. È possibile attivare la modalità numerica, ma non per la prima tabella.

In questo caso, si invierà il risultato OCR direttamente al foglio elettronico di Microsoft Excel. Selezionare quindi Excel come applicazione di destinazione utilizzando il tasto "Formato".



Il foglio elettronico viene avviato automaticamente con questo risultato: viene ricreata la tipica struttura a tabella con righe e colonne. È possibile ora passare all'elaborazione dei dati.



Microsoft Excel - Table

File Edit View Insert Format Tools Data Wind

Arial

G3

	A	B	C	D	E
1	123.985	69	3137	2.395	
2	987.745	129	24	5.583	
3	19	149	915	91.54	
4	287.410	49	526	9.812	
5	429.000	0	17	17.42	
6	499.123	149	28	112.9	
7					

È possibile che vi siano tabelle “senza griglia” non individuate dall’analisi della pagina come zona tabella perché le colonne sono troppo distanziate. Readiris cerca di evitare confusione con i blocchi di testo organizzati in colonne. Per creare manualmente una finestra di tabella, fare clic sullo strumento "Finestra di Tabella" nella barra strumenti dell’immagine e procedere come d’abitudine; la descrizione degli strumenti indica di nuovo il numero di finestre di tabelle.

Disegna una finestra di tabella: 0

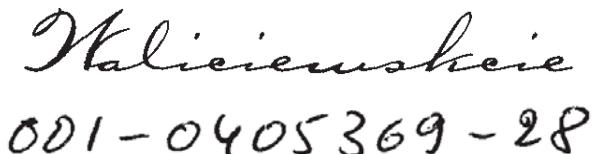


RICONOSCIMENTO DELLA SCRITTURA MANUALE

Si è visto come eseguire il riconoscimento di documenti acquisiti, tabelle, fax, istantanee scattate con macchine fotografiche digitali, nonché salvare gli elementi grafici e convertire documenti PDF. Readiris offre un’altra opzione di lettura: il riconoscimento di testi scritti a mano.

In realtà, si dovrebbe parlare di testi in stampatello, non di testi scritti a mano. La scrittura a mano (in minuscolo) è un tipo di scrittura manuale “corsiva” e continua. I simboli che compongono una parola o una stringa di caratteri si toccano

ed è impossibile definire dove finisce un simbolo e inizia il successivo. Con la scrittura manuale in *stampatello*, le lettere sono separate e per il software che effettua il riconoscimento è più facile individuare i singoli caratteri.



Kalicienuskeie
001-0405369-28

È necessario un software specializzato, come “ICR” (“Intelligent Character Recognition”), per il riconoscimento di simboli scritti in stampatello. La potente tecnologia ICR di I.R.I.S. si basa su oltre un milione di campioni di scrittura! Readiris supporta tutti gli stili naturali di scrittura, sia europei che americani. Non è necessario seguire uno stile particolare.

Il riconoscimento di testo in stampatello è limitato ai simboli numerici (0-9), alle maiuscole (A-Z) e ai simboli di punteggiatura “,” (virgola), “.” (punto) e “-” (trattino).

Nel caso i risultati ottenuti non siano ottimali, è possibile migliorarli utilizzando il modulo di scrittura I.R.I.S. e adattarlo al proprio stile di scrittura. Consultare la guida in linea di Readiris per conoscerne le regole. Una volta acquisita la padronanza di alcune regole base, si comprenderà cosa determina errori di riconoscimento e come evitarli. Il modulo di scrittura vuoto I.R.I.S. può essere utilizzato come una pagina “modello” per garantire che la distanza tra le lettere maiuscole sia mantenuta costante e che le lettere abbiano la stessa dimensione! Il modulo vuoto per la ristampa e la modifica è disponibile nel CD-ROM di Readiris.



Forma e regole per la scrittura manuale

Nel caso i risultati ottenuti non siano ottimali, è possibile migliorarli utilizzando il **modulo di scrittura I.R.I.S.** e **adattandolo** al vostro **stile di scrittura**.

- Il **modulo di scrittura** vuoto I.R.I.S. può essere utilizzato come una pagina "modello" per assicurare che la distanza tra le **lettere maiuscole** sia mantenuta costante e che le lettere abbiano la stessa dimensione!

Nota: il modulo di scrittura I.R.I.S. vuoto è disponibile nel **CD-ROM Readiris** e può essere **ristampato** (in formato Adobe Acrobat PDF) e **modificato** (in formato Word (*.doc)). Stampare il modulo a **colori** e utilizzarlo per la scrittura. In alternativa, stamparlo in **bianco e nero** e utilizzarlo come una "guida" sotto la carta sulla quale state scrivendo.

- Una volta acquisita la padronanza di alcune **regole** base, si comprenderà cosa determina errori di riconoscimento e come evitarli.

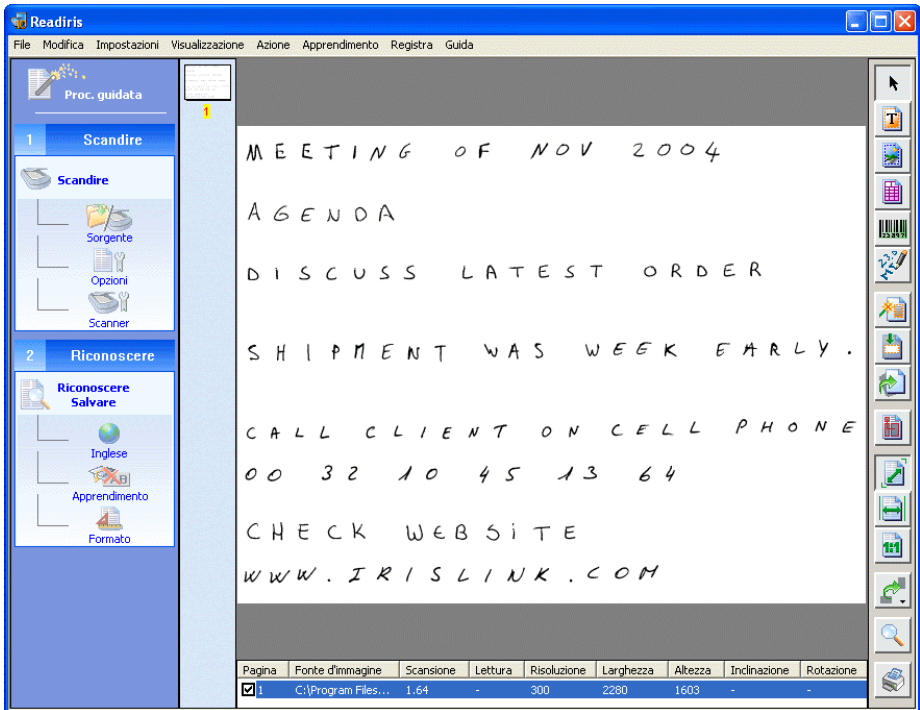
Lo scrittore è la penna
Utilizzare una **penna a sfera** con un tratto ben

Questa eccezionale funzione consentirà di leggere in modo automatico gli appunti presi durante una riunione. Il riconoscimento in Readiris è semplicissimo. Sarà sufficiente tracciare a mano una finestra intorno al testo da elaborare ed eseguire il riconoscimento. (Per fare una prova, utilizzare l'immagine HANDPRINTING.TIF.)

Disegna una finestra di codici a barra: 0



Le caratteristiche del documento (come lingua, tipo di font e spaziatura dei caratteri) non si applicano alla scrittura in stampatello. Il riconoscimento si limita a una lingua di base “standard”, ossia a un set di caratteri latini di lettere (maiuscole). Non risulterà attivo neppure l'apprendimento interattivo. L'apprendimento non ha senso in un ambiente nel quale ciascuno ha un proprio stile di scrittura. (Inoltre, come è stato già precisato, la potente tecnologia ICR si basa su oltre un milione di campioni di scrittura...)





LETTURA DELLE BARRE E DEGLI SPAZI

Readiris è tanto eccezionale da riuscire a leggere anche i codici a barre: i codici a barre presenti nelle immagini acquisite possono essere letti e inseriti nei documenti elaborati come informazioni riconosciute.

I codici a barre sono costituiti da barre parallele separate da spazi. Combinazioni predefinite di barre e spazi rappresentano caratteri specifici. Esistono diversi sistemi standardizzati (“simbologie”) di codici a barre. Sono supportate tutte le simbologie di codici a barre più diffuse.

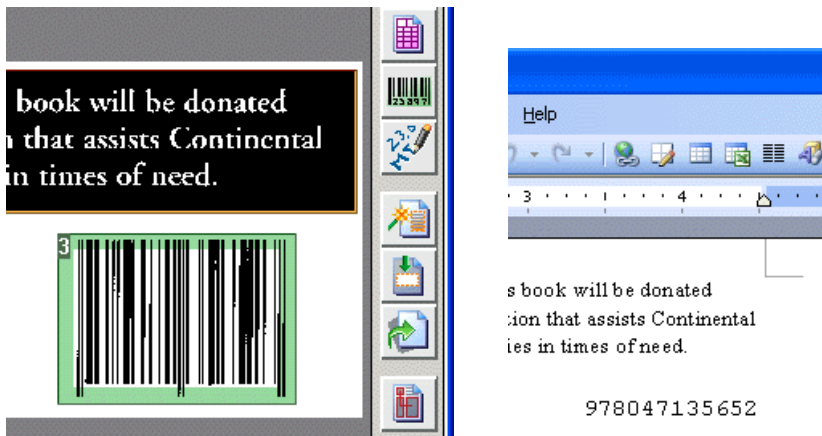


La qualità dei codici a barre è sufficiente solo se essi sono stampati con una stampante laser o una stampante a getto di inchiostro. Sono da escludere i codici a barre stampati con una stampante a matrice di punti: essi non hanno un contrasto sufficiente e la loro risoluzione non supera quasi mai i 60 dpi! Readiris riconosce in modo ottimale i codici a barre con un buon contrasto; le barre nere su sfondo bianco assicurano i risultati migliori. Per la maggior parte dei tipi di codici a barre è necessaria una “zona tranquilla” attorno al codice vero e proprio. I codici a barre non producono risultati parziali; un carattere iniziale o finale mancante o una cifra di controllo non corretta provocano sempre un errore di lettura, ossia un risultato nullo!



Disegnare una finestra di codici a barre attorno a ciascun codice, per fare in modo che l’analisi della pagina non li rilevi, ed eseguire il riconoscimento. I codici a barre vengono letti e inseriti nei documenti di testo generati. È anche possibile

fare clic con il pulsante destro del mouse su una zona di codici a barre e selezionare il comando "Copia come Dati" dal menu di scelta rapida; il codice a barre viene letto e inviato agli appunti. (I caratteri di controllo di alcune simbologie di codici a barre vengono verificati, ma poi eliminati dal risultato della lettura.) L'immagine di esempio BARCODE.TIF ne mostra il funzionamento.



RICONOSCIMENTO DEI BIGLIETTI DA VISITA

Si è visto come acquisire documenti, tabelle, fax, istantanee scattate con fotocamere digitali, nonché come salvare elementi grafici e convertire documenti PDF. Readiris Corporate offre un'ulteriore funzione di lettura: il riconoscimento dei biglietti da visita ("BCR" - "Business Card Reading").

Con Readiris è possibile acquisire biglietti da visita, riconoscerli e convertirli in un **database di indirizzi**. In questo contesto, **OCR** permette di codificare i biglietti da visita senza doverli ridigitare. Si pensi all'ultima fiera commerciale dalla quale si è tornati con una pila di biglietti da visita, che hanno richiesto alla segretaria due giorni di lavoro per essere classificati.



I dati dei biglietti vengono estratti automaticamente dall'immagine e quelli riconosciuti vengono assegnati ai campi specifici del database. Readiris utilizza un "knowledge database" che consente di acquisire l'intelligenza necessaria per distinguere nome e cognome, città e stato, numero di telefono e fax ecc. Ogni paese presenta uno "stile" diverso di composizione dei biglietti da visita; gli americani dispongono l'indirizzo in modo diverso dai francesi e così via.

Il sistema funziona con almeno 28 **paesi**: sono supportati i biglietti da visita dell'America del Nord e del Sud e i biglietti da visita dei paesi europei comprese le nazioni dell'Europa dell'Est. (A scelta, è possibile leggere i biglietti da visita asiatici di paesi come Cina, Giappone, Corea e Taiwan.)

I dati risultanti sono disponibili per l'**esportazione**. È possibile salvare e strutturare i propri indirizzi in un file di testo, ad esempio nel formato di dati delimitati da virgola e nel formato vCard, nonché importarli in qualsiasi database d'indirizzi come Microsoft Access.

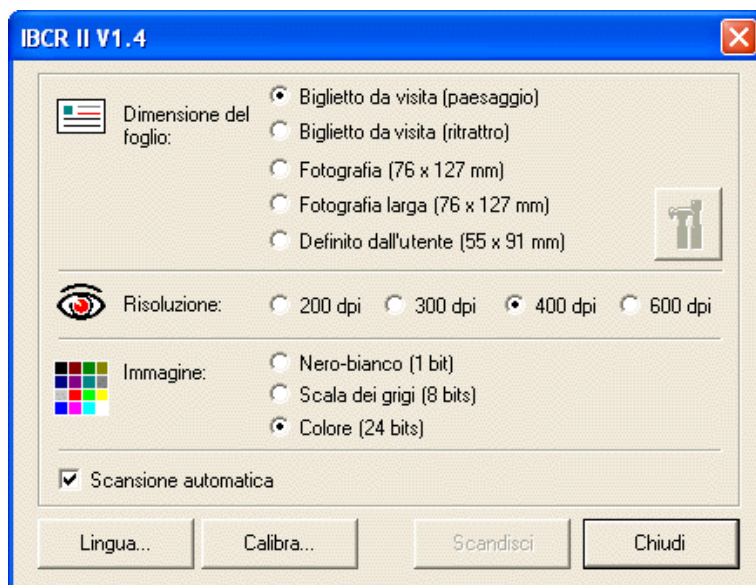
In alternativa, è possibile inviare gli indirizzi direttamente al software di gestione indirizzi di Microsoft Outlook (Express) e al software **PDA** Palm Desktop. La lettura dei biglietti da visita completa applicazioni come gestioni contatti, database o anche programmi di elaborazione di testi le cui funzioni di "stampa unione" consentono di stampare lettere, buste ed etichette.

SCANSIONE DEI BIGLIETTI DA VISITA

Come funziona la lettura dei biglietti da visita? Innanzitutto, lo scanner deve essere impostato in modo adeguato.

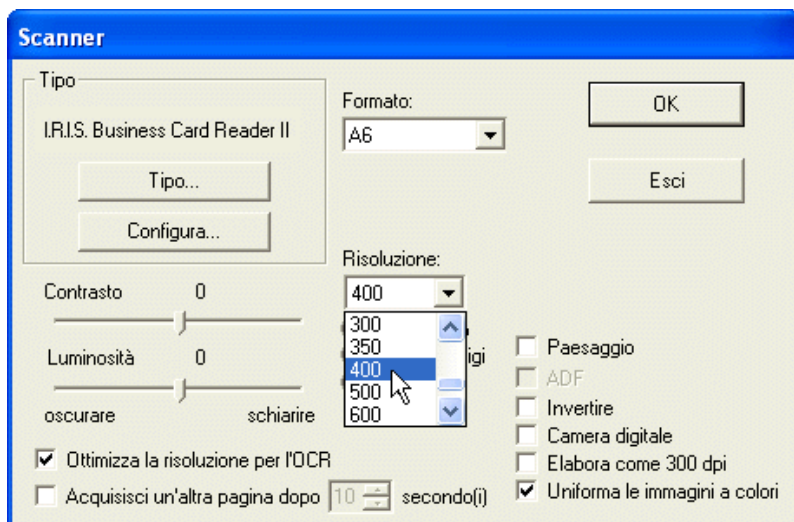
Se la lettura dei biglietti da visita è utilizzata come applicazione principale, quando si acquisisce una grande quantità di biglietti da visita per uno stand in fiera, è possibile ricorrere allo scanner **IBCR-II** di I.R.I.S. per i biglietti da visita. Gli **scanner** dedicati sono ottimizzati per la scansione **dei biglietti da visita** e presentano numerosi vantaggi rispetto agli altri tipi di scanner: occupano poco spazio sulla scrivania, gestiscono biglietti con diversi tipi di carta e qualità di stampa e l'utente non si deve preoccupare di selezionare il formato corretto.

Utilizzare l'interfaccia Twain dell'IBCR-II per attivare la modalità di **scansione automatica**. Inviare una volta il comando di scansione e attivare l'opzione "Scansione Automatica" nell'interfaccia utente Twain.



È possibile inserire un biglietto da visita dopo l'altro: una volta posizionato il biglietto da visita nello scanner, la scansione viene avviata...

Si consiglia di eseguire la scansione a una **risoluzione** di 400 dpi per ottenere una buona lettura dei biglietti.



Quando si utilizza uno **scanner piano**, è possibile acquisire contemporaneamente vari biglietti da visita ponendoli sul piano dello scanner e frazionandoli tramite il software. Per eseguire la scansione di più biglietti da visita, lo sfondo dovrà essere nero. Esiste un modo molto facile per raggiungere questo risultato: eseguire la scansione con il coperchio aperto!

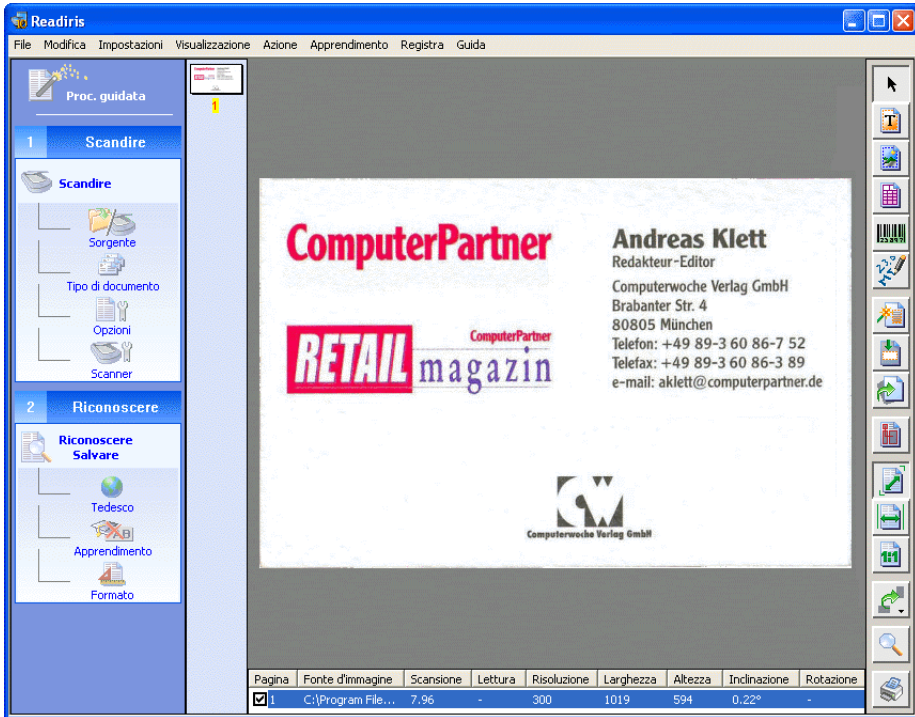


Selezionare l'opzione "Molteplici Biglietti da Visita in Una Sola Immagine" dal pulsante "Tipo di Documento" sulla barra degli strumenti principale (o con il comando "Tipo di Documento" del menu "Impostazioni"), quindi eseguire la scansione dei biglietti desiderati.



Con gli scanner piani, è sufficiente fare clic sul pulsante "Scandire" per avviare l'acquisizione dell'immagine. (È anche possibile selezionare il comando "Scansione" dal menu "Azione".)

La pagina completa acquisita non apparirà come tale sullo schermo del computer; saranno visualizzati solo i biglietti da visita segmentati.



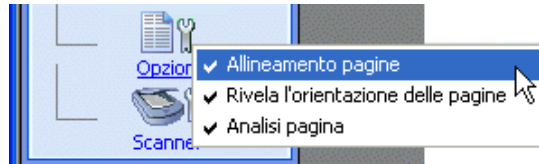
Se si è dimenticato di attivare il “frazionamento” con la modalità biglietti da visita multipli, fare clic sul comando "Estragga Biglietti da Visita" nel menu "Azione" per segmentare l'immagine grande nelle immagini effettive dei vari biglietti, eliminando il bordo nero non necessario.

Estragga biglietti da visita



Ma indipendentemente dal tipo di scanner, alcune opzioni vengono attivate in modo invisibile per ottimizzare il processo di lettura: l'allineamento pagine e la

rilevazione dell'orientamento pagine. Fare clic sul pulsante "Opzioni" della barra degli strumenti principale per individuarle.



È difficile pensare che i biglietti da visita inseriti in uno scanner piano siano posizionati (tutti) perfettamente dritti, lasciare quindi che il software si occupi del problema. Nel caso in cui si sia incidentalmente disattivata questa opzione, utilizzare il tasto "Allinea Pagina" sulla barra strumenti dell'immagine per allineare i biglietti da visita. (Potrebbe diventare noioso dover allineare ogni biglietto da visita manualmente!)

Lo stesso vale per la rilevazione dell'orientamento: lasciare che il software si occupi del lavoro, altrimenti sarà necessario capovolgere i biglietti da visita o ruotarli di 90° sullo scanner piano. (Utilizzare i tasti di rotazione della barra strumenti dell'immagine nel caso in cui questa opzione sia stata disattivata...)



ATTIVARE LA MODALITÀ LETTURA DEI BIGLIETTI

Le singole immagini del biglietto sono state correttamente inserite nella finestra dell'immagine. Assicurarsi di aver attivato la **modalità di lettura del biglietto** dal pulsante "Tipo di Documento" (o con il comando equivalente nel menu "Impostazioni").



Selezionare l'opzione "Biglietti da Visita" quando si desidera eseguire la scansione dei biglietti da visita una alla volta oppure selezionare l'opzione "Molteplici



Biglietti da Visita in Una Sola Immagine" per acquisire contemporaneamente diversi biglietti da visita dallo scanner a piano. Questa opzione non è sufficiente per frazionare l'immagine dello scanner: è necessario selezionare una delle due opzioni disponibili per il *riconoscimento* dei biglietti da visita! Selezionare "Pagine di Ttesto" per disattivare la lettura dei biglietti.

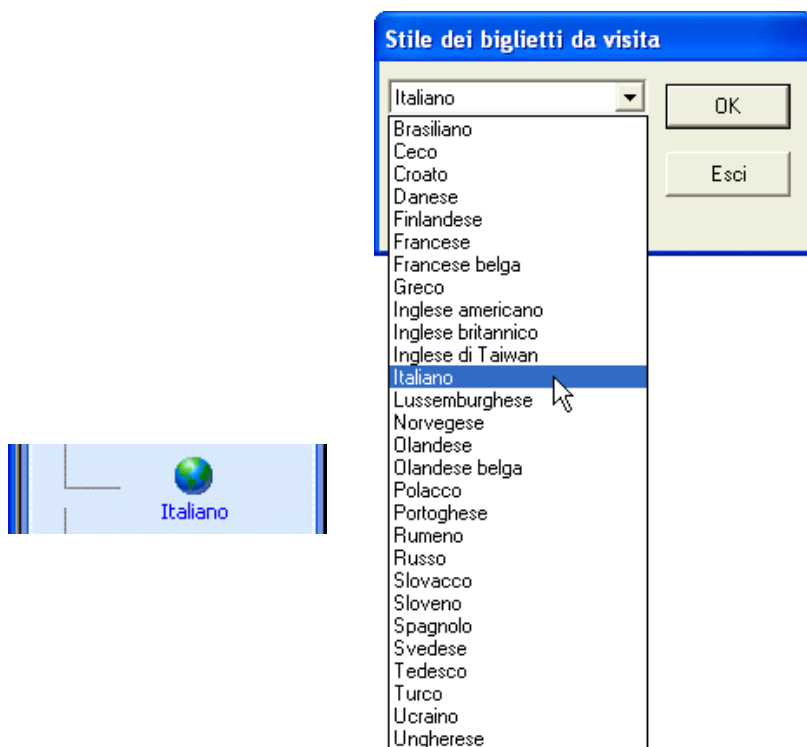
Il pulsante "Tipo di Documento" conferma se la modalità di lettura dei biglietti da vista è attivata.



È necessario abilitare queste modalità specifiche perché le routine OCR speciali vengano utilizzate per consentire al software di assegnare i dati riconosciuti ai campi specifici del database, come nome dell'azienda, indirizzi e-mail e di siti web ecc. Se Readiris riconosce la stringa di caratteri "Marco Rutelli", "sa" che Marco rappresenta il nome e Rutelli il cognome. Allo stesso modo, il sistema riconosce le qualifiche come "vice presidente" o "ingegnere", le città come Genova e Milano, le provincie come Rimini e Latina ecc.

Il sistema funziona con almeno 28 paesi: sono supportati i biglietti da visita dell'America del Nord e del Sud e i biglietti da visita dei paesi europei comprese le nazioni dell'Europa dell'Est. (Per leggere i biglietti da visita in cinese, giapponese, coreano e taiwanese, è necessaria l'opzione software "Modulo BCR asiatico". I biglietti da visita dei paesi asiatici scritti in inglese sono supportati dal software Readiris Corporate "standard"!)

Una volta abilitata questa modalità di lettura, il contenuto dell'elenco scorrevole "Lingua" cambia. Selezionare il paese corretto.



Selezionando il paese del biglietto da visita nella barra degli strumenti principale, non solo si seleziona la lingua del testo da riconoscere, ma si indica anche il layout generale del biglietto da visita. Ha un aspetto americano o italiano e si adatta a questo? Ogni paese presenta uno “stile” diverso di composizione dei biglietti da visita, gli americani dispongono l’indirizzo in modo diverso dai francesi, i codici di avviamento postale e i numeri di telefono olandesi presentano una sintassi diversa da quelli inglesi ecc. (Alcuni stili di biglietti corrispondono a diverse



lingue: Belgio e Canada hanno due lingue ufficiali, la Svizzera tre. Readiris individua automaticamente la lingua selezionata...)

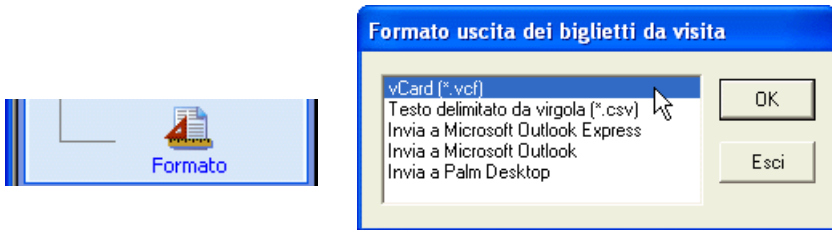
 <p>Recognition Systems</p> <p>Recognition Systems, Inc. 1520 Dell Avenue Campbell, CA 95008 USA www.handreader.com</p>	<p>Helen Thomas International Sales</p> <p>(408) 341-4100 (408) 341-4101 Fax (408) 341-4135 Direct</p> <p>helen_thomas@irco.com</p>	<p>APACABAR La Référence Logicielle</p> <p>Stéphan Mallavergne Coordinateur Technique</p> <p>10 rue Vercingétorix 75680 Paris cedex 14 Tel 01.56.54.80.22 Fax 01.56.54.80.01 E-mail : smallavergne@apacabar.fr Web : http://www.apacabar.fr</p> <p></p>
---	---	---

L'analisi dei campi comprende un terzo compito: Readiris non si limita all'analisi ma *formatta* anche il testo riconosciuto. Il sistema filtra i dati non necessari di un biglietto da visita, anche se rilevanti. Se nel biglietto da visita è indicata l'informazione "Telefono: 06-33 88 85 14", il risultato indicato al campo telefonico sarà "0633888514". Viene eliminata la parola "Telefono", anche se Readiris se ne serve in modo utile per riconoscere l'ubicazione del numero di telefono. Vengono eliminate anche le parentesi, il trattino e gli spazi compresi nel numero di telefono.

RICONOSCIMENTO DEI BIGLIETTI DA VISITA

Elaborare i biglietti da visita come di consueto e fare clic sul pulsante "Riconoscere-Salvare".

Tuttavia, si consiglia di definire il formato di output prima di questa fase. A questo punto, fare clic sul pulsante "Formato".



Sono compatibili direttamente i diversi popolari Personal Information Manager ("PIM"): Microsoft Outlook (Express) e il software PDA Palm Desktop. Per interfacciare con altre applicazioni, utilizzare i formati "universali" vCard e i dati delimitati da virgola. (I file vCard sono importati in qualsiasi applicazione compatibile vCard: se si fa doppio clic su un file vCard i contatti vengono aggiunti alla Rubrica Indirizzi di Windows.)

Ad esempio, il risultato del testo appare come questo, quando si inviano i risultati direttamente a Microsoft Outlook (Express).



Kirk Y. Matsuo - Contact

File Edit View Insert Format Tools Actions Help

Save and Close

General Details Activities Certificates All Fields

Full Name... Kirk Y. Matsuo

Job title: Product Manager Business & Home/Office Proc

Company: Seiko Instruments USA Inc.

File as: Matsuo, Kirk Y.

E-mail... kirk.matsuo@seikosi.com

Display as: Kirk Y. Matsuo (kirk.matsuo@seikosj.com)

Web page address: <http://www.seikosmart.com>

IM address:

Phone numbers

Business... 40843210

Home...

Business Fax... 408 4333261

Mobile...

Addresses

Business... 1130 Ringwood Court
95131 San Jose California

This is the mailing address

Contacts... Categories... Private

USO DELLA GUIDA IN LINEA

La panoramica di Readiris è conclusa. Non è possibile includere nel presente manuale alcune informazioni dell'ultimo minuto. Pertanto, per ulteriori informazioni su Readiris si consiglia di consultare la guida in linea.

A questo scopo, passare al menu "Guida". Il comando "Argomenti della Guida" e il relativo tasto di scelta rapida F1 consentono di spostarsi attraverso i numerosi argomenti della guida.

Benvenuti nella guida Readiris™!

- Utilizzare la guida in linea per maggiori informazioni su Readiris.
- La guida consente di trovare rapidamente una risposta alle vostre domande.
- Collegarsi al sito web I.R.I.S. per gli ultimi suggerimenti e le più recenti versioni dei prodotti.

©2005 Copyright [I.R.I.S.](#). Tutti i diritti riservati

Gli altri comandi del menu "Guida" indicano come ricevere il supporto sul prodotto e come contattare I.R.I.S., dando l'accesso alla pagina iniziale di I.R.I.S.

