

# Readiris<sup>TM</sup>PRO9

## MANUALE DELL'UTENTE



Readiris Pro

© 2004 I.R.I.S. Tutti diritti riservati

Tecnologia OCR di I.R.I.S.

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

© 2004 I.R.I.S. Tutti diritti riservati

---

## **RISPARMIA TEMPO SENZA RIBATTERE IL TESTO!**

---

Congratulazioni per l'acquisto di Readiris. Questo pacchetto software vi sarà indubbiamente di grande aiuto per riprendere testi, tabelle ed elementi grafici.

Efficiente quanto un computer, richiede innanzitutto l'immissione a tastiera delle vostre informazioni. Se vi è mai capitato di ribattere una relazione di 15 pagine o un'ampia tabella numerica, sapete bene quanto possa essere noioso e dispendioso. Utilizzate questo pacchetto OCR, che rappresenta lo stato dell'arte, per inserire automaticamente i vostri testi nelle vostre applicazioni e acquisirete un livello di efficienza e comodità senza precedenti!

Eseguite la scansione di un documento stampato o dattiloscritto, indicate le aree d'interesse - o lasciate che il sistema le individui per voi - ed effettuate il riconoscimento dei caratteri. I documenti costituiti da diverse pagine vengono elaborati dall'inizio alla fine in un'unica volta. 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 battitura.

L'assistenza alla procedura ("wizard") vi guida con comodamente nel processo di OCR: rispondete ad alcune semplici domande e otterrete facilmente e rapidamente dei risultati con Readiris. Potete inviare i risultati direttamente al vostro sistema di videoscrittura e al foglio elettronico. Per riconoscere i fax e convertire documenti PDF, basta trascinare il archivio da Esplora Risorse di Windows alla finestra di applicazione Readiris o cliccare con il tasto di destra 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 del sistema di videoscrittura. 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 dal font 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 riconosce fino a 107 lingue: supporta tutte le lingue americane ed europee, incluse le lingue dell'Europa centrale, le lingue baltiche, il greco e le lingue cirilliche ("russi"). (Opzionalmente, potete leggere anche quattro lingue asiatiche - il giapponese, il cinese tradizionale e quello semplificato e il coreano.) Readiris gestisce persino testi in più alfabeti: il software individua le parole "occidentali" che compaiono nei documenti in lingua greca, cirillica e asiatica - numerosi nomi propri non translitterabili, marche ecc. vengono scritti usando le lettere dell'alfabeto latino.

Readiris utilizza criteri di linguistica *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 scansionati o copiati di qualità scadente nei quali i corpi del font sono troppo chiari o troppo scuri. I caratteri uniti ("legature") vengono scomposti e le forme frammentate, come ad esempio i simboli di matrice di punti, vengono ricomposte.

La verifica utente sotto forma di messaggi a comparsa ("popup") non solo individua i caratteri dubbi, ma aumenta anche 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 vi consente di insegnare a Readiris caratteri particolari come ad esempio simboli matematici e segni grafici decorativi ("dingbats"), ma anche di gestire font distorti presenti nei documenti reali.

Al fine di aumentare ulteriormente la vostra produttività, Readiris non solo riconosce i vostri testi, ma è in grado di *formattarli* al vostro posto! Con "l'autoformattazione" Readiris ricrea una copia facsimile del documento scansionato, 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. Poiché Readiris supporta senza sforzo la scansione di immagini con scala dei grigi e a colori, è possibile riprendere qualsiasi elemento grafico, sia esso grafica al tratto, foto in bianco e nero o illustrazioni a colori. Se un documento contiene delle tabelle, Readiris le riorganizza in celle reali e ricrea i bordi delle celle delle tabelle originali.

In altre parole, Readiris vi consente di archiviare una copia conforme dei vostri documenti, pur trattandosi ora di file di testo compatti editabili e non di un'immagine scansionata! Sono disponibili vari livelli di formattazione, la scelta è lasciata all'utente.

Readiris supporta un'ampia gamma di scanner molto diffusi: numerosi scanner da tavolo, scanner con alimentazione a foglio e periferiche “tutto in uno” o “multifunzionali” (“MFP” - “multifunctional peripherals”) oppure camere digitali possono essere utilizzati. Inoltre Readiris supporta lo standard di scansione Twain e alcune piattaforme di scansione.

## **SOMMARIO**

---

Risparmia tempo senza ribattere il testo! .....	III
Sommario .....	V
Diritti e copyright .....	VII

### **Capitolo 1: Installazione**

Requisiti del sistema .....	1-1
Installazione del software Readiris .....	1-1
Disinstallazione del software Readiris .....	1-4
Programma di disinstallazione di Readiris .....	1-5
(Dis)installazione assistita di Windows .....	1-5
Installazione delle opzioni software .....	1-6
Installazione di prodotti correlati .....	1-8
File installati .....	1-10
Archivio “Leggimi” e documentazione .....	1-10
Piloti dello scanner .....	1-10
Registrarsi per far valere i propri diritti! .....	1-10
Assistenza al prodotto .....	1-12



## **Capitolo 2: Informazioni guidate**

Avviare il software .....	2-1
Il primo avvio .....	2-2
Scoprire l'interfaccia Readiris .....	2-3
Una prima esercitazione .....	2-6
Zoomare sulle immagini .....	2-11
Uno, scomporre un'immagine scandita .....	2-14
Uno e mezzo, selezione delle finestre .....	2-16
Due, organizzare manualmente in finestre un'immagine scandita .....	2-18
Tre, salvare modelli di organizzazione in finestre .....	2-23
Readiris permette di viaggiare in tutto il mondo .....	2-24
Se necessario, Readiris cambia le lingue .....	2-30
Definire le caratteristiche del documento .....	2-33
Readiris diventa ogni volta più intelligente! .....	2-35
Apprendi .....	2-37
Non Apprendi .....	2-38
Elimina .....	2-39
Ripristina .....	2-39
Termina .....	2-39
Interrompi .....	2-39
Il ruolo dei dizionari dei font .....	2-39
Inviare il risultato all'applicazione .....	2-42
Salvare i risultati in un file di testo .....	2-46
Per creare un documento esportabile .....	2-50
... O leggerli .....	2-54
Riconoscimento di pagine multiple .....	2-56
Modifica di documenti multipagina .....	2-64
Creazione di un nuovo documento .....	2-66
Riconoscimento di zone di testo .....	2-67
Organizzazione dell'uscita di testo .....	2-68
Installazione dello scanner .....	2-70
Colorate le vostre scansioni! .....	2-72
Apparecchi diversi, diverse risoluzioni .....	2-75
Salvataggio di impostazioni di difetto .....	2-79
Salvataggio di impostazioni specifiche .....	2-80
Scansione di documenti .....	2-81
Regolazione dell'immagine scandita .....	2-84

La procedura guidata fa tutto per voi! .....	2-88
Readiris ricrea il layout del documento .....	2-90
Colonne, non cornice! .....	2-94
Formattazione di testo, parte 2 .....	2-96
Esportare il testo diverse volte .....	2-97
Salvataggio separato degli elementi di grafica .....	2-98
Lettura di fax e riconoscimento posticipato .....	2-100
Riconoscimento di tabelle .....	2-101
Uso della guida in linea .....	2-107

## **DIRITTI E COPYRIGHT**

---

Il software Readiris è stato progettato e realizzato da I.R.I.S. Tecnologia OCR, Connectionist, AutoFormat e Linguistica di I.R.I.S. I.R.I.S. possiede il copyright del software Readiris, della tecnologia OCR, della tecnologia linguistica, del sistema di guida in linea e del presente manuale.

AutoFormat, Cardiris, Connectionist, 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](http://www.apache.org)).

Acrobat e Reader sono marchi (registrati) di Adobe. AsianBridge è un marchio di TwinBridge. AsianSuite è un marchio di UnionWay. Excel, Windows e Word sono marchi registrati di Microsoft. Intel è un marchio registrato di Intel.





# Capitolo 1

## I N S T A L L A Z I O N E

Questo capitolo spiega i requisiti del sistema e l'installazione del software Readiris.

### REQUISITI DEL SISTEMA

---

Di seguito viene indicata la configurazione del sistema strettamente necessaria per utilizzare Readiris:

- ☐ un PC Intel 486 o compatibile. Si consiglia un PC Pentium.
- ☐ 32 MB di RAM. Si consigliano 64 MB di RAM per l'elaborazione di immagini con scala dei grigi e a colori.
- ☐ 110 MB di spazio libero su disco. Sono sufficienti 95 MB di spazio su disco se le immagini d'esempio vengono lasciate sul CD-ROM.
- ☐ il sistema operativo Windows XP, Windows ME, Windows 2000, Windows 98 o Windows NT 4.0.

È possibile che alcuni **piloti di scanner** non sono compatibili con la (le) ultima(e) versione(i) di Windows. Consultare la documentazione fornita in dotazione unitamente allo scanner.

### I N S T A L L A Z I O N E D E L S O F T W A R E R E A D I R I S

---

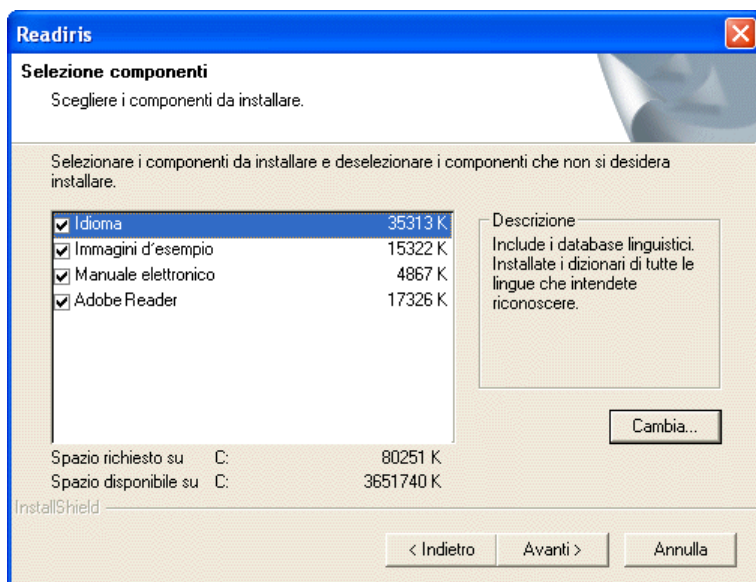
Il software Readiris viene fornito esclusivamente su **CD-ROM ad esecuzione automatica**. Per installarlo è sufficiente inserire il CD-ROM nel drive per CD-ROM e attendere l'avvio del programma di installazione. Seguire le istruzioni sullo schermo.



Nel caso in cui, inserito il CD-ROM nell'apposito drive, l'installazione non venga avviata, avviare il programma di setup MENU.EXE per installare il software.

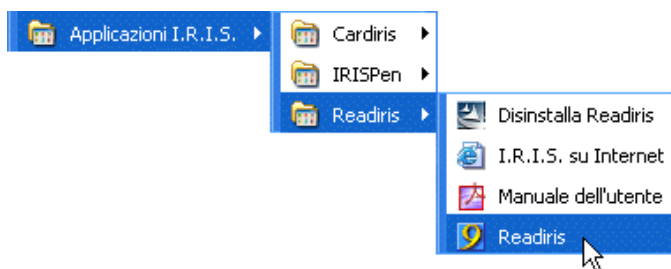
Gli utenti di Windows XP, Windows 2000 e di Windows NT devono accertarsi di possedere i **diritti di accesso** appropriati - se necessario contattare l'amministratore del sistema.

Per l'installazione vengono offerte alcune opzioni. Assicurarci di installare i **database linguistici** di tutte le lingue che si intende leggere. Tutti i vocabolari sono installati per difetto. 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 **manual elettronico** viene copiato per difetto sul disco fisso. È possibile anche lasciarlo su CD-ROM.

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



In tal modo è possibile avere il **collegamento** di Readiris direttamente della scrivania. Di questa maniera, si può lanciare Readiris direttamente della scrivania.



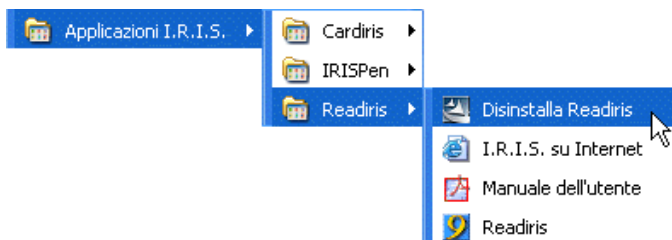
## DISINSTALLAZIONE DEL SOFTWARE READIRIS

---

Esistono solo due modi corretti per disinstallare Readiris: utilizzare il programma di disinstallazione di Readiris e la (dis)installazione assistita di Windows. Si consiglia vivamente di *non* disinstallare Readiris o i relativi moduli del software cancellando manualmente i file del programma.

## **Programma di disinstallazione di Readiris**

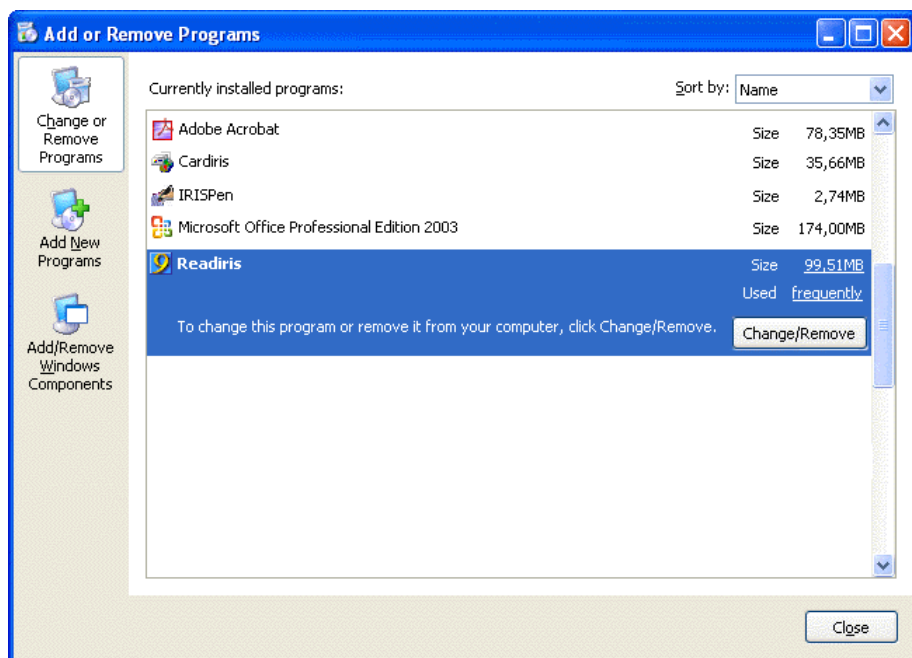
Per avviare il programma di disinstallazione di Readiris, selezionare "Disinstalla Readiris" nel sottomenu "Applicazioni I.R.I.S. - Readiris" e seguire le istruzioni sullo schermo.



## **(Dis)installazione assistita di Windows**

Per utilizzare wizard di (dis)installazione di Windows, eseguire le seguenti operazioni.

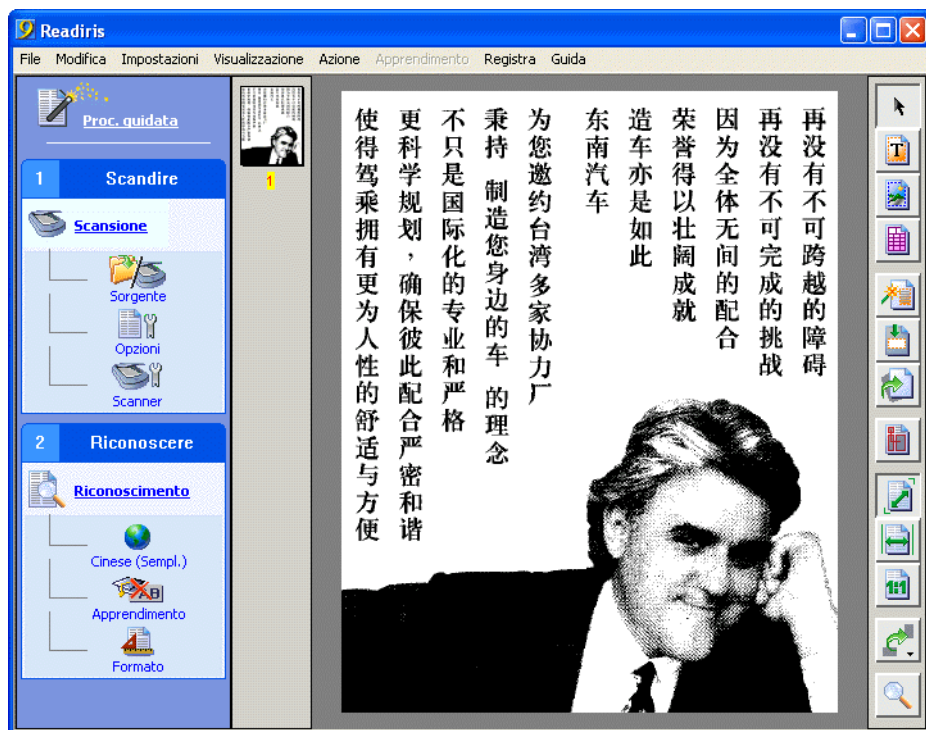
- ☐ Fare clic su "Impostazioni" nel menu di Windows "Avvio" e passare al "Pannello di Controllo".
- ☐ Fare doppio clic sull'icona "Aggiungi/Rimuovi Programmi" nel pannello di controllo.



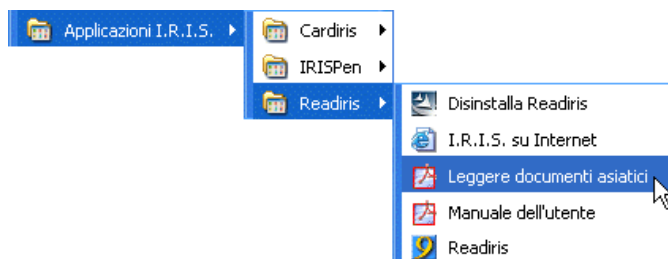
- ☐ Seguire le istruzioni sullo schermo per rimuovere il software Readiris.

## INSTALLAZIONE DELLE OPZIONI SOFTWARE

Esiste una singola opzione software disponibile per il software Readiris: il software aggiuntivo “Modulo d’OCR asiatico”. Esso consente di leggere il giapponese, il cinese tradizionale e quello semplificato e il coreano. Questo software è fornito anche esso su un CD-ROM con avviamento automatico.



Installando questa opzione sarà disponibile una documentazione specifica su come riconoscere i documenti asiatici.



## INSTALLAZIONE DI PRODOTTI CORRELATI

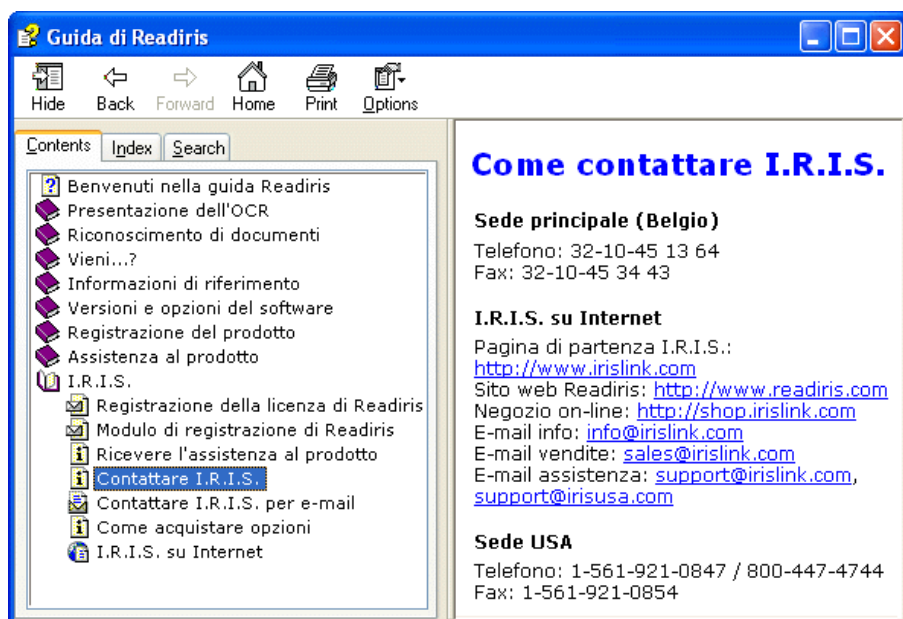
---

In base al pacchetto software acquistato, Readiris può essere fornito di una versione di valutazione del prodotto correlato Cardiris, l'**organizzatore elettronico di biglietti da visita**.

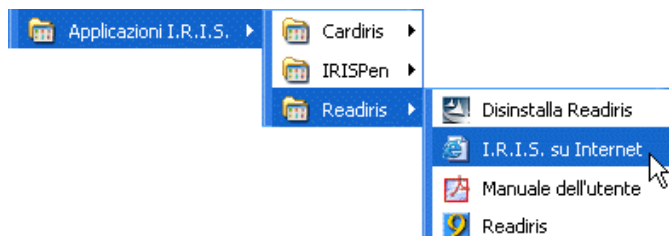
Qualora questo pacchetto software è incluso gratuitamente nel CD-ROM Readiris, la loro installazione viene ugualmente eseguita utilizzando il CD-ROM ad avvio automatico e seguendo le istruzioni sullo schermo.

Per ulteriori informazioni su questo pacchetto software correlato, contattare I.R.I.S.; il comando "Contatta I.R.I.S." nel menu "Guida" di Readiris spiega come mettersi in contatto con I.R.I.S.





Un'icona di applicazione nel sottomenu "Applicazioni I.R.I.S. - Readiris" nel menu "Programmi" porta direttamente alla **home page** I.R.I.S. Lo stesso avviene dalla videata di avvio di Readiris e selezionando il comando "I.R.I.S. su Internet" nel menu "Guida" di Readiris.





## 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 (dis)installazione assistita di Windows. Vedere sopra.

### Archivio “Leggimi” e documentazione

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

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

### Piloti dello scanner

Non esitate a contattare il produttore o il rappresentante dello scanner qualora i problemi con i piloti (“driver”) dello scanner persistano. La maggior parte dei produttori consente di scaricare le versioni aggiornate dei piloti dello scanner dai loro siti web.

## REGISTRARSI PER FAR VALERE I PROPRI DIRITTI!

---

Non dimenticare di registrare la licenza Readiris! Questo consentirà a I.R.I.S. di tenersi aggiornati 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** al prodotto e le **offerte speciali**.

È possibile effettuare la registrazione in numerosi modi: inviando la scheda di registrazione o trasmettendo via fax il duplicato elettronico, contattando I.R.I.S. durante l'orario di lavoro e compilando un modulo di registrazione sul sito web I.R.I.S.!



L'assistenza per la registrazione di Readiris che si trova nel menu "Registra" di Readiris vi guiderà agevolmente nella procedura di registrazione.

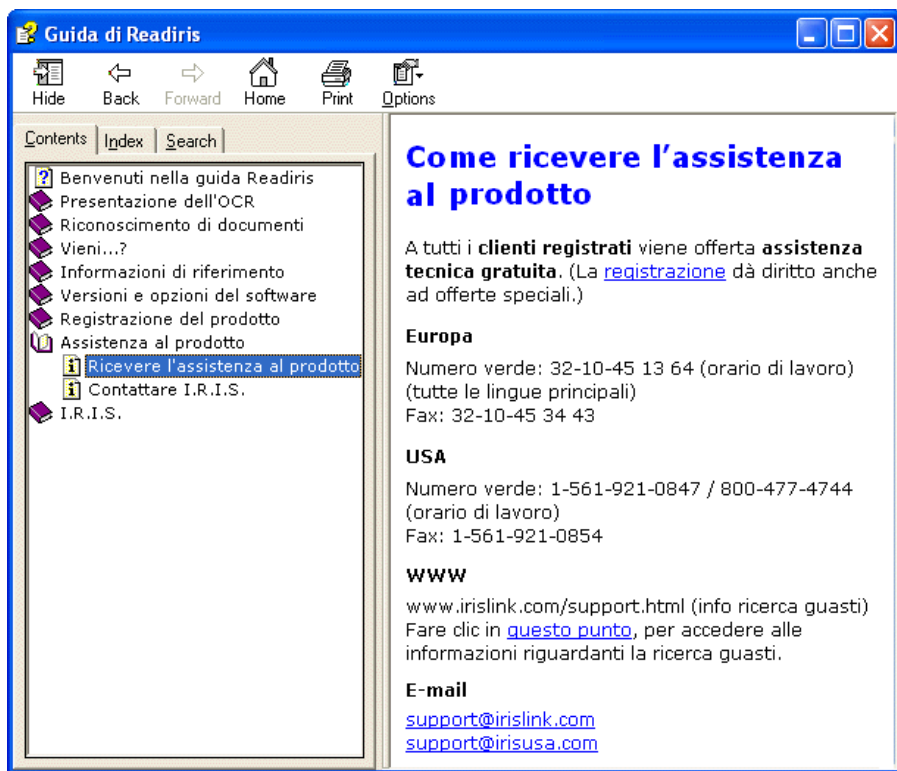


In base alla versione acquistata, riceverete il **tasto personalizzabile** ("softkey") necessario per continuare ad utilizzare il software Readiris dopo un mese.

## **ASSISTENZA AL PRODOTTO**

---

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





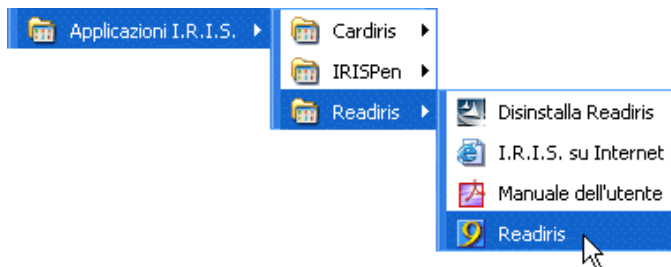
# Capitolo 2

## INFORMAZIONI GUIDATE

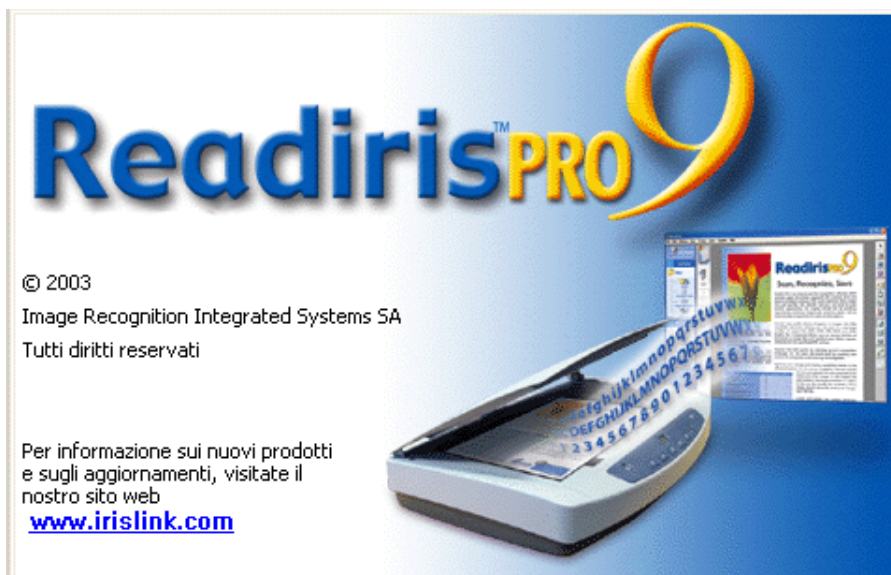
Readiris è un pacchetto OCR che rappresenta lo stato dell'arte, dotato di numerose caratteristiche avanzate. Questo capitolo illustra tutte le principali caratteristiche, informazioni utili e consigli relativi all'uso di Readiris.

### AVVIARE IL SOFTWARE

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



Vengono visualizzati la videata di avvio e la finestra dell'applicazione Readiris. La videata di avvio visualizza la versione e il copyright del software Readiris. Consente inoltre di accedere direttamente alla **home page** I.R.I.S. - basta cliccare sulla URL per visitare il sito web I.R.I.S. Facendo clic con il mouse in qualsiasi punto, la videata scompare.



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

## IL PRIMO AVVIO

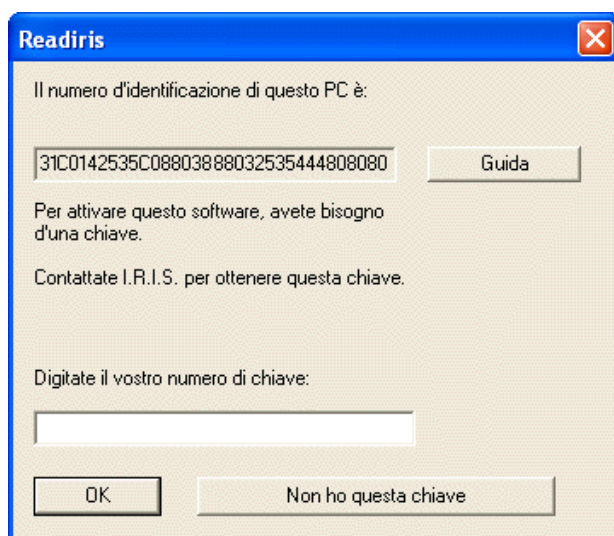
---

In base al 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 ad un mese. In seguito alla registrazione, I.R.I.S. vi fornirà un **tasto personalizzabile** ("softkey") che consente di continuare ad utilizzare Readiris dopo il primo mese.



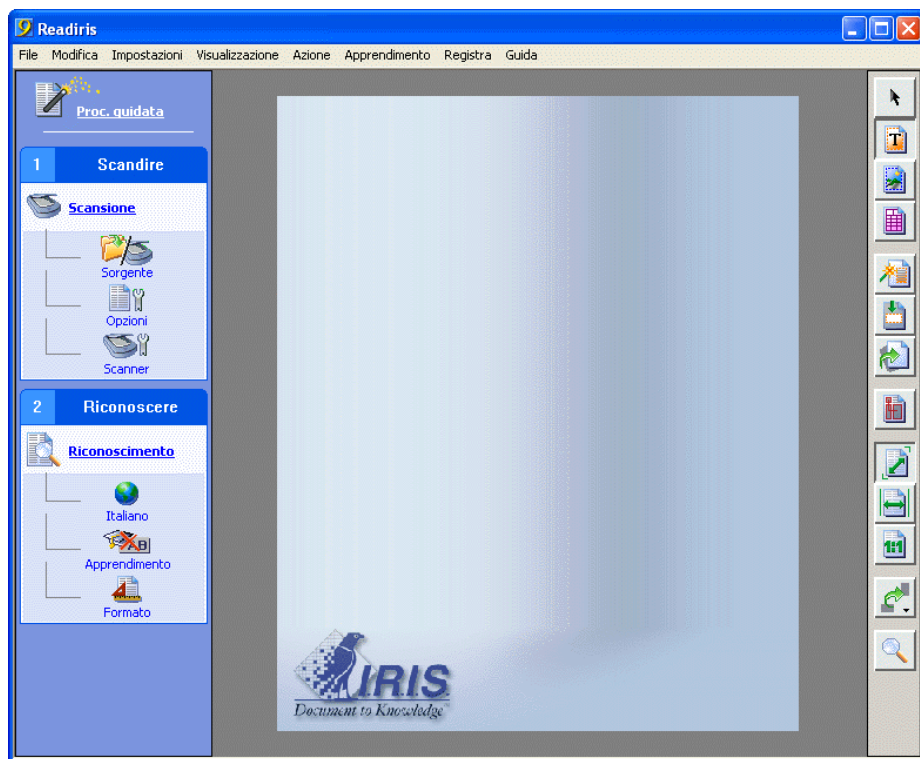
Per creare il tasto personalizzabile è necessario il **numero di identificazione** personale; durante la registrazione della licenza assicurarsi che tale numero sia disponibile o indicato.



## SCOPRIRE L'INTERFACCIA READIRIS

---

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



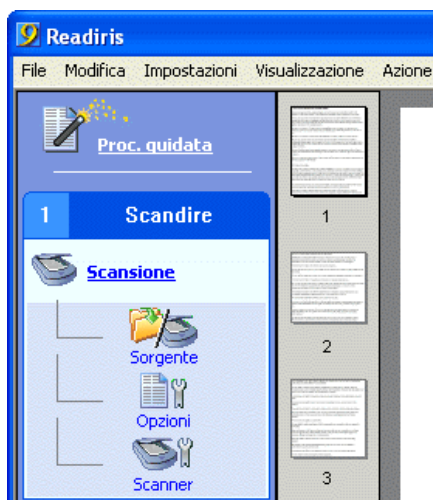
Lo stesso vale per la **barra degli strumenti immagine** posta sul lato destro della finestra dell'applicazione: questa contiene tutti i comandi necessari durante l'anteprima dell'immagine. La **barra degli strumenti principale** posta sul lato sinistro consente un rapido accesso a tutti i comandi generali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

Per scoprire quale comando corrisponde ad un determinato pulsante, tenere per qualche secondo il puntatore del mouse sul pulsante: una **casella informativa** indicherà all'utente la funzione del comando.



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

Durante l'elaborazione pagine, viene aggiunta sul lato sinistro una barra addizionale, la **barra strumenti della pagina**. Rappresenta le diverse pagine del documento e consente l'accesso ai comandi pagina cliccando a destra (menu "Contesto").



## UNA PRIMA ESERCITAZIONE

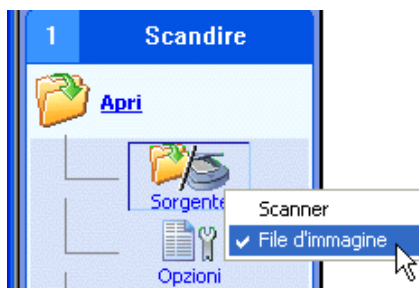
Il modo migliore per acquisire dimestichezza con il funzionamento di Readiris è indubbiamente utilizzarlo. Unitamente al software viene fornita una serie di **immagini prescandite** che consentono di iniziare persino quando lo scanner non è collegato al computer. Segue la spiegazione.

Il tasto "Sorgente" della barra strumenti principale consente di indicare se si intende utilizzare un'immagine scandita o prescandita come fonte di immagini.

Immagini a colori, con scala dei grigi e in bianco e nero sono supportate in eguale misura. Readiris consente di aprire documenti PDF Adobe Acrobat, immagini JPEG, immagini Paintbrush (PCX), immagini fax DCX (una versione multipagina del formato Paintbrush), immagini PNG, immagini TIFF (non compresse, compresse di LZW, PackBits, Gruppo 3 e Gruppo 4), immagini TIFF multipagina e bitmap Windows (BMP).

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

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



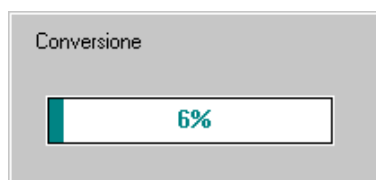
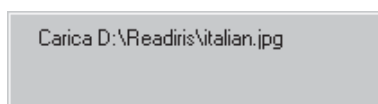
Cliccare quindi sul pulsante "Apri". (Selezionando il disco come sorgente d'immagine, il tasto "Scansione" viene sostituito dal tasto "Apri", mentre il corrispondente comando "Scansione" del menu "Azione" viene sostituito dal comando "Apri".)



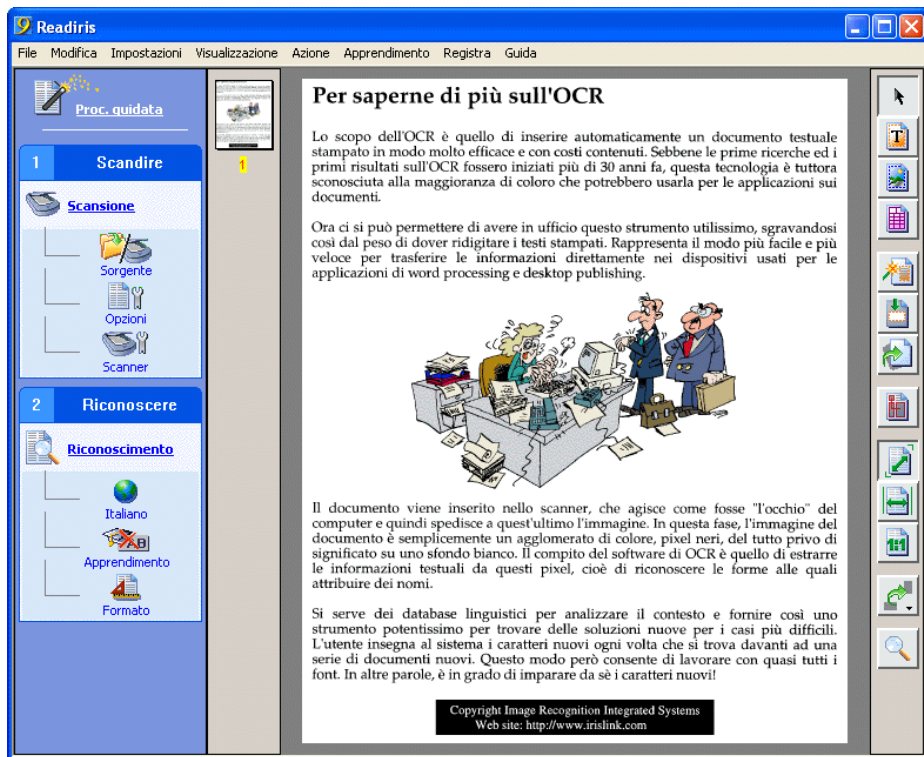
È possibile inoltre selezionare il comando "Apri" dal menu "File" e aprire direttamente un'immagine prescandita - questa operazione è possibile anche quando la fonte di immagini corrente è lo scanner.



Il sistema chiederà di selezionare un file di 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.



L'immagine viene quindi visualizzata nella zona di immagine. La barra strumenti immagine indica che una singola pagina è caricata in Readiris.

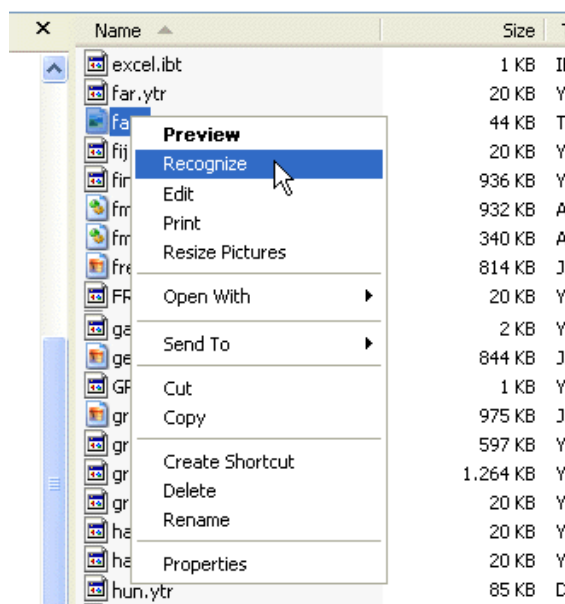


Un terzo modo per aprire immagini prescandite è costituito dall'utilizzo di **“trascinare selezione”**: trascinando immagini da Esplora Risorse di Windows alla zona d'immagine di Readiris o sull'icona Readiris potrete aprire le immagini immediatamente.



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





Non significa che l'OCR è stato esaguito immediatamente: per dare all'utente piena flessibilità, Readiris solamente è lanciato e l'immagine è aperta.

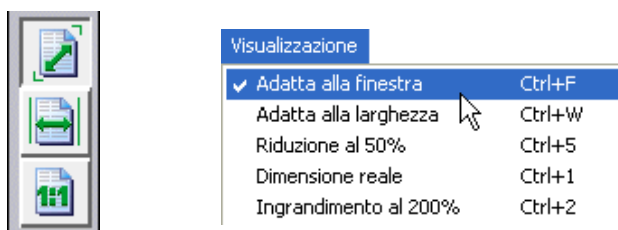
La barra degli strumenti immagine posta sul lato destro della finestra dell'applicazione Readiris contiene tutti i comandi necessari durante l'anteprima dell'immagine, i strumenti relativi all'indicazione delle zone di interesse, alla rotazione dell'immagine, al suo zoom ecc.

## ZOOMARE SULLE IMMAGINI

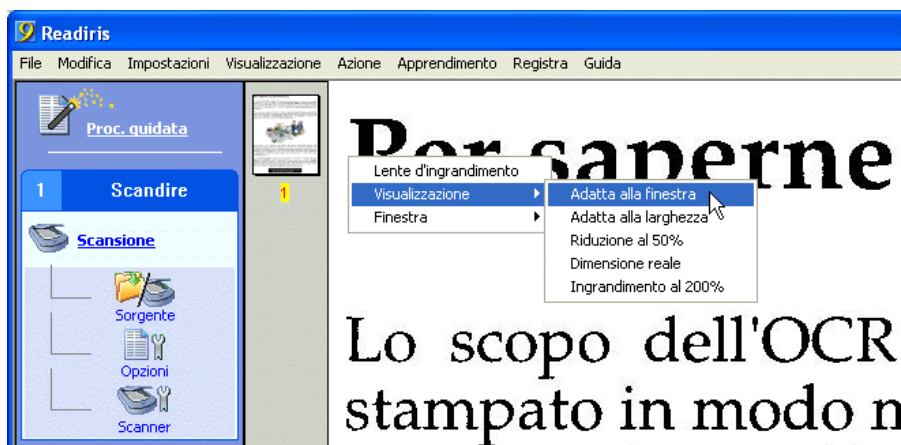
Readiris è dotato di numerosi comandi che consentono di **zoomare** sull'immagine scandita, ad esempio per verificare la qualità della scansione.



La barra degli strumenti immagine contiene pulsanti che consentono di zoomare 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!)



Prendete nota pure che i livelli dello zoom sono ottenibili mediante il clic di destra. Cliccate il tasto destro del mouse per entrare nel menu "Contesto" e selezionate il livello di zoom che desiderate.

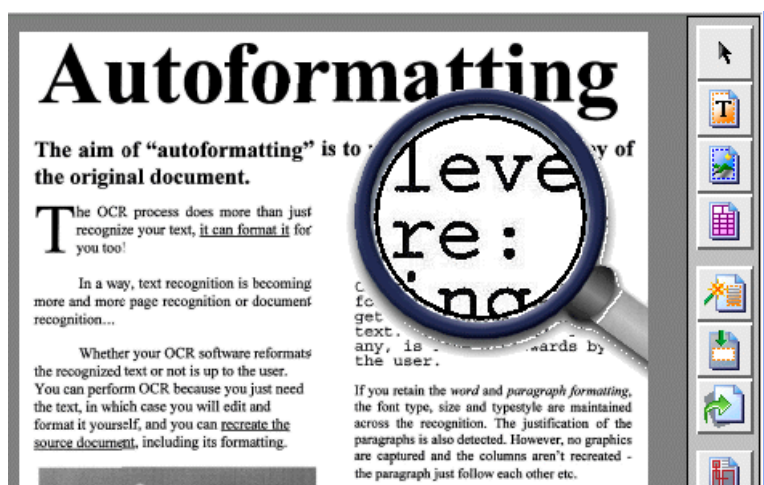
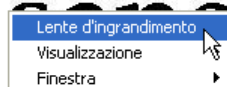


Potete, inoltre, cliccare *due volte* con il tasto destro del mouse su una zona dell'immagine scandita per zoomare istantaneamente in dimensione reale. Ripetete l'operazione di nuovo per zoomare in allontanamento.

Potete utilizzare, infine, la **lente di ingrandimento** per zoomare in avvicinamento i dettagli di un documento scandito. La lente di ingrandimento è anche disponibile nel menu "Contesto" cliccando con il lato destro del mouse sull'immagine.



# Per saperne



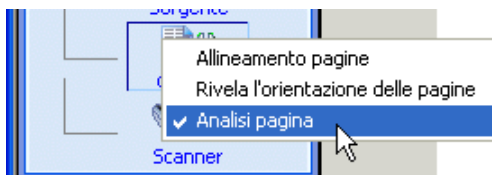


## UNO, SCOMPORRE UN'IMMAGINE SCANDITA

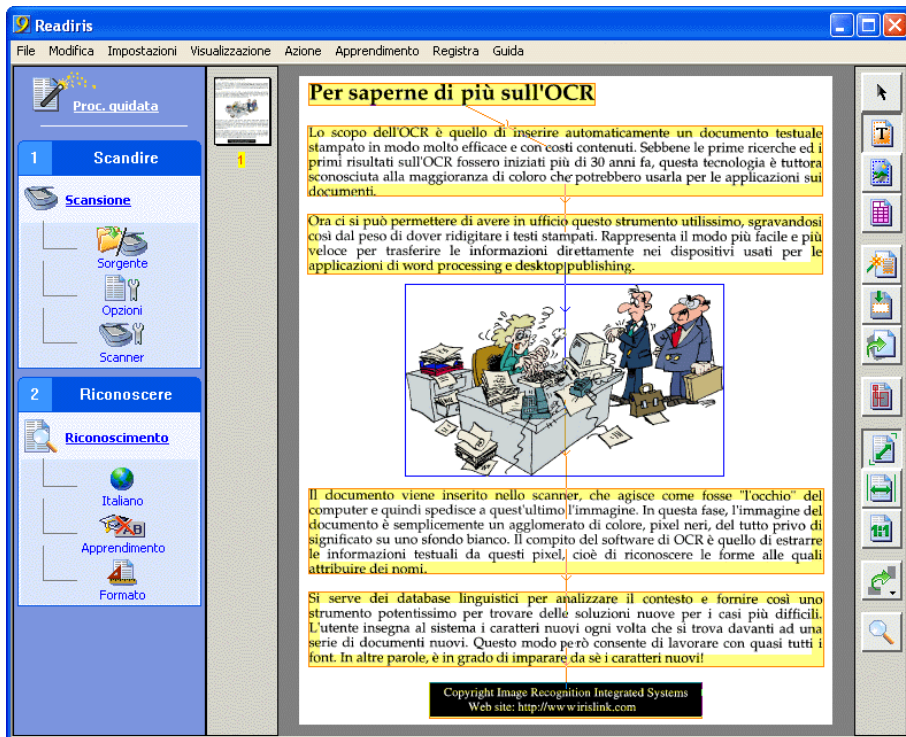
---

Una volta scandita l'immagine, è necessario indicare quali parti si desiderano convertire in testo editabile, tracciando strutture definite “finestre” intorno alle zone di interesse.

In realtà, abilitando l'opzione "Analisi Pagina" sotto il tasto "Opzioni" della barra strumenti principale (o sotto il menu "Impostazioni"), Readiris lo farà automaticamente.



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



La scomposizione della pagina utilizza tre **tipi di finestra**: finestra di testo, finestra grafica e finestra di tabella. Readiris riconosce i blocchi di testo, le tabelle e le zone di grafica contenenti foto, illustrazioni ecc. nella pagina. (Il salvataggio dei grafici e il riconoscimento delle tabelle sarà illustrato più avanti.)

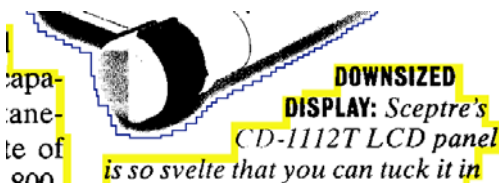
Un **codice colore** indica il tipo di finestra: le zone di testo hanno un bordo giallo, le finestre di grafica hanno un bordo blu e le tabelle, un bordo bordò.



Il numero delle finestre è sempre indicato nelle caselle informative degli strumenti "Finestra di Testo", "Finestra Grafica" e "Finestra di Tabella".



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 alcun problema: mentre l'anteprima visualizza sullo schermo il documento correttamente scandito, 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.)

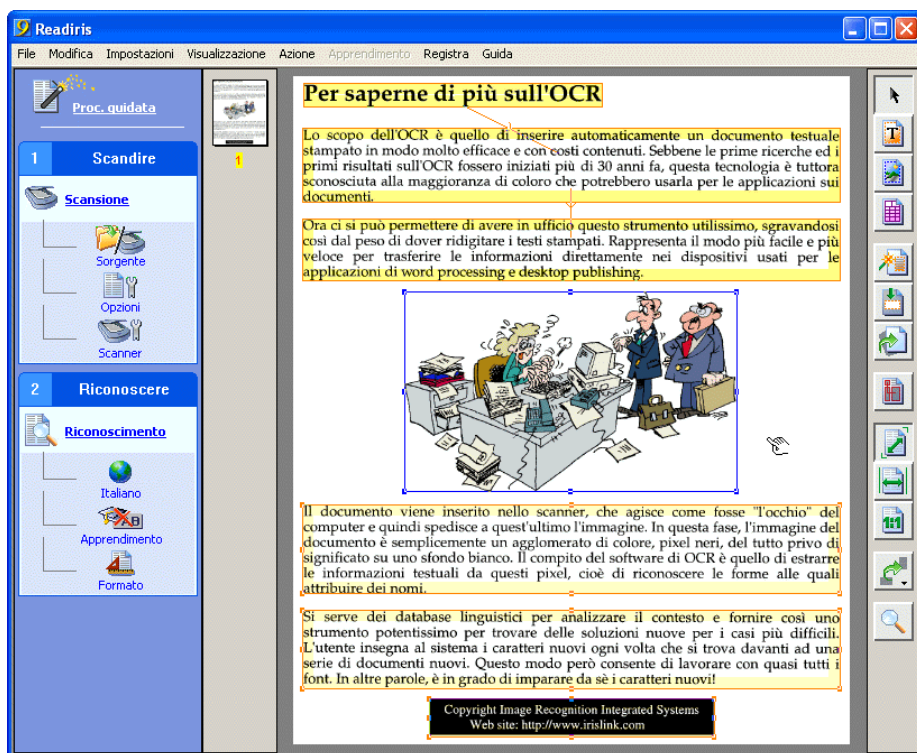
## UNO E MEZZO, SELEZIONE DELLE FINESTRE

Readiris non solo individua i vari blocchi ma li *ordina*: le zone vengono selezionate per difetto a testa in giù, da sinistra a destra, per adattarsi ai documenti organizzati in colonne.

È possibile modificare l'**ordine di selezione**. Cliccate sul tasto "Ordina" della barra strumenti immagine. Il cursore del mouse assume la forma del puntatore non appena il modo d'ordinamento viene abilitato.



Cliccare sulle finestre che si desidera includere. Le finestre *non* selezionate vengono semplicemente ignorate ed escluse dal riconoscimento. È facile vedere quali finestre sono selezionate o meno: le finestre selezionate hanno un colore più intenso, quelle non selezionate un colore più tenue.



L'analisi della pagina è abilitata per difetto. Qualora si sia involontariamente disabilitata l'analisi della pagina o cancellato alcune finestre e si intenda 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.



Selezionate la lingua del documento *prima* d'eseguire l'analisi della pagina quando trattate dei documenti asiatici. Delle routine specifiche sono utilizzate per queste lingue: lo spazio tra righe è d'abitudine più grande nei documenti asiatici che nei documenti occidentali, il testo è composto di piccoli grafici ("ideogrammi") che potrebbero facilmente essere visti come delle zone grafiche nei documenti occidentali ed il testo si legge d'alto in basso, di destra a sinistra. E se avete dimenticato di selezionare la lingua voluta, potete farlo in seguito. Readiris esegue di nuovo automaticamente l'analisi della pagina!

Alcuni documenti presentano puntini "sporadici" sulla pagina, che possono determinare un bordo nero sulla pagina tutto intorno all'immagine reale ecc. Per cancellare tutte le piccole finestre - si presume che esse non contengano alcun testo - e riprendere la finestra restante, potete cliccare sul comando "Sopprimi le Piccole Finestre" che si trova nel menu "Modifica".

Sopprimi le piccole finestre ... Ctrl+M

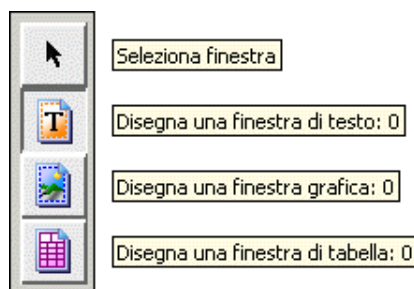


## **DUE, ORGANIZZARE MANUALMENTE IN FINESTRE UN'IMMAGINE SCANDITA**

---

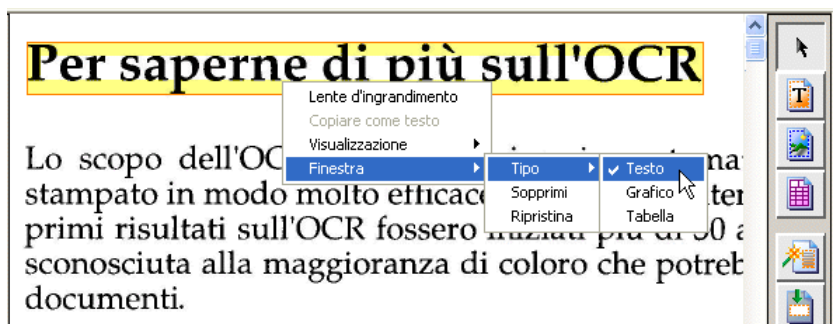
L'analisi della pagina consente di organizzare automaticamente in finestre una pagina scandita. In alternativa, è possibile dividere manualmente in zone un'immagine mediante gli **strumenti di organizzazione in finestre** di Readiris.





Per **disegnare** un rettangolo intorno a una zona di interesse, selezionare lo strumento corrispondente nella barra degli strumenti immagine, cliccare con il cursore nell'angolo superiore sinistro della finestra, allungare la finestra spostando il mouse nell'angolo inferiore destro e cliccare nuovamente. (Non sono consentiti lati inferiori a 1 mm. perché non potrebbero contenere neppure un singolo carattere.)

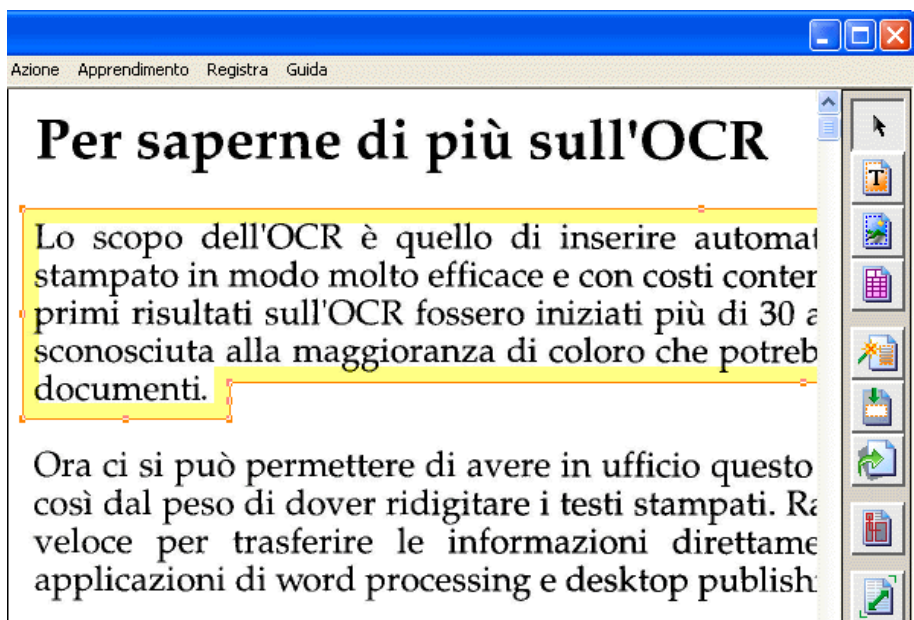
Non preoccupatevi se avete selezionato il tipo zona errato: è possibile modificare rapidamente il tipo cliccando con il lato destro del mouse su una finestra e selezionando il comando "Finestra - Tipo" dal menu "Contesto".



Le finestre vengono ordinate automaticamente sulla base dell'ordine di creazione: le frecce 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 di tabelle poligonali.)



Inoltre, l'organizzazione in finestre manuale può essere associata con la selezione delle finestre: è possibile disegnare nuove finestre anche quando il modo d'ordinamento è attivato. Utilizzare la funzione d'ordinamento per includere un numero di finestre individuate 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 nel modo d'ordinamento tutte le finestre non selezionate saranno immediatamente cancellate!

Per modificare, spostare ed eliminare finestre, è necessario innanzi tutto **selezionarle**. Per fare ciò, selezionare lo strumento "Selezione Finestra" (lo strumento "freccia") della barra strumenti immagine e fare clic all'interno di una finestra. Compiono così marcatori rettangolari in ciascun angolo e al centro dei lati della finestra.



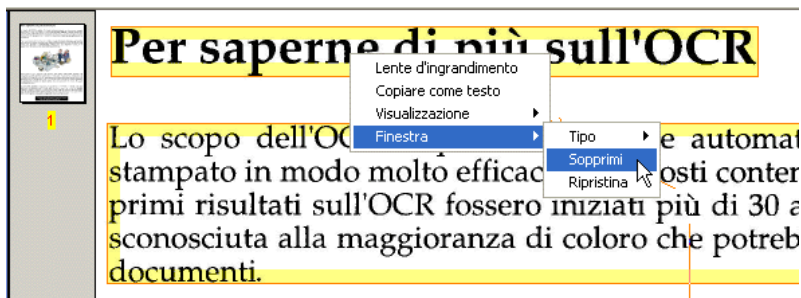
**Per saperne di più sull'OCR**

Per **delesezionare** finestre, fare clic con il pulsante del mouse in qualche altro posto. Per selezionare **ulteriori finestre**, tenere premuto il tasto Mai facendo clic su tali finestre supplementari. Per selezionare una finestra e le **finestre incluse** (dello tipo differente), 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 del mouse su un marcatore e trascinare il lato per cambiare le dimensioni della finestra.

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

Per **eliminare** finestre, selezionatele, cliccate con il lato destro del mouse e scegliete il comando "Finestra - Sopprimi" dal menu "Contesto". In questo modo si cancelleranno tutte le finestre selezionate nonché la finestra sotto il cursore del mouse.



In alternativa, è possibile selezionare zone e scegliere il comando "Taglia" o "Cancella" dal menu "Modifica". Il comando "Taglia" taglia la (le) finestra(e) trasferendola(e) in una memoria temporanea interna, mentre "Cancella" elimina definitivamente la (le) finestra(e). 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 ed 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 cancellare l'ultima operazione.

Modifica	
Ripristina	Ctrl+Z
Taglia	Ctrl+X
Copia	Ctrl+C
Incolla	Ctrl+V
Cancella	Delete
Sopprimi le piccole finestre ...	Ctrl+M
Seleziona tutto	Ctrl+A

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.

## TRE, SALVARE 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".



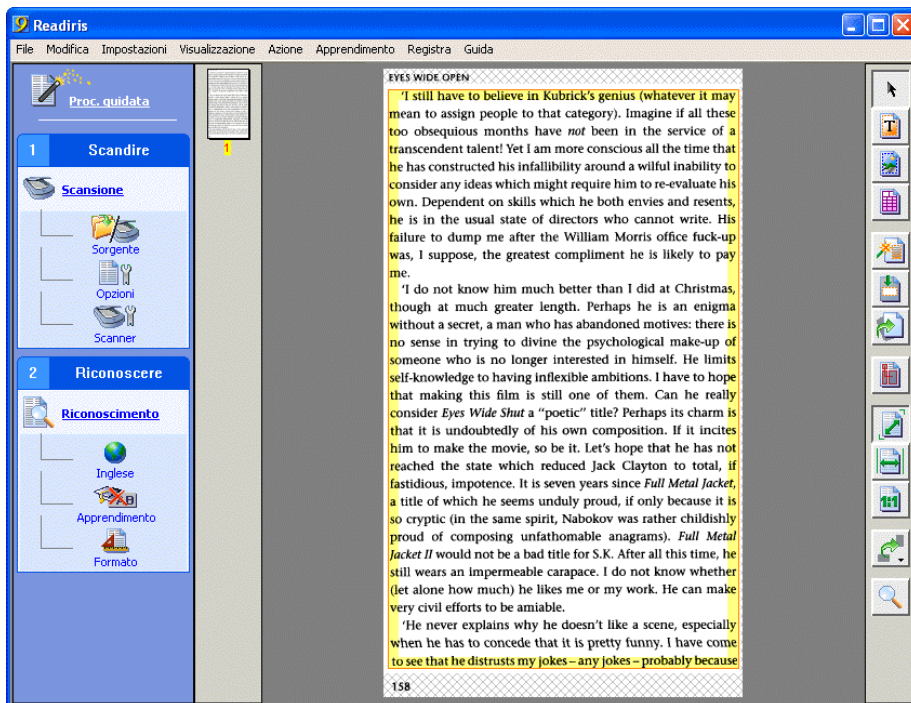
Qualora si renda necessario riconoscere documenti con un layout simile, ad esempio una relazione di 50 pagine in cui, per ovvie ragioni, sarebbe necessario escludere l'intestazione e il piè di pagina, è possibile utilizzare un singolo modello per dividere in zone tutte le 50 pagine.

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.

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



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



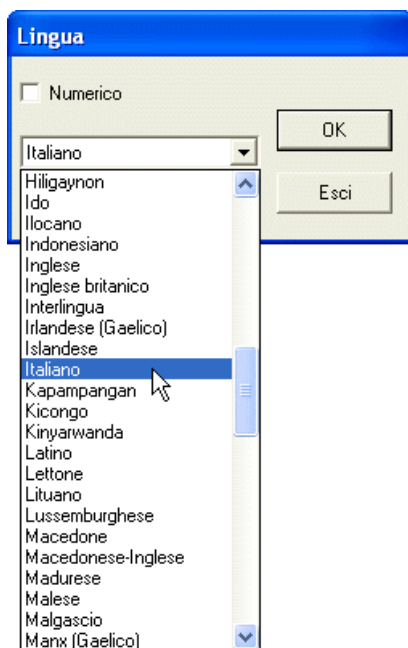
## READIRIS PERMETTE DI VIAGGIARE IN TUTTO IL MONDO

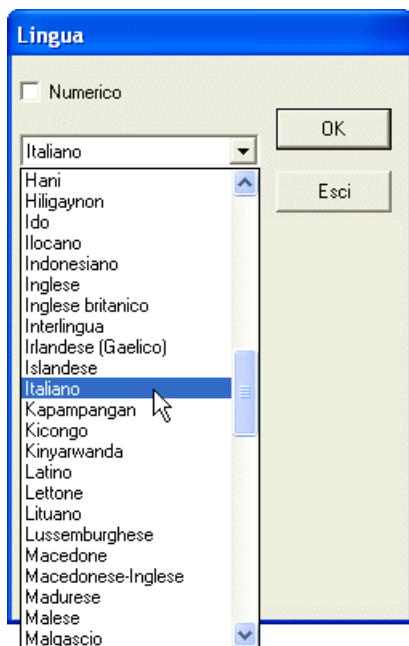
Ammesso che le finestre siano definite correttamente, ora il testo è quasi pronto per il riconoscimento dei caratteri. “Quasi” perché la lingua e le impostazioni del documento non sono ancora state verificate.

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



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





Per accedervi direttamente, premere un tasto lettera: qualora la lingua attualmente selezionata sia l'italiano e si desideri selezionare per esempio il danese, è possibile cliccare sul tasto "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 al occitano. (Per passare a un'altra lettera, ad esempio T, premere il tasto BackSpace prima di digitare il carattere "T".)



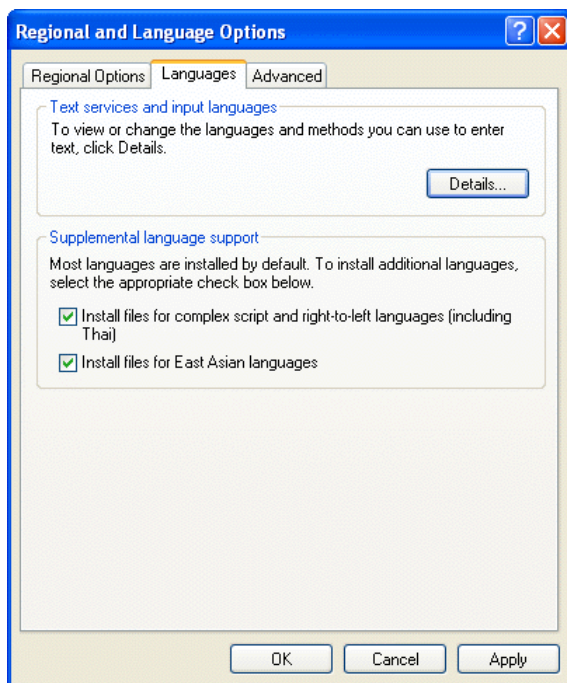
Readiris non si limita all'italiano ma è in grado di gestire fino a 107 **lingue!** Supporta infatti tutte le lingue americane ed europee, incluse le lingue dell'Europa centrale, le lingue cirilliche ("russi"), le lingue baltiche, il greco e il turco.

Potete leggere opzionalmente dei **documenti in caratteri asiatici**: il modulo aggiuntivo "Modulo OCR asiatico" consente il riconoscimento del giapponese, del cinese tradizionale e di quello semplificato, nonché del coreano. (Il cinese semplificato è usato nella Cina continentale e Singapore, mentre il cinese tradizionale è usato a Hong Kong, Taiwan, Macau e dalle comunità cinesi d'oltremare.)

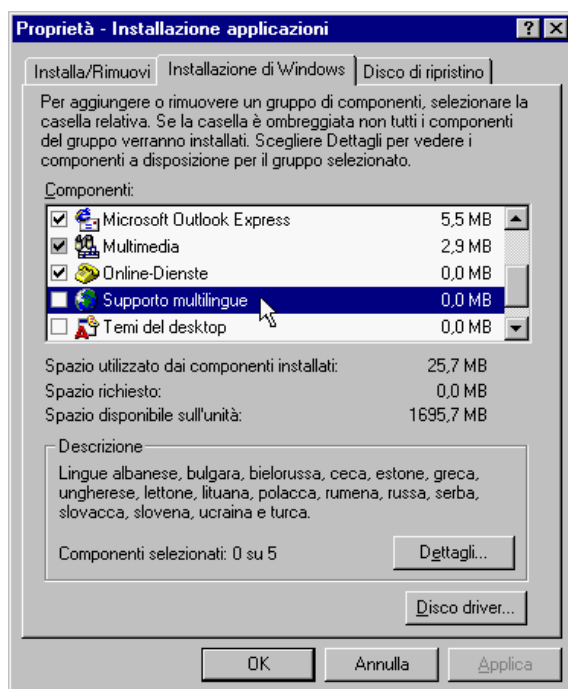
Inoltre l'inglese britannico viene distinto dall'inglese americano - o si dovrebbe dire "internazionale"?

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, selezionate l'icona "Parametri Regionali (e Lingue)" del "Pannello di Controllo".



Su sistemi operativi Windows ME e 98, per verificare che il modulo di Windows "Supporto Multilingue" sia installato nel PC, selezionare l'icona "Aggiungi/Rimuovi Programmi" el "Pannello di Controllo".



Per visualizzare e modificare documenti asiatici, occorre installare una versione asiatica del sistema operativo Windows. Alternativamente, è possibile utilizzare un software “d'emulazione” specializzato (quali ad esempio UnionWay AsianSuite o TwinBridge AsianBridge) su una versione di Windows occidentale per rappresentare correttamente gli ideogrammi di queste lingue asiatiche. È possibile, infine, utilizzare Word 2003, 2002 o 2000 per visualizzare e modificare questi documenti: Office 2003 System, Office XP ed Office 2000 sono stati appositamente studiati per trattare questi documenti in molte lingue diverse.

Consulti il file “**Leggimi**” di Readiris per maggiori informazioni sull'argomento.



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 d'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 lo 'O'. 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 "l". 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

## **SE NECESSARIO, READIRIS CAMBIA LE LINGUE**

---

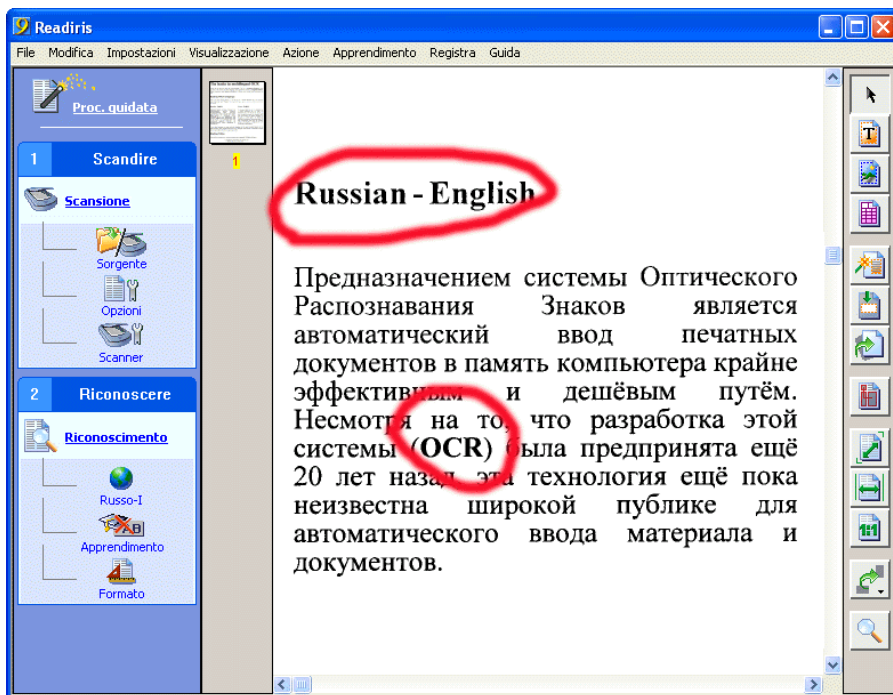
Ma Readiris fa molto di più: Readiris è in grado di commutare tra più lingue all'interno di una frase senza alcun aiuto da parte dell'utente! Quando parole “occidentali” compaiono in documenti in lingua greca, cirillica o asiatica - molti nomi propri non traslitterabili, marche ecc. vengono scritti con comuni simboli occidentali - Readiris è in grado di adattare automaticamente l'alfabeto corretto.

In altre parole, è possibile attivare un **alfabeto misto** di caratteri greci, cirillici o asiatici con caratteri occidentali.

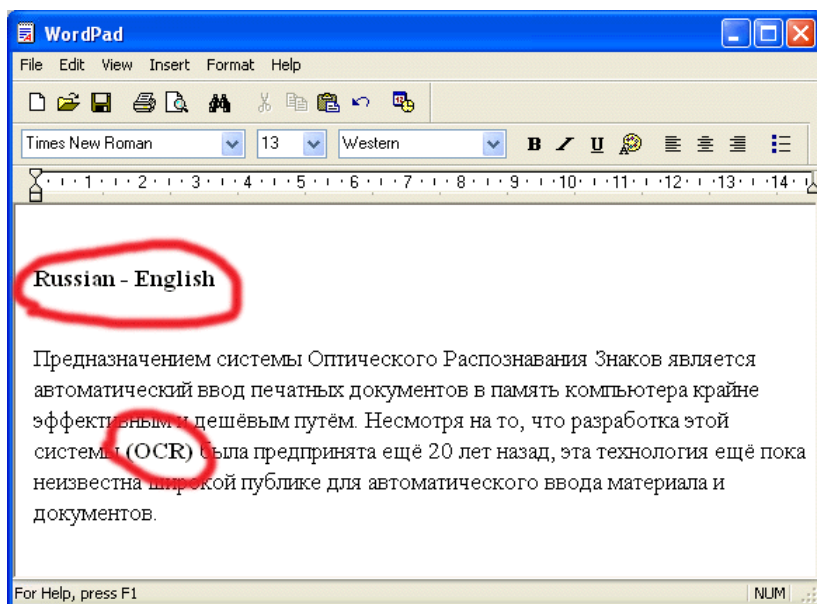
Assicurarsi di selezionare "Greco-Inglese" o la lingua cirillica appropriata - ad esempio "Bielorusso-Inglese". In altre parole: non tentare di selezionare solo "Greco" o "Bielorusso" come lingua del documento sperando che i simboli occidentali risultino chiari!



Segue l'esempio di un testo russo contenente alcune parole inglesi - nel caso in cui si desideri tentare autonomamente, aprire il file d'immagine ALPHABETS.TIF!



Quando viene aperto in un programma di videoscrittura, il risultato finale è il seguente - per visualizzare correttamente il testo russo, può essere necessario selezionare un **font** cirillico.



Per **combinare tra loro altre lingue**, è sufficiente selezionare la lingua che presenta il set di caratteri più esteso. In caso di un documento nel quale ad esempio la traduzione francese è posta accanto al testo inglese, è necessario selezionare la lingua francese per garantire che caratteri quali ç, é e ù vengano riconosciuti correttamente.

## 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 il passo dei caratteri, impostando con precisione il riconoscimento del documento stesso. (Questi comandi non si applica ai documenti asiatici.) Segue ora la spiegazione.



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

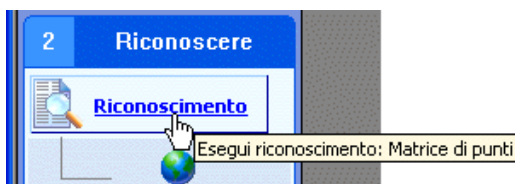
## ape-descended life

La stampa a matrice di punti "qualità corrispondenza", chiamata anche matrice di punti "25 pin", richiede l'impostazione normale, proprio come le **qualità di stampa** di materiale tipografico, dattiloscritto, stampato su laser e a getto d'inchiostro.

L'impostazione "Automatico" vuol significare che Readiris individua il modo di font automaticamente. Lasciate Readiris "auto-individuare" il modo di font in tutte le situazioni - almeno che voi siete sicuri che soltanto i documenti a matrice siano letti. (Ovviamente "Automatico" è il valore di difetto.)



Il tipo di font è indicato nella casella strumenti del tasto "Riconoscimento"; nel caso nessun messaggio è aggiunto nella casella di strumenti, l'autorilevazione della qualità di stampa s'applica, quando il messaggio "Matrice di punti" compare nella casella di strumenti, la lettura per matrici di punti è attivata.



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





Con font *fisso* “monospaziati”, 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 dattilografato.

Con passo *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 ai caratteri “stretti” “l” o “j”. In genere, tutti i libri, riviste e giornali sono stampati con passo proporzionale.

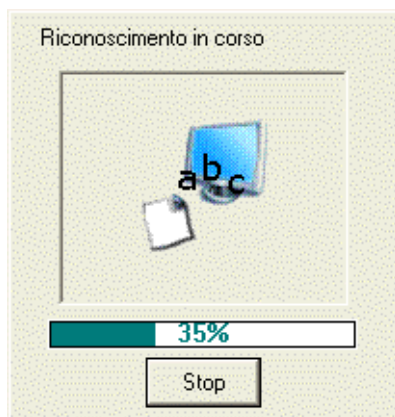
La soluzione più semplice consiste nel lasciare ogni volta questa opzione impostata sul valore di difetto “Automatico”, in tal modo Readiris individuerà automaticamente il passo dei caratteri.

## READIRIS DIVENTA OGNI VOLTA PIÙ INTELLIGENTE!

Una volta selezionata la lingua del documento e impostate le caratteristiche del documento, attivi l'apprendimento interattivo e clicca il tasto “Riconoscimento”.



L'avanzamento del riconoscimento viene indicata sullo schermo. Per interrompere il riconoscimento del testo, fare clic sul pulsante “Stop”.



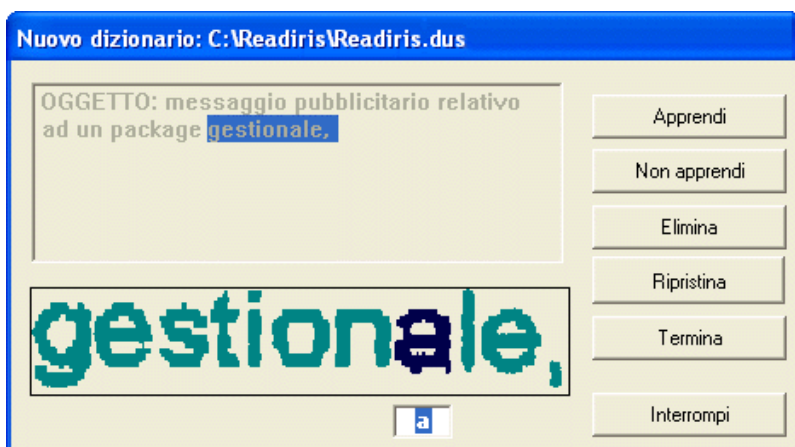
Al termine del riconoscimento, Readiris passa alla fase di apprendimento interattivo in cui l'apprendimento è abilitato dal pulsante "Apprendimento" della barra strumenti principale.

(L'apprendimento interattivo non si applica in documenti asiatici: l'apprendimento non ha senso per queste lingue che usano migliaia di simboli differenti - e di più dovete essere capace di introdurre "l'ideogrammi", cosa non facile con delle tastiere occidentali!)

L'**addestramento ai 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'addestramento consente di superare questo temporaneo "insuccesso".

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



Verifichi innanzitutto di aver attivato il dizionario di font e la modalità di dizionario appropriati - sempre indicati nel titolo della finestra di apprendimento. In caso contrario, fate clic sul pulsante "Interrompi" - viene così rivisualizzata l'immagine del documento con la divisione in zone iniziale -, abilita 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, sostituisca la forma errata o sconosciuta con un carattere (o stringa di caratteri) e fate clic su uno dei seguenti pulsanti.

### **Apprendi**

Accorda la soluzione proposta o correggerla. Il programma salva questo carattere dubbio nel dizionario di font come "certo", definitivo. Per un



riconoscimento successivo non verrà più richiesto l'intervento dell'utente, in quanto la forma viene considerata definitivamente appresa.

Nell'esempio sopra indicato, 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

Accorda 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 ad 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 sui documenti - chiazze, macchie di caffè ecc. - che potrebbero essere riconosciute come punti, virgole ecc. - e per cancellare ogni altro simbolo non desiderato.

### **Ripristina**

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 uscita si ricomincia tutto da capo, con "Termina" si crea un testo che però non viene controllato dettagliatamente!

## **IL RUOLO DEI DIZIONARI DEI FONT**

---

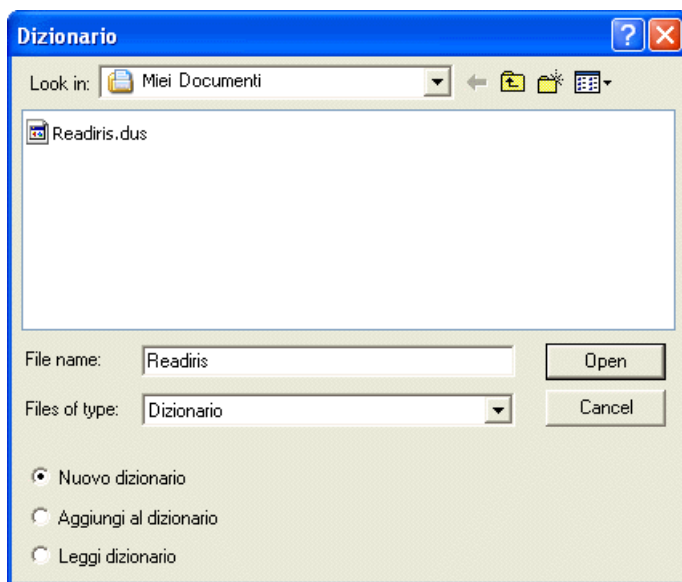
I risultati di ciascuna sessione di addestramento vengono conservati temporaneamente nella memoria del computer ma, per impieghi successivi, è possibile e necessario memorizzarli in file chiamati "dizionari dei font".

(Non confondete i dizionari dei font con i lessici (d'utente)! I dizionari dei font contengono forme carattere apprese durante la fase OCR interattiva, i lessici sono database linguistici di ausilio nel riconoscimento.)



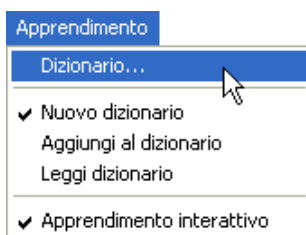
Quando si desidera riconoscere documenti simili, è necessario caricare in memoria questi dizionari dei font per utilizzare l'intelligenza supplementare che contengono. In questo modo Readiris tiene conto dell'intelligenza memorizzata in queste biblioteche dei font. Si potrebbe affermare che ad ogni impiego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

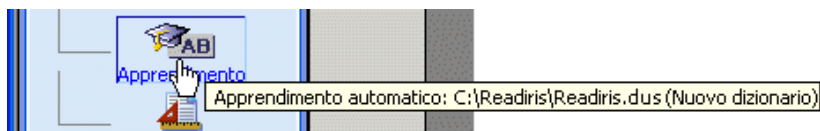
Con la modalità "aggiungi" i risultati del processo di apprendimento verranno salvati in un dizionario *esistente*: il riconoscimento si avvale di un'intelligenza supplementare già presente nel dizionario, a cui vengono aggiunte nuove forme di font. In altre parole, questa opzione consente di creare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

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

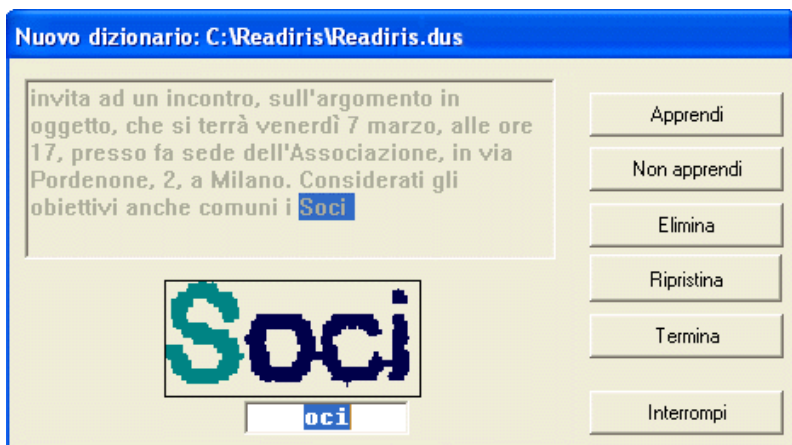
Con l'ultima opzione, "Leggi Dizionario", il dizionario funziona nella modalità di sola lettura: si utilizza il dizionario senza aggiungervi *nuove* forme di font.

Selezionare la modalità nuovo quando si esegue il riconoscimento di una sola pagina. Per riconoscere numerose pagine dello stesso tipo - 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 (di) documento(i).

La casella informativa del pulsante "Apprendimento" indica sempre quale dizionario di font è attualmente attivo e in quale modalità opera.



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.



## INVIARE IL RISULTATO ALL'APPLICAZIONE

L'addestramento interattivo pone termine al riconoscimento dei caratteri. Perché Microsoft Word agisce per difetto come applicazione destinataria, il sistema di videoscrittura 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.)



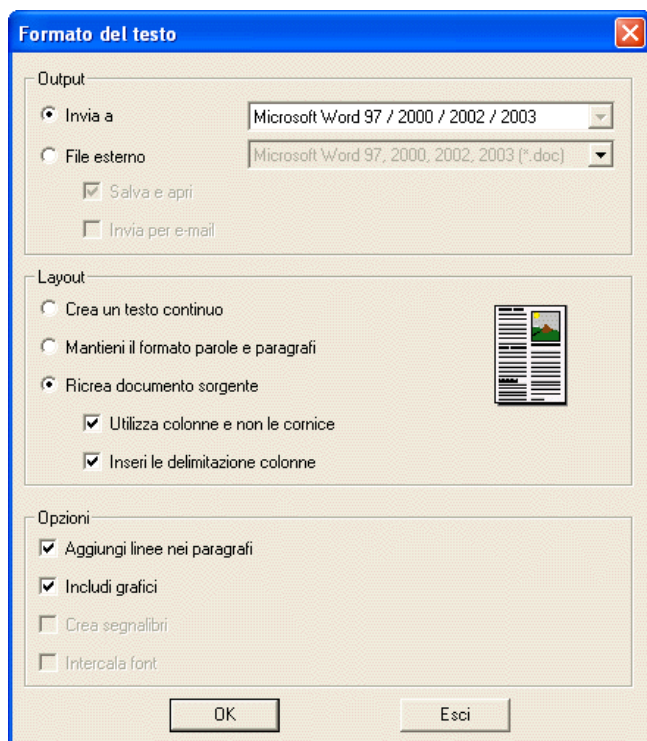


L'immagine scandita viene nuovamente visualizzata con la divisione in zone creata, per essere disponibile per ulteriori elaborazioni, restando visualizzata fino a quando non si esegue la scansione di un'altra pagina.

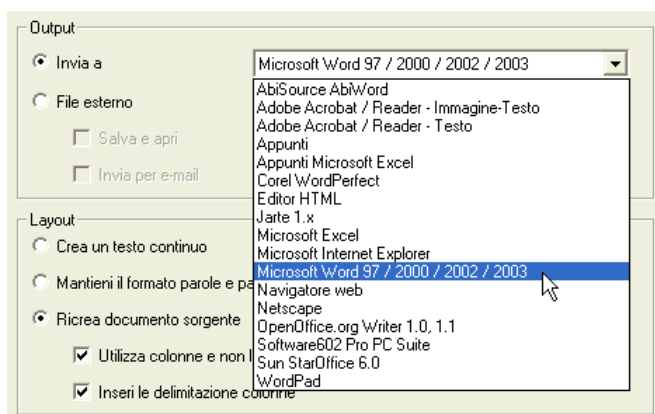
In questo modo si è convertito un documento cartaceo in un file editabile, impiegando tempi fino a 40 volte inferiori rispetto a quelli necessari per la ribattitura manuale! Procedere e confrontare il testo con l'immagine posta all'interno della finestra di Readiris.

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

Utilizzando il tasto "Formato" sulla barra strumenti principale (o il comando "Formato Testo" nel menu "Impostazioni") viene selezionato il **target output**.



La funzione "Invia a" consente un collegamento diretto OCR fra lo scanner e le applicazioni Windows: **inviare** i documenti scansionati direttamente al sistema di videoscrittura, al foglio elettronico, o al navigatore web, ad Adobe Reader ecc.!



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



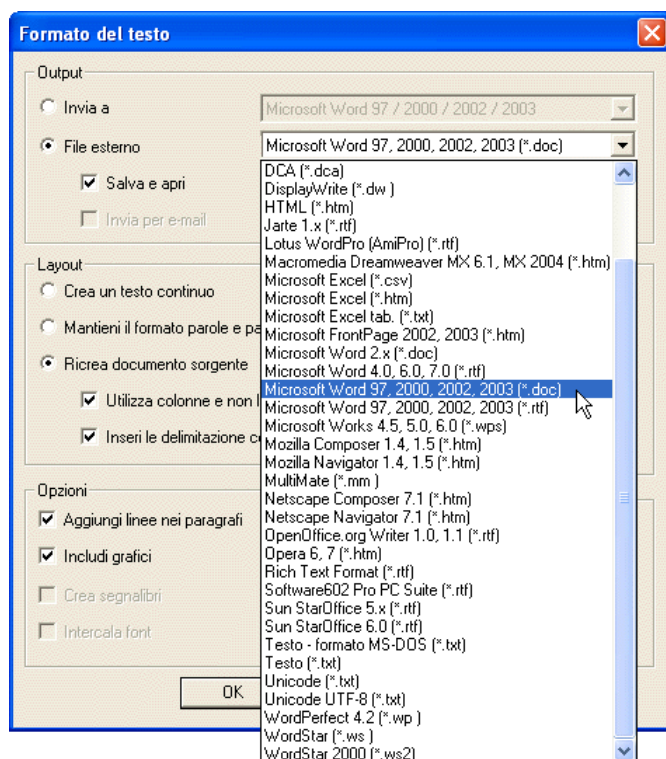


Non dimenticare che l'opzione "Invia a" consente anche di copiare il testo riconosciuto negli **appunti** di Windows. Non vi è quindi alcuna necessità di esportare il risultato... o di salvarlo in un file esterno!

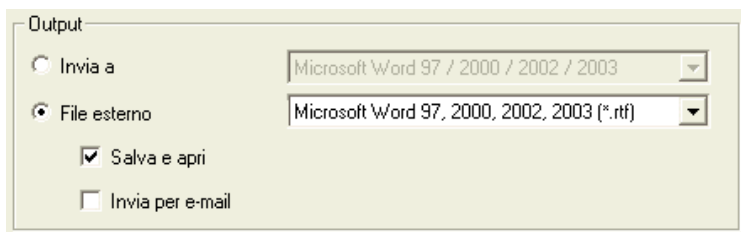
## **SALVARE I RISULTATI IN UN FILE DI TESTO**

---

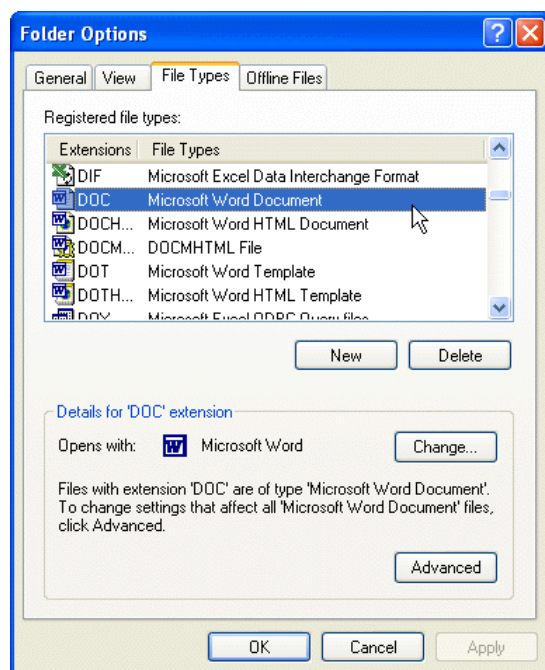
È possibile scrivere il risultato OCR in un file “esterno”. Di nuovo, Readiris supporta un'ampia gamma di formati file comprendenti tutti i sistemi di videoscrittura, fogli elettronici, applicazioni web ecc.



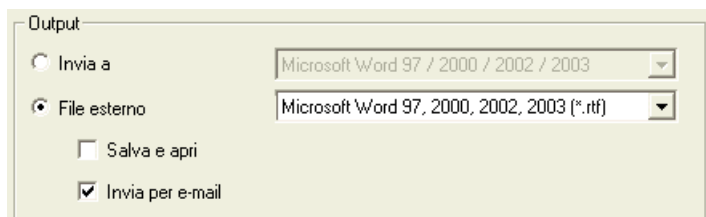
L'opzione "Salva e Apri" è simile alla funzione "Invia a": si apre il documento riconosciuto una volta salvato.

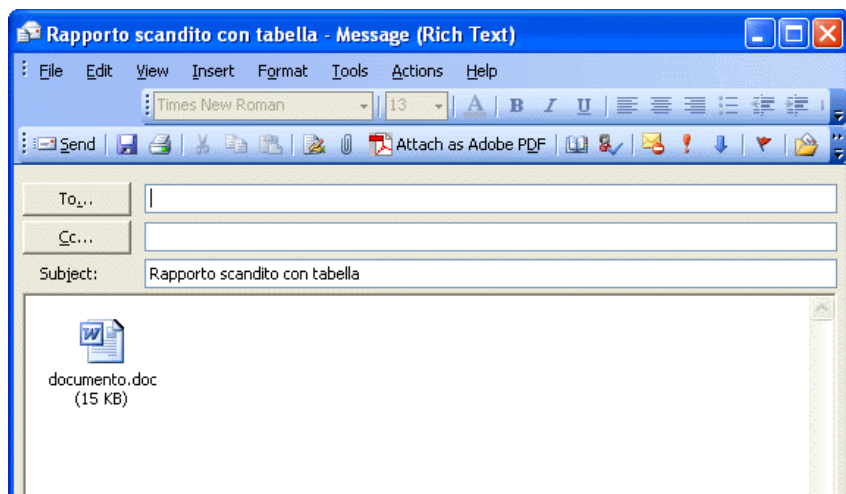


Tuttavia il metodo utilizzato per inviare le applicazioni target è diverso. In questo caso i **tipi file di Windows** stabiliscono le applicazioni da avviare. È come se si cliccasse due volte sul file di uscita in Esplora Risorse di Windows... (Con l'opzione "Invia a", Readiris invia direttamente applicazioni target specifiche.)



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

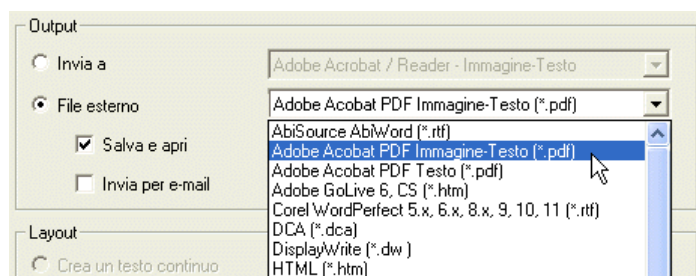
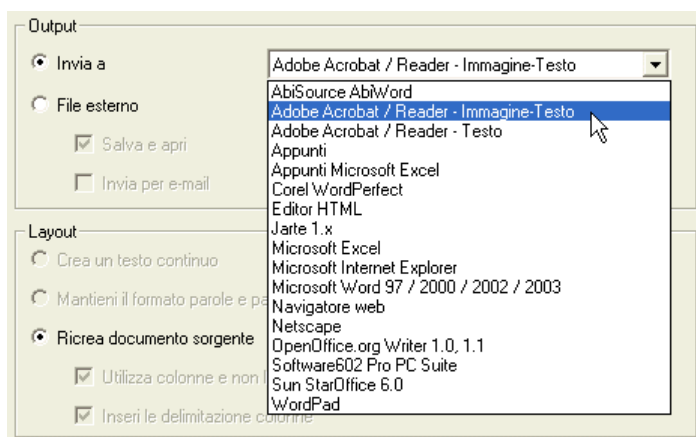




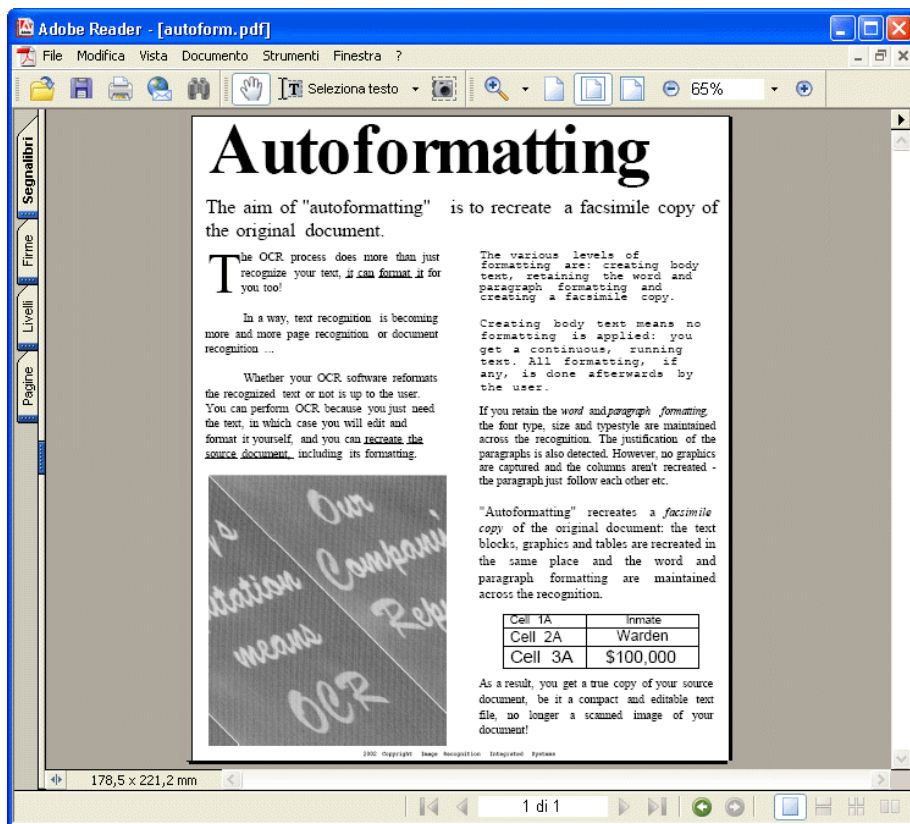
## PER CREARE UN DOCUMENTO ESPORTABILE

Esaminiamo più nel dettaglio un formato: **Adobe Acrobat PDF**. Readiris vi consente di creare documenti PDF di due tipi - PDF Testo e PDF Immagine-Testo.

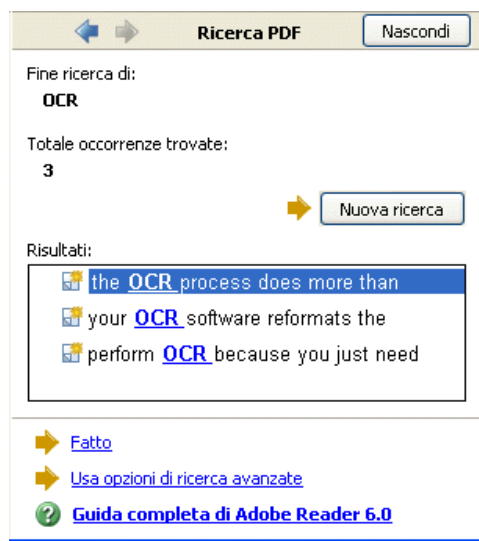




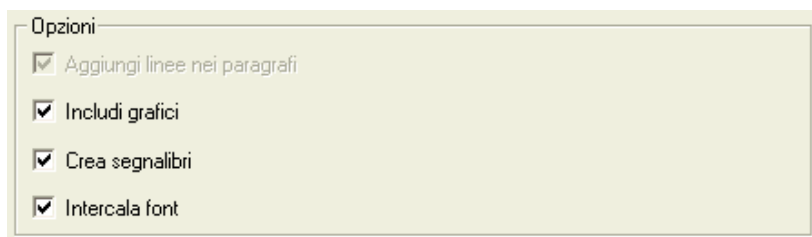
Quale è la differenza fra i due? Quando selezionate il formato "PDF Testo", Readiris crea un archivio PDF che contiene il risultato del testo. (Può anche esserci grafica ma solo quando zone di grafica sono presenti nella pagina - fotografie, materiale illustrativo ecc.) In altre parole: l'immagine della pagina *non* è contenuta nel file PDF mostrato!



Il formato "PDF immagine-Testo" produce un risultato diverso. Readiris crea un file ricercabile PDF che contiene il testo riconosciuto e l'immagine pagina. Questa è contenuta sopra il testo in un file PDF a doppio strato. Usate la funzione di ricerca dell'Adobe Reader e tutto diventerà semplice!



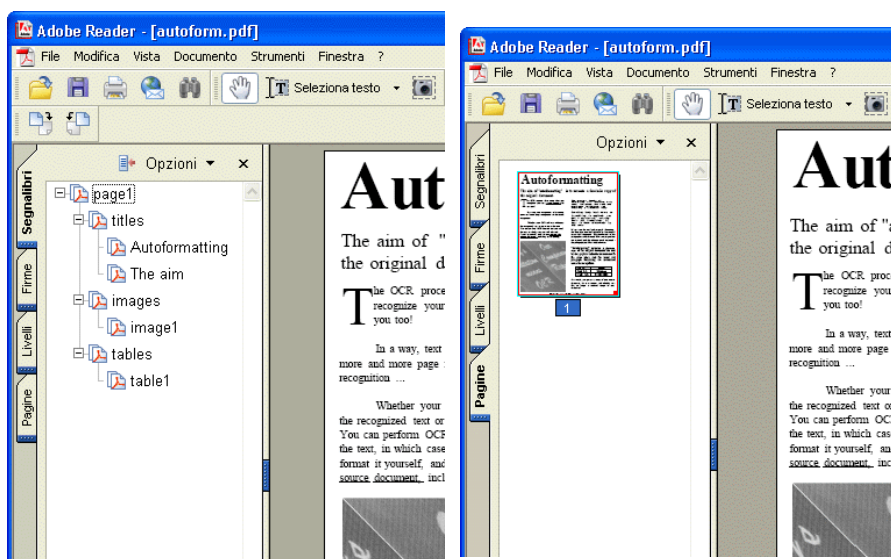
Cliccate il tasto "Formato" per scoprire due opzioni che riguardano il formato Acrobat PDF: "Crea Segnalibri" e "Intercala Font".



La prima opzione fa sì che un **segnalibro** sia creato per ogni elemento di documento - per i grafici come per i blocchi di testo e le tabelle. Per le zone di testo, Readiris applica un algoritmo intelligente per esprimere un titolo, un "sommario" per ciascuna zona; le tabelle e i grafici sono semplicemente numerati.



(È possibile creare dinamicamente con il software Adobe Reader un altro elemento della navigazione dei documenti PDF, le **miniature**.)



L'opzione "Intercala Font" intercala i font nei file PDF. Intercalando 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, intercalando i font, si aumenta (leggermente) la dimensione del file dei documenti riconosciuti!

## ... O LEGGERLI

Consideriamo anche l'altro aspetto. 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 maggiori applicazioni di Readiris. Diversi sono i motivi. Prima di tutto, si tratta di un modo di convertire immagini in testo: aprite i documenti PDF con le immagini, eseguite il riconoscimento e salvate il risultato OCR in un documento testo (in qualsiasi formato testo supportato). I file di testo sono editabili, i file immagini non lo sono!

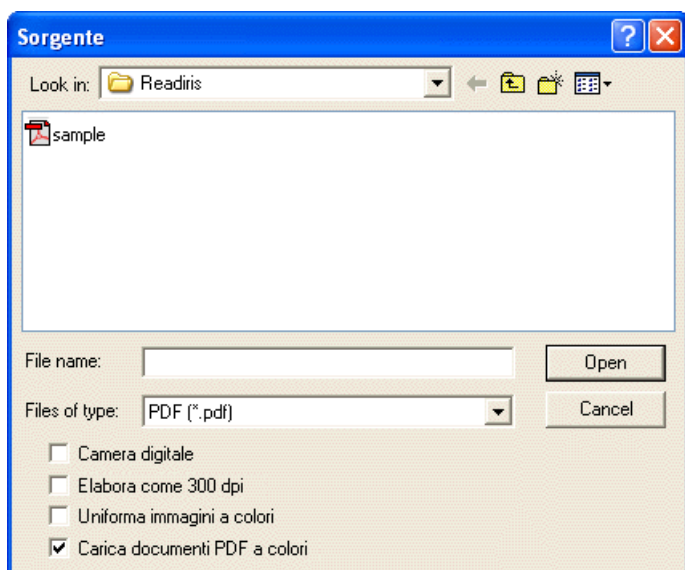
Secondo caso: è possibile convertire file PDF con immagini in documenti PDF di testo. Si esegue in tal caso il riconoscimento sui file PDF “solo immagini” e si salvano i risultati OCR... come documenti PDF di testo. I file PDF di testo sono ricercabili e modificabili, i file PDF... “solo immagini” non 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), con i file di “sola lettura”, il contenuto non può essere estratto. Questi documenti possono essere solo visualizzati e stampati!

Una caratteristica importante: Readiris non apre documenti PDF protetti con password anche se tutte le altre barriere di sicurezza vengono superate con Readiris!

Procedete come di solito: memorizzate i file PDF aprendo le immagini prescandite - fax, istantanee con macchina digitale ecc. Cliccate il tasto "Stop" o premete Escape per interrompere il processo di caricamento fra due pagine.

Vi è 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!



## RICONOSCIMENTO DI PAGINE MULTIPLE

---

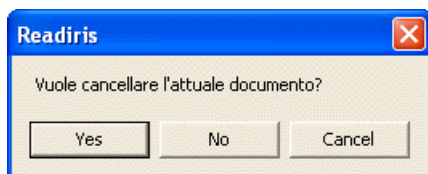
Dopo l'OCR l'immagine scandita viene rivisualizzata con la suddivisione in zone creata per la successiva elaborazione.

Ora 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.

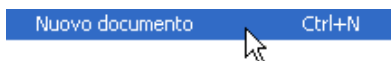
Ma come salvare il testo di pagine successive? O in altri termini: come elaborare documenti costituiti da molteplici pagine? In realtà è molto semplice: proseguite con il riconoscimento delle pagine e salvate i risultati nello stesso file! (Assicuratevi che il file non sia aperto al momento, poiché in tal caso non sarebbe

possibile scrivervi!) In secondo luogo ricordarsi di impostare il dizionario di font in modalità appendi, così da potere aggiungere il testo e continuare comodamente l'addestramento ai font.

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

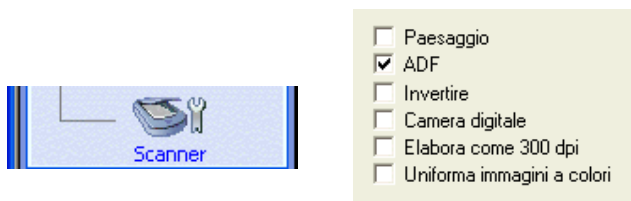


Rispondere "No" per aggiungere le pagine al documento corrente, rispondere "Sì" per creare un nuovo documento. Questa risposta ha lo stesso effetto del comando "Nuovo Documento" nel menu "File".



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

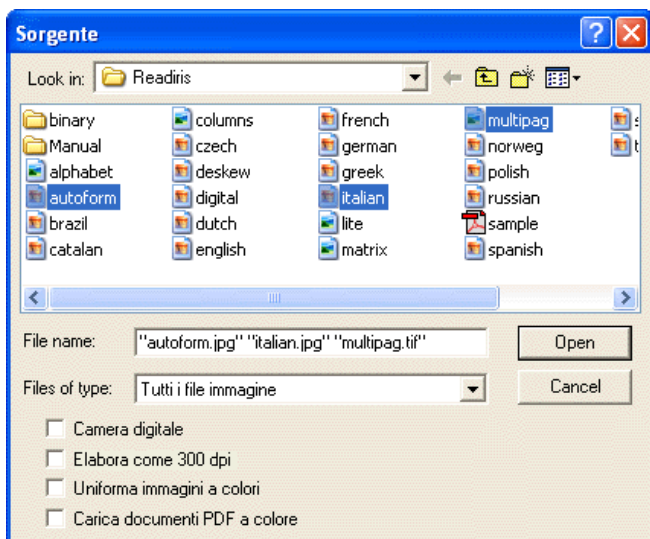
Per scandire un documento costituito da diverse pagine in un'unica operazione, attivare l'alimentatore fogli dello scanner con l'opzione "ADF" sotto il tasto "Scanner".





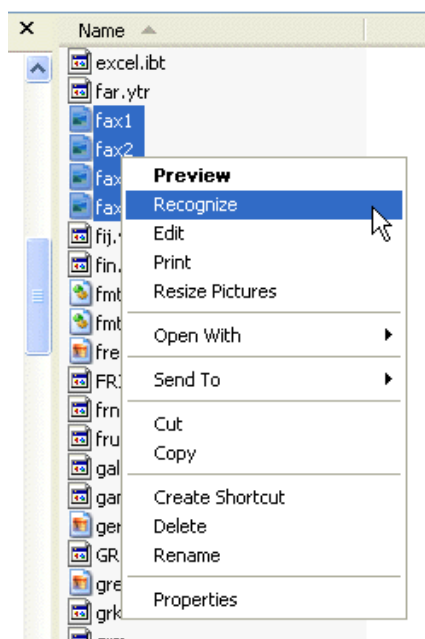
Inserire le pagine del documento nell'alimentatore fogli automatico ed iniziare la scansione: tutte le pagine vengono scandite fin quando l'alimentatore fogli risulta vuoto.

È anche possibile *aprire* immagini prescandite multiple. Per caricare diverse immagini, selezionare la prima immagine e tenere premuto il tasto Ctrl durante la selezione di immagini aggiuntive. Per caricare una gamma continua d'immagini, selezionare la prima immagine e tenere premuto il tasto Mai durante la selezione dell'ultima immagine.



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

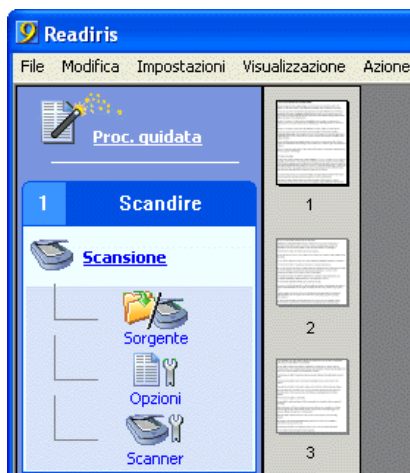




È possibile anche *trascinare* diverse immagini prescandite da Esplora Risorse di Windows alla finestra di Readiris! Vale lo stesso argomento: tutte le immagini trascinate sulla finestra Readiris vengono aggiunte al documento corrente fino a quando si cliccherà sul comando "Nuovo Documento".

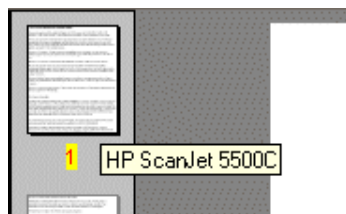
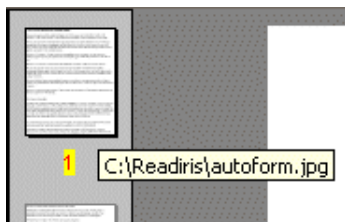
Readiris ordina le immagini automaticamente: l'immagine 001.tif precede 002.tif che precede 003.tif ecc.

La **barra strumenti pagina** che compare sul lato sinistro - visualizzata durante l'elaborazione delle pagine - rappresenta le diverse pagine del documento e consente di accedere ai comandi della pagina (cliccando a destra).



La pagina attuale viene evidenziata nella barra strumenti pagina e menzionata nella barra titoli di Readiris.

La barra strumenti della pagina supporta una casella informativa: puntare il mouse sulla miniatura della pagina per sapere quale immagine è stata caricata in memoria. (Se è stata aperta un'immagine multipagina, esiste ovviamente un unico file per tutte le immagini.) Durante la *scansione* di documenti multipagina, la casella informativa menziona semplicemente il modello dello scanner.



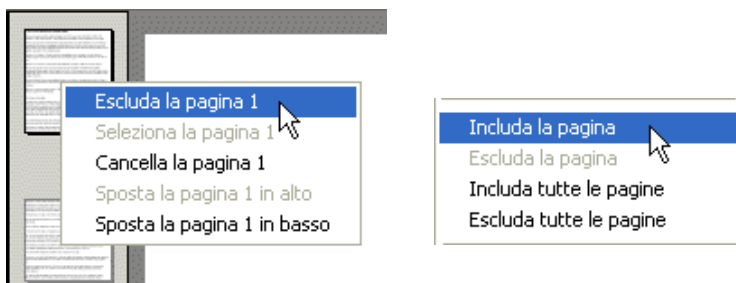
Voi potete velocemente stampare le immagini scandite con il comando "Stampa Immagine" nel menu "File" se volete una visione d'insieme del vostro documento.



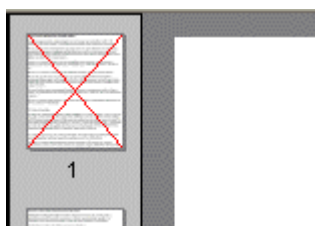
Potete stampare la pagina corrente o tutte le pagine. Selezionate il numero di pagine o miniature che desiderate stampare su una pagina.



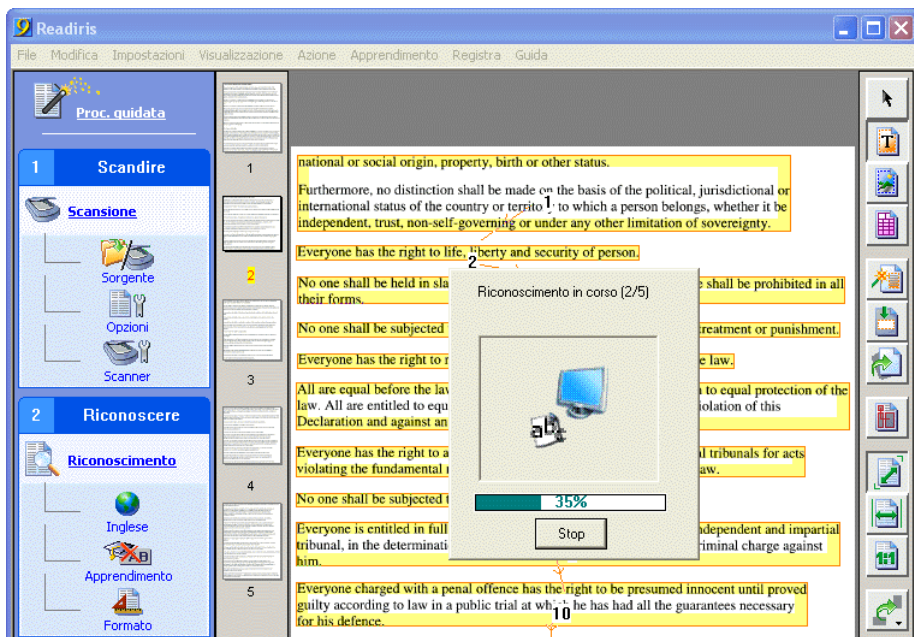
Ma non è necessario stampare tutte le pagine; la barra strumenti della pagina (e il menu "Modifica") vi consentono di escludere (temporaneamente) delle pagine. Cliccate con il lato destro del mouse su una "miniatura" pagina e selezionate il comando "Escluda la Pagina" nel menu "Contesto" o visualizzate una pagina e selezionate il comando "Escluda la Pagina" dal menu "Modifica" per escluderla dal processo di stampa (e riconoscimento). Selezionate il comando "Includa la Pagina" per includerla di nuovo. Per maggiore flessibilità, il menu "Modifica" offre comandi equivalenti che si applicano a *tutte* le pagine.



Le miniature (“thumbnails”) delle pagine esclusi vengono sbarrate. Ricordate che è sempre possibile stampare la pagina corrente - anche se è “disattivata” per il momento.



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



Se l'apprendimento interattivo è attivato, le fasi di riconoscimento ed apprendimento sono eseguite pagina per pagina. Il modo dizionario "nuovo" viene utilizzato per la prima pagina mentre il modo "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 è annullato per *tutte* le pagine; l'OCR di questo documento continua nel modo automatico.

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



Perlomeno, questo è il caso quando l'opzione "Crea Un Documento per Pagina" è disattivata quando si salva il documento riconosciuto. Questa opzione prevede che ogni pagina di un documento multipagina sia salvata in un archivio separato. Se l'utente dà il nome al file testo.doc, i file saranno chiamati testo-1.doc, testo-2.doc ecc.

File name:

Save as type:

☐ Salva e apri

☐ Invia per e-mail

☒ Crea un documento per pagina  
(Multipage-1, Multipage-2...)

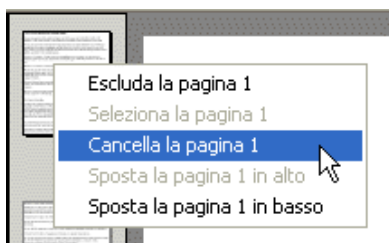
## MODIFICA DI DOCUMENTI MULTIPAGINA

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

Dapprima la navigazione. Per *spostarsi su una pagina*, cliccare sulla sua miniatura nella barra strumenti della pagina o tenere il cursore sulla sua icona, richiamare il menu "Contesto" cliccando a destra e utilizzare il comando "Seleziona la Pagina". Per andare alla pagina precedente, premete PageUp, per andare alla pagina successiva, premete PageDn. Premete Home per andare alla prima pagina, premete End per andare all'ultima pagina. O utilizza i comandi corrispondenti del menu "Visualizzazione".

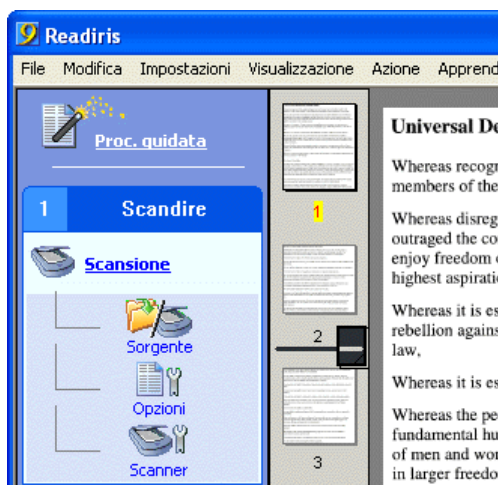
Prima pagina	Home
Pagina precedente	PageUp
Pagina successiva	PageDown
Ultima pagina	End

Modifichiamo ora il documento. Per *cancellare una pagina* dal documento, tenere il cursore sulla sua miniatura, cliccare a destra ed utilizzare il comando "Cancella la Pagina". Ricordiamo che potete escludere temporaneamente le pagine - non *cancellarle* - dal processo di riconoscimento (e stampa immagine): la barra strumenti della pagina (e il menu "Modifica") offrono i comandi necessari.



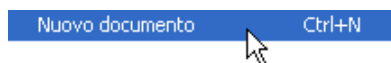
Per *spostarsi alla pagina precedente*, utilizzare il comando "Sposta la Pagina in Alto" e per *spostarsi alla pagina successiva*, utilizzare il comando "Sposta la Pagina in Basso".

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" nel menu "File" per chiudere il documento corrente.



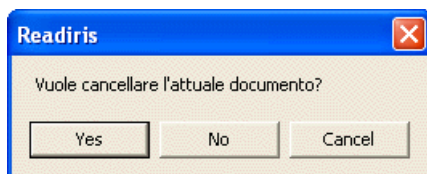
Questo comando "cancella la lavagna". Qualsiasi documento caricato in memoria - contenente una pagina singola o pagine multiple - viene cancellato. A questo punto è possibile creare un nuovo documento.

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 scansionare tutte le pagine contenute nell'alimentatore automatico dello scanner, riempire nuovamente l'alimentatore e ripartire. Tutte le pagine scandite comporranno un singolo documento. In alternativa è possibile scansionare una serie di pagine e



aggiungere alcuni file di immagine, ad esempio fax. Queste pagine formano a loro volta un singolo documento. Tutto quello che occorre fare è cambiare la sorgente di immagine utilizzando il tasto "Sorgente".

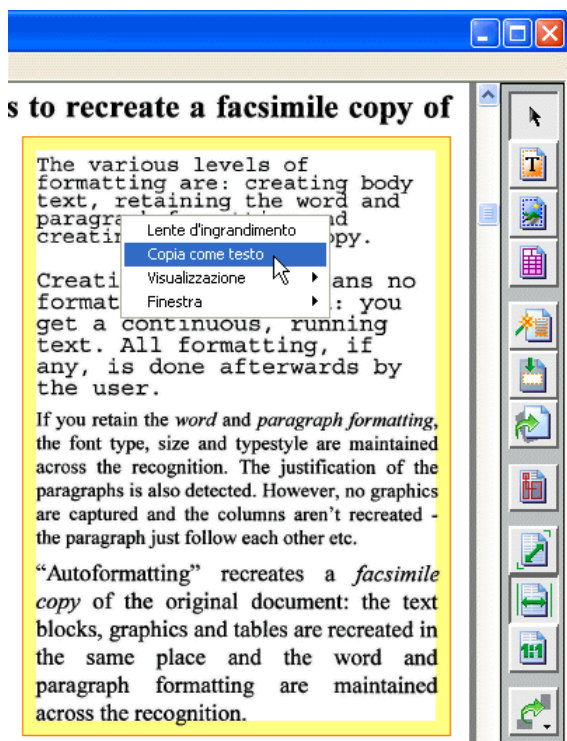
Se una volta *ultimato* l'OCR viene riavviata la scansione (o il caricamento delle immagini), il sistema chiede di creare un nuovo documento o di completare il documento corrente.



## **RICONOSCIMENTO DI ZONE DI TESTO**

---

Sappiamo ora come riconoscere le pagine e come elaborare un documento multipagina. Ma possiamo riconoscere con la stessa facilità meno di una pagina? È possibile! Cliccate con il pulsante destro del mouse e selezionate il comando "Copia come Testo" dal menu "Contesto": la finestra di testo sotto il mouse viene riconosciuta e inviata negli appunti.



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

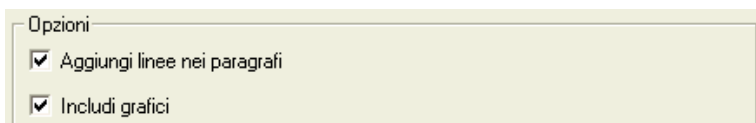
## ORGANIZZAZIONE DELL'USCITA DI TESTO

Salvare o esportare il testo significa molto più che selezionare un metodo di output o assegnare un nome di file al file d'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 dove si desidera prendere il testo.

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

Il **flusso di testo** del documento d'uscita è direttamente influenzato dall'opzione "Aggiungi Linee in Paragrafi".



Per fare in modo che Readiris individui i paragrafi, mantenere questa opzione abilitata: Readiris applicherà così il normale **ritorno a capo automatico** tipico dei programmi di videoscrittura, altrimenti dopo ciascuna riga viene aggiunto un ritorno del carrello e le parole separate da trattino rimangono tali! L'individuazione dei paragrafi è abilitata per difetto.

Segue un esempio chiarificativo. 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 da trattino della parola "presidente" sono state "re-incollate" ed è stato aggiunto uno spazio alla fine della frase, creando un testo fluente.

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

L'opzione *non* è valida quando il formato PDF è selezionato: i file Adobe Acrobat PDF archiviano sempre il testo linea per linea.

(Il pulsante "Formato" contiene alcune opzioni di formattazione non ancora trattate - l'argomento verrà trattato tra breve.)



## INSTALLAZIONE DELLO SCANNER

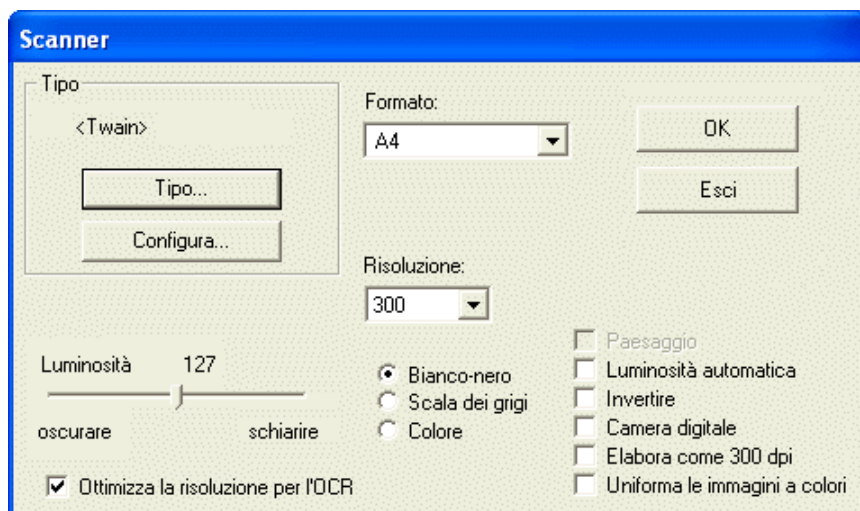
Si proceda ora all'installazione dello scanner. Si presume che l'hardware dello scanner ed i piloti necessari siano installati correttamente.

Nel caso in cui la licenza del software Readiris sia allegata ad 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.

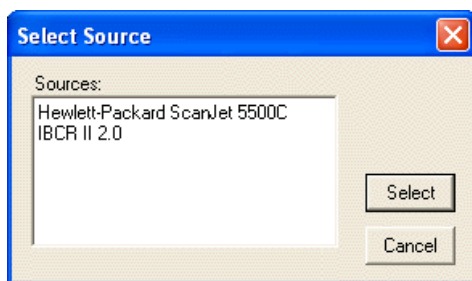


Per determinare il **modello dello scanner** in dotazione, fare clic sul pulsante "Tipo".



Quando si seleziona "<Image>" (immagine) come "scanner", le immagini prescandite funzionano sempre come fonte di immagini - non è neppure necessario selezionare il disco come fonte d'immagini sulla barra strumenti principale con il botone "Sorgente".

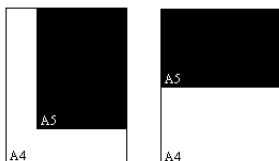
Il pulsante "Configura" è disponibile solo quando lo scanner ne permette l'utilizzo. Tale pulsante consente l'accesso a parametri di scansione avanzati; con gli scanner Twain, facendo clic sul pulsante "Configura" è possibile selezionare l'origine Twain. (È possibile utilizzare anche il comando "Seleziona Origine" del 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, e può permettere d'indicare se sarà utilizzato il sistema di lettura automatica di documenti dello scanner o no. Con scanner Twain compatibili, tutti i parametri di scansione vengono spesso impostati all'interno dell'interfaccia utente.

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

Abilitando l'opzione "Paesaggio", la pagina selezionata sarà orientata in senso orizzontale, anziché verticale. In realtà, l'orientamento della pagina viene applicato ai formati di pagina ridotti: con uno scanner da tavolo 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 piena A4 solo in una direzione!



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

## COLORATE LE VOSTRE SCANSIONI!

---

Readiris supporta su uguale base immagini in bianco e nero, nella scala dei grigi e a colori, e pertanto si è liberi di scegliere la **modalità colore** che meglio risponde alle proprie esigenze. Per includere grafici al tratto nei documenti riconosciuti, scansionare in bianco e nero, per includere fotografie in bianco e nero scansionare nella scala dei grigi e per includere immagini a colori scansionare a colori.

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

La scansione, con scala dei 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 le 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 un "rumore" che rende 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 potete fare per regolare la luminosità dello scanner...

**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.

**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.

Readiris crea una versione in bianco e nero per ogni immagine nella scala dei 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 scandita in bianco e nero deselezionare l'opzione "Visualizza Documento a Colori" nel menu "Visualizzazione".

✓ Visualizza documento a colori Ctrl+O





Quando questa opzione è attivata, non si vedranno immagini in bianco e nero sul video, anche se scandite immagini a due livelli. Questo succede perché l'opzione "Visualizzazione Ottimale" del menu "Visualizzazione" ottimizza le immagini per una leggibilità video ottimale.

✓ Visualizzazione ottimale



Questa particolare tecnica di visualizzazione ad alta risoluzione converte le immagini in bianco e nero in immagini in scala dei 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 dei 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 vi è alcuna necessità di zoomare, anche nei laptop con schermo LCD o computer desktop con video di 14". La visualizzazione ad alta qualità è attiva per difetto ma può essere superflua negli schermi ad alta risoluzione.

## APPARECCHI DIVERSI, DIVERSE RISOLUZIONI

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

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

6 punti

72 punti

Readiris riconosce inoltre “**capolettere**”, ampie maiuscole che coprono più righe. (Naturalmente queste non possono essere di dimensioni superiori a 72 punti!)



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

Dal momento che 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. Tuttavia, Readiris può correggere scansioni con troppi dettagli! Attivare a tale scopo l'opzione "Ottimizza la Risoluzione per l'OCR" dalle impostazioni di scansione. Se la risoluzione dell'immagine delle scansioni supera i 600 dpi, la risoluzione viene ridotta per il processo OCR.

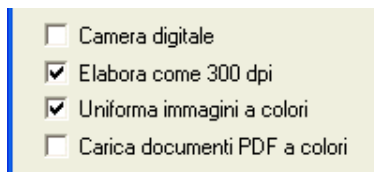
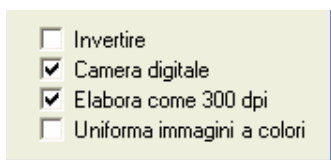
☒ Ottimizza la risoluzione per l'OCR

Vi sono due modi per evitare questo avvertimento: quando si leggono **fax** (con una risoluzione di 100 o 200 dpi), quando si creano immagini con una macchina fotografica digitale (dove la risoluzione non è nota) e quando si aprono immagini in cui l'intestazione del file contiene una risoluzione errata. Per processare queste immagini senza problemi, utilizzate l'opzione "Elabora come 300 dpi". Questa impostazione va bene sia per la scansione diretta sia per l'apertura di immagini già scandite.

☐ Invertire  
☐ Camera digitale  
☒ Elabora come 300 dpi  
☐ Uniforma immagini a colori

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

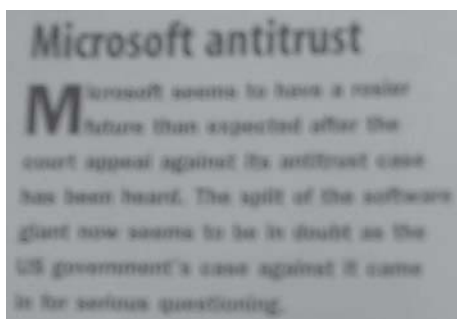
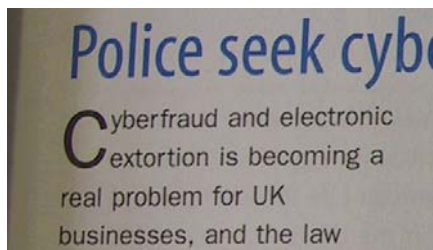
Se le vostre immagini sono acquisite da una **telecamera digitale** invece che mediante uno scanner, è tassativo che attivate una speciale opzione. Questo parametro è di nuovo valido per la scansione diretta e per le immagini già scanndite.



Operando in questo modo, intensificate l'immagine prima di riconoscerla. Ci sono specifici aspetti da tener presenti se si tratta di telecamere digitali: esse producono immagini a bassa risoluzione - anche se tenete la telecamera molto vicina ai documenti - e la risoluzione dell'immagine è in ogni caso sconosciuta.

Ci sono altresì altri aspetti più delicati da prendere in conto per riconoscere con successo le immagini riprese con una telecamera digitale.

Prima di tutto, selezionate la risoluzione di immagine più alta possibile. Create per 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 andate in modalità "macro" per riprendere primi piani - che è sempre il caso di fare quando si fotografano documenti. (Questa modalità è stata studiata ad hoc per la ripresa di fiori, insetti ecc). Altrimenti otterrete immagini non nette e illeggibili.



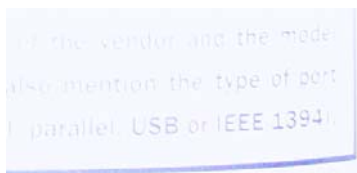
Attenzione alla compressione, che dovrà essere limitata o nulla. Una compressione importante riduce la nitidezza del testo ripreso. Zoomate a mano per riprendere il vostro documento - alcune telecamere sono fornite di software



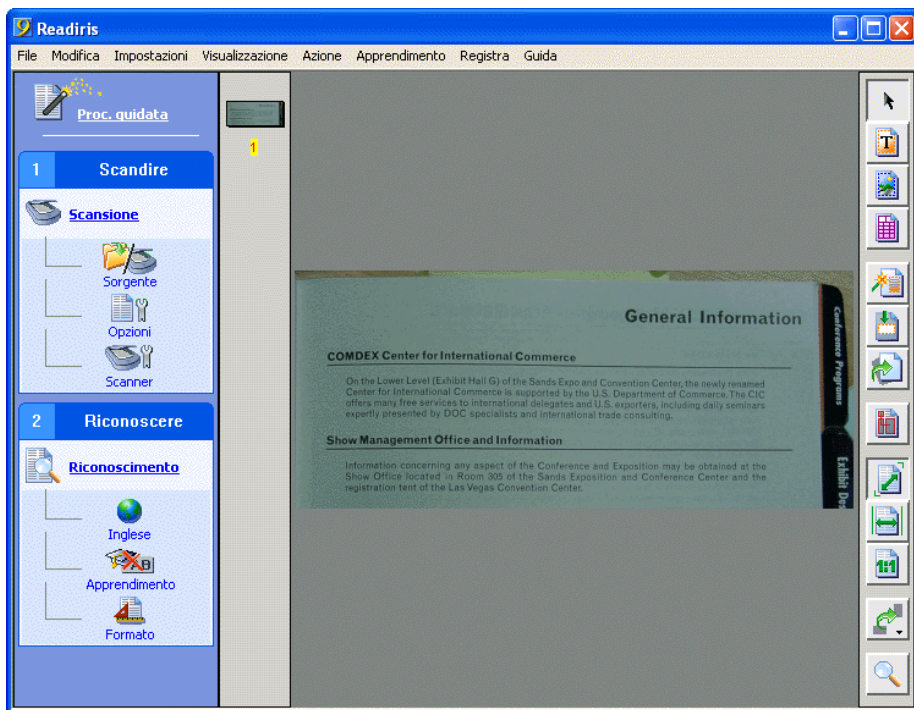
per la ripresa di immagini fisse (“photo stitching”) ma non preoccupatevi di farne uso per le vostre finalità.

Tenete la telecamera direttamente sul documento per evitare la ripresa del documento da un angolo. Tuttavia evitate che vi siano ombre sul documento determinate dalla vostra mano o dall’apparecchio medesimo! Filmate immagini stabili. Pensate a montare un treppiedi se del caso.

Non fate uso di flash se filmate immagini lucide, per evitare i riflessi. Parlando in generale, adattate la lucentezza ed il contrasto al contesto intorno a voi - luce del giorno, luce da lampada, neon ecc. (Alcune telecamere possono essere calibrate per filmare un documento bianco.)



Per fare un tentativo, aprite l’immagine **DIGITAL.JPG** nella cartella Readiris e eseguite quanto previsto.



## SALVATAGGIO DI IMPOSTAZIONI DI DIFETTO

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 di default per impieghi successivi.

Salva impostazioni standard



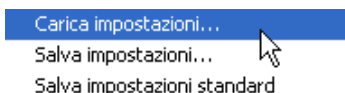


I file di impostazione non solo contengono le **impostazioni** dello scanner: consentono di determinare anche l'eventuale impiego dell'apprendimento interattivo, la lingua, il modo d'uscita (ad esempio il testo inviato a WordPad) ecc. In breve, *tutte* le impostazioni di funzionamento di Readiris vengono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

## **SALVATAGGIO DI IMPOSTAZIONI SPECIFICHE**

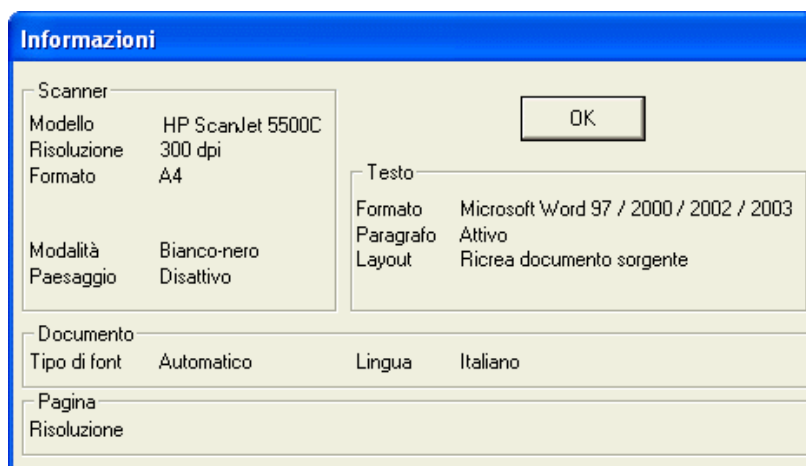
---

Ad ogni avvio si utilizzeranno ovviamente le impostazioni di difetto, 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.



Segue un esempio: qualora si renda spesso necessario 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 di font nella modalità "leggi", in quanto si utilizzano sistematicamente gli stessi tipi di caratteri ecc.

Qualora non si sia certi delle impostazioni correnti, per scoprire quali esse siano non è necessario "precipitarsi" in ogni menu e comando. Per averne una panoramica, è possibile utilizzare il comando "Informazioni" dal menu "File".

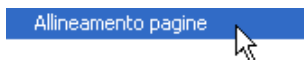


## SCANSIONE DI DOCUMENTI

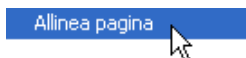
Una volta impostato lo scanner è possibile cominciare a scandire i documenti. Ecco alcuni elementi da ricordare.

In primo luogo prestare attenzione al disallineamento. Sebbene l'analisi ed il riconoscimento pagine siano tolleranti al 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.

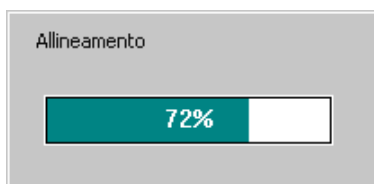
L'opzione "Allineamento Pagine" nel tasto "Opzioni" (e sotto il menu "Impostazioni") definisce se le pagine scandite storte verranno **raddrizzate** automaticamente - un disallineamento limitato viene ignorato. Questa opzione è disattivata per difetto.



Quando dimenticate di attivare questa opzione, utilizzare il tasto "Allinea Pagina" sulla barra strumenti immagine (ed il comando "Allinea Pagina" nel menu "Azione") per "raddrizzare" le pagine scandite storte.

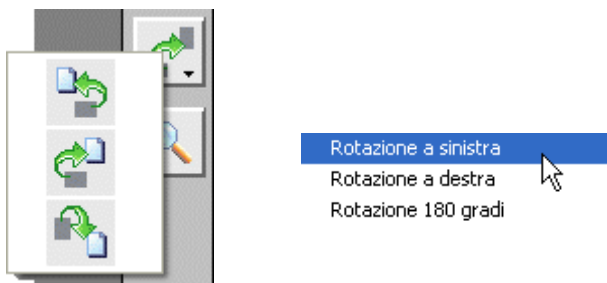


Il raddrizzamento richiede alcuni secondi: l'immagine viene analizzata per rilevare l'angolo di disallineamento e, se necessario, l'immagine a colori o nella scala dei grigi e la sua versione in bianco e nero sono raddrizzate e l'analisi della pagina viene ripetuta.



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.

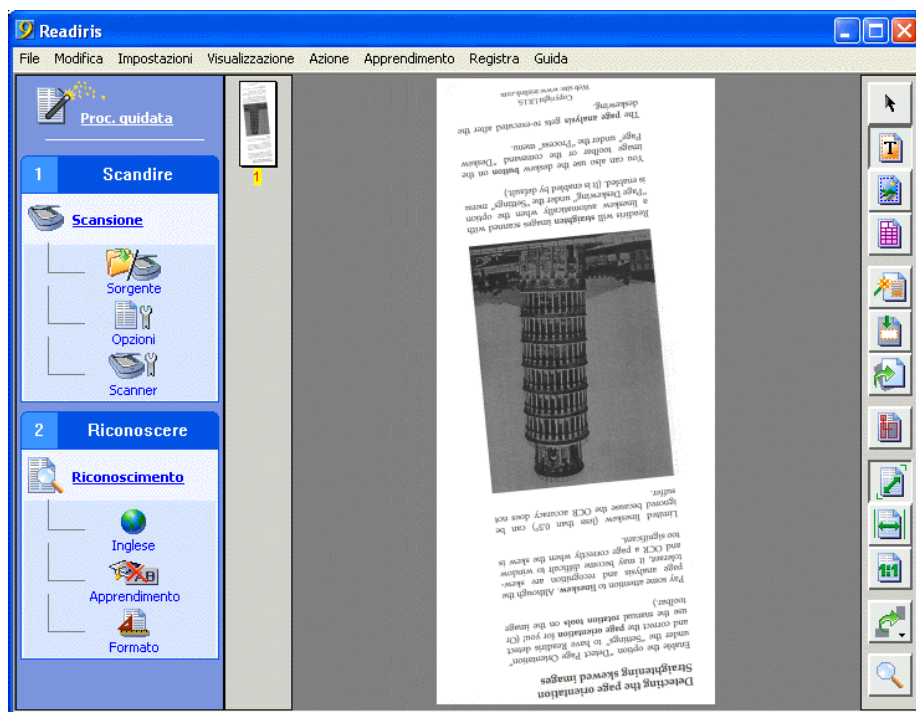




Tuttavia, Readiris può correggere le pagine male orientate. Attivate l'opzione "Rivela l'Orientamento delle Pagine" nel tasto "Opzioni" (o sotto il menu "Impostazioni") e Readiris correggerà l'orientamento della pagina ove necessario.



Se lo desiderate, voi potete fare un buon uso dell'immagine DESKEW.JPG della cartella Readiris. Attivate le opzioni "Allineamento Pagine" e "Rivela l'Orientamento delle Pagine" prima di aprire l'immagine, e lasciate che Readiris riaggiuste la Torre di Pisa... come noi la preferiamo!

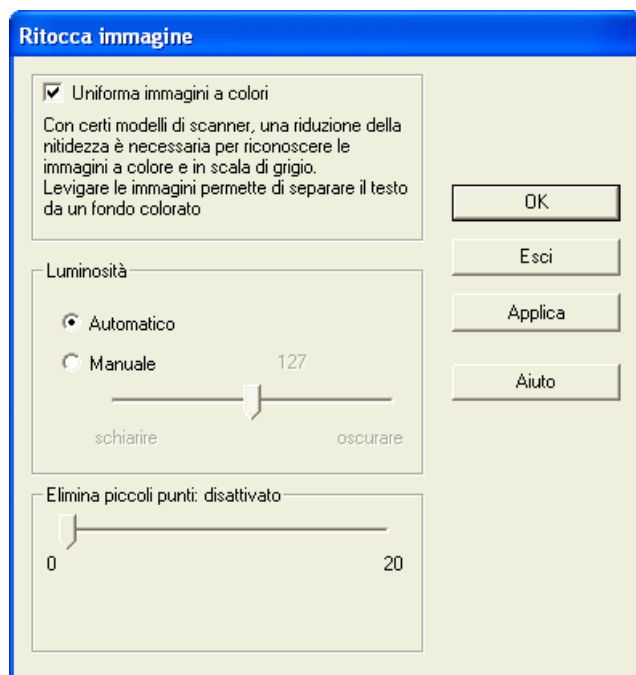


## REGOLAZIONE DELL'IMMAGINE SCANDITA

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



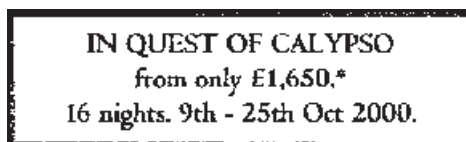
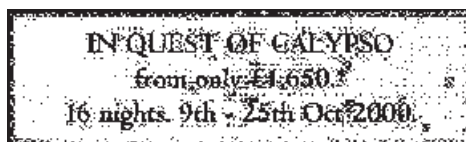
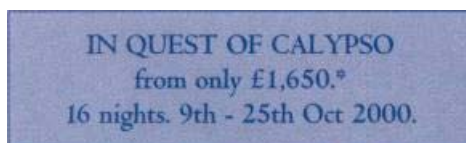
Quando si accede a questo comando, la versione in bianco e nero viene automaticamente visualizzata. (Come se fosse stata disabilitata l'opzione "Visualizza Documento a Colori"! ). A questo punto subentrano alcuni concetti complicati che dovranno essere spiegati dettagliatamente.



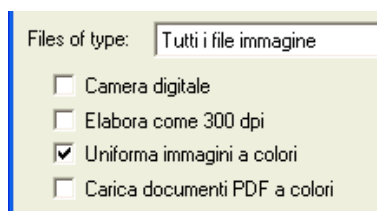
L'opzione "Uniforma Immagini a Colori" rende le immagini con scale di grigio e a colori più omogenee "appiattendolo" e uniformando le differenze di intensità relative. Come risultato si ottiene un contrasto più netto tra il primo piano (il testo) e lo sfondo (un colore, motivo ecc.).



Questa funzione di **pre-elaborazione** può sembrare molto tecnica e di difficile comprensione ma sicuramente ha il suo ruolo: con alcuni modelli di scanner, la riduzione dell'intensità è necessaria per riconoscere le immagini a colori e in scala dei grigi. L'uniformazione è talvolta il solo modo per separare il testo dallo sfondo colorato. Di seguito un'immagine campione semplicemente illeggibile senza uniformazione dell'immagine.



Questa opzione può essere attivata quando si memorizzano immagini prescandite!



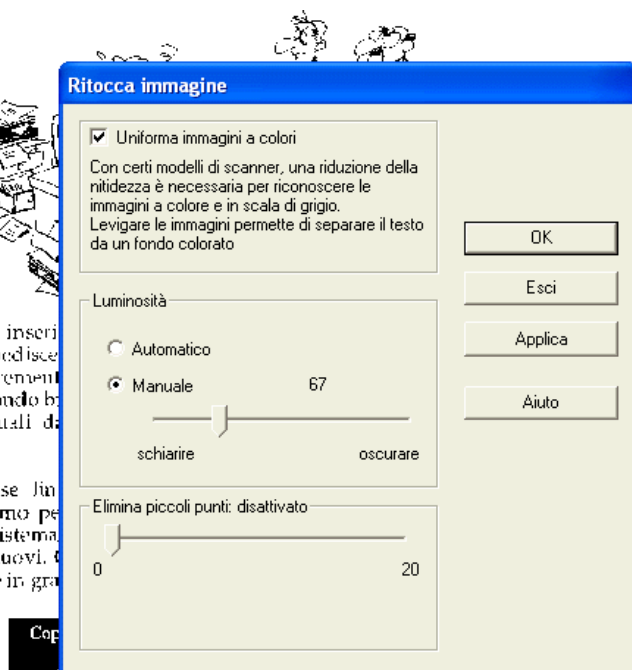
Passiamo ora alla luminosità. Per "luminosità" intendiamo la soglia in bianco e nero. L'impostazione "Automatico" determina automaticamente la soglia bilivello. Se necessario applicare una soglia differente scurendo o schiarendo l'immagine

in bianco e nero: quando l'immagine viene scurita, un maggior numero di pixel diventa nero nella versione in bianco e nero, mentre quando l'immagine viene schiarita, meno pixel diventano bianchi nella versione in bianco e nero.

Si noti che non viene attuata alcuna regolazione dell'immagine finché non si clicca sul tasto "Applica"! Cliccando su "OK" viene eseguita la regolazione e si chiude la finestra. Ecco un esempio in cui l'immagine in bianco e nero è stata fortemente schiarita - senza tenere conto della precisione del processo OCR!

Il documento viene inserito nel computer e quindi spedito al documento. Il documento è semplicemente significato su uno sfondo bianco. Le informazioni testuali del documento attribuite dei nomi.

Si serve del database. In questo strumento potentissimo per l'utente insegna al sistema a serie di documenti nuovi. Con font. In altre parole, è in grado



Le prime due opzioni riguardano le immagini a colori e nelle scale di grigio, mentre l'ultima opzione, "Elimina Piccoli Punti", è utilizzabile esclusivamente su

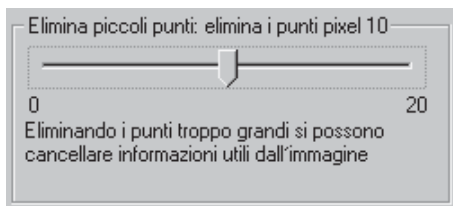


immagini in bianco e nero. La smacchiatura significa che i pixel “parassiti” verranno rimossi dalle immagini in bianco e nero.

**If computers can't  
adapt easily, then  
maybe the people  
using them can.**

**If computers can't  
adapt easily, then  
maybe the people  
using them can.**

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

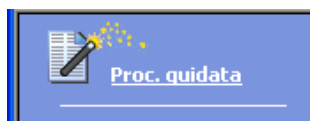


Il modo migliore per ottimizzare le immagini per il processo OCR è il seguente: posizionare la finestra di regolazione in modo tale che non impedisca di valutare la regolazione eseguita sull'immagine. Adattare i parametri - cliccando ogni volta su "Applica" - finché l'immagine risulterà nitida e chiara.

## **LA PROCEDURA GUIDATA FÀ TUTTO PER VOI!**

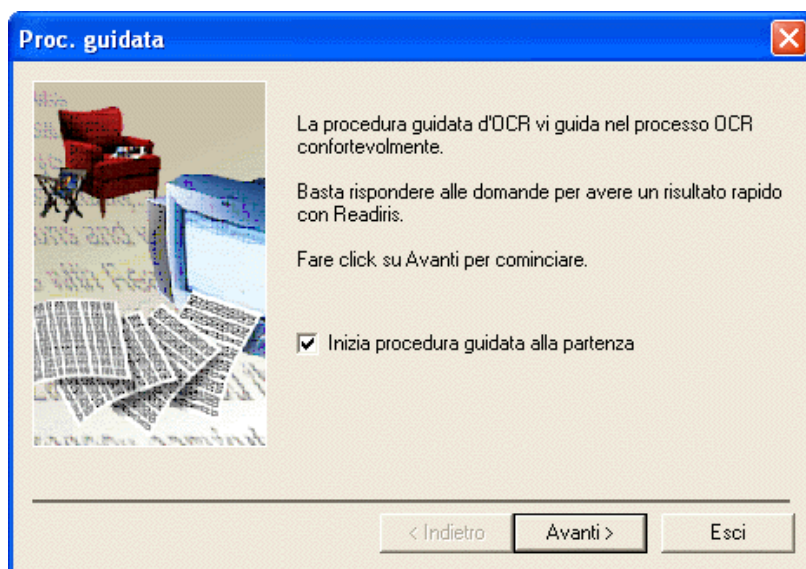
E ora cominciamo a catturare i documenti. Invece di scorrere tutti i parametri, utilizzare la **procedura guidata d'OCR** (“wizard”), un modo decisamente agevole di riconoscimento delle pagine.

Cliccare il comando "Proc. Guidata" sulla barra degli strumenti principale (o selezionare il comando "Procedura Guidata" nel menu "Azione").



Procedura guidata... Ctrl+D

L'assistenza vi guida agevolmente attraverso la procedura OCR: basta rispondere a qualche domanda e si otterranno facilmente risultati rapidi con Readiris.



Si noti che la procedura guidata d'OCR inizia ogni volta che usate Readiris; potete evitarlo disattivando l'opzione "Inizia Procedura Guidata alla Partenza" nel primo schermo dell'assistenza (e con l'opzione equivalente nel menu "Impostazioni").



## READIRIS RICREA IL LAYOUT DEL DOCUMENTO

---

L'assistente 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 scandito, applicando la formattazione di parola, paragrafo e pagina del documento originale.

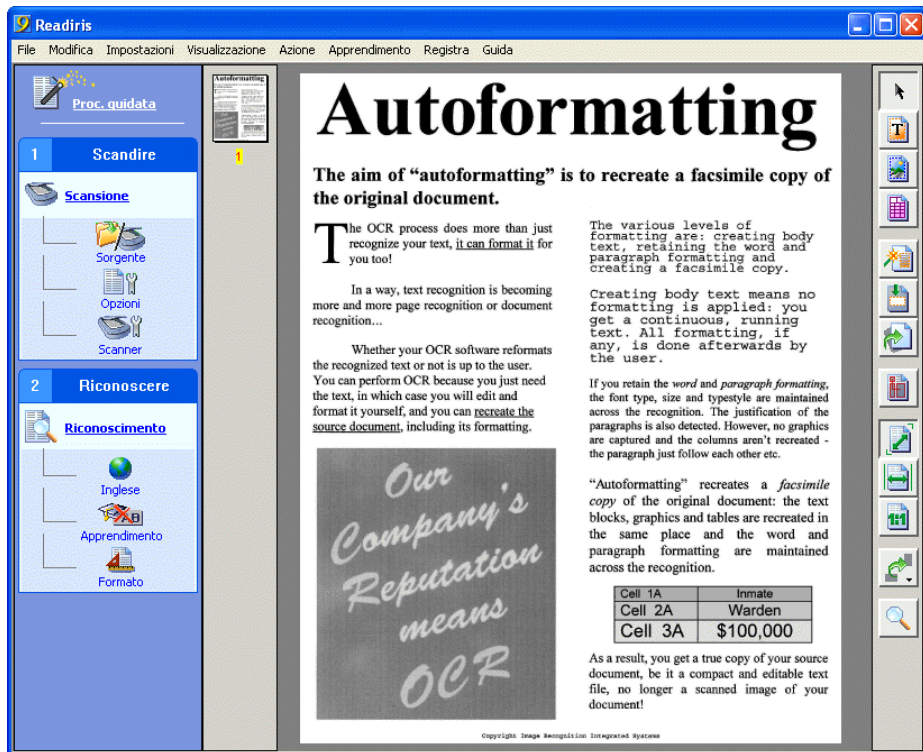
Durante il riconoscimento vengono utilizzati tipi di carattere (serif e sans serif, proporzionale e fisso, normale e condensato) simili a quelli del documento di partenza e vengono mantenuti i corpi e gli stili dei caratteri (grassetto, corsivo, sottolineato, apice e pedici). Si ricreano le tabulazioni e l'allineamento (a sinistra, centrato, a destra e giustificato) di ciascun blocco di testo. Lo stesso avviene per gli elenchi puntati ("bullets") e numerati. Gli indirizzi e-mail e le URL delle pagine web vengono rilevati e ricreati come hyperlink nell'uscita. Il posizionamento delle colonne, dei blocchi di testo e degli 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 scansionata!

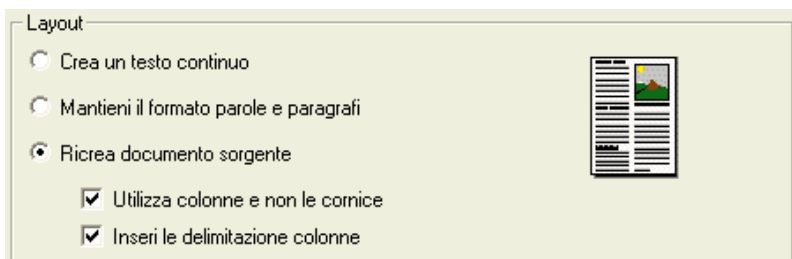
Pertanto, quando si utilizza "l'autoformattazione", viene *parzialmente* applicato l'ordinamento di finestre: è possibile includere ed escludere zone, ma qualsiasi riordinamento delle zone viene semplicemente ignorato!

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



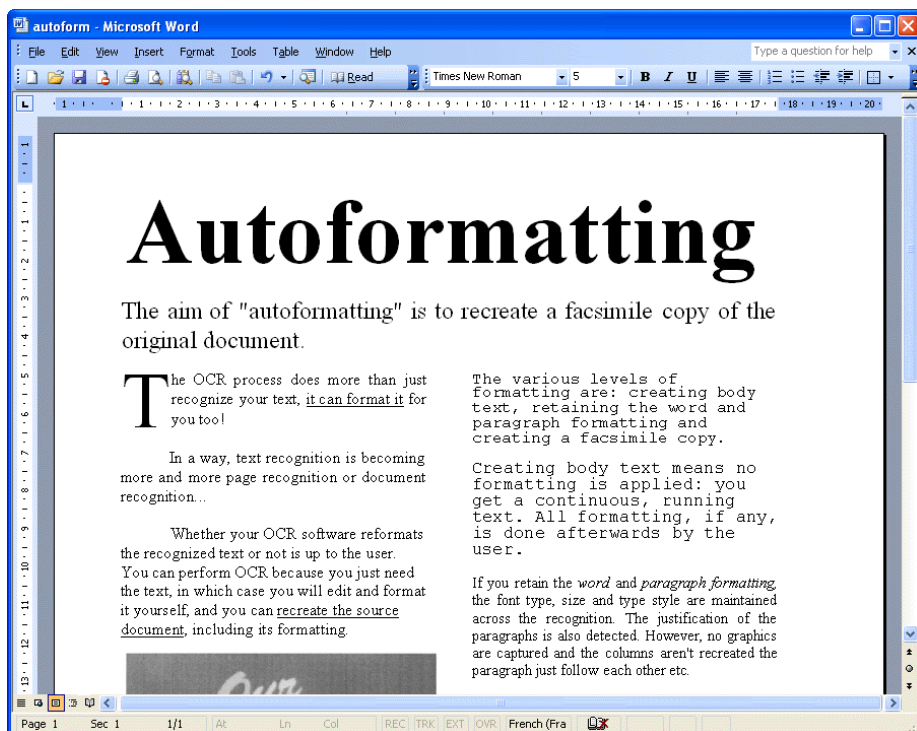


Cliccare sul pulsante "Formato" della barra strumenti principale e inviare il risultato OCR a Microsoft Word o selezionare il formato "RTF" (Rich Text Format) o Word (DOC). Selezionare "Ricrea Documento Sorgente" come opzione di layout. (Viene attivata per difetto l'opzione "Aggiungi Linee in Paragrafi" per attivare il ritorno a capo all'interno dei paragrafi.)

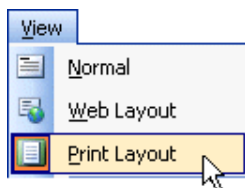


La ricostruzione del layout è disponibile a seconda del modo di output selezionato. Alcuni formati “poveri” che producono testi “semplici” quali Testo (ANSI), Testo formato MS-DOS (ASCII) ecc., *non* supportano codici di formattazione avanzati e quindi non sono in grado di offrire l’autoformattazione. 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 programma di videoscrittura, il testo appare in questo modo senza *alcun* intervento da parte dell’utente.



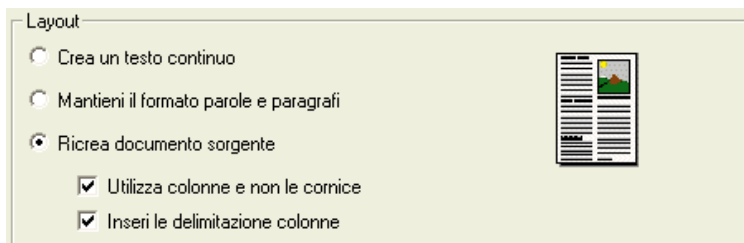
Per visualizzare correttamente l'effetto, è necessario abilitare il modo "WYSIWIG" del programma di videoscrittura, generalmente chiamato modo "layout di pagina". Tuttavia, inviando il documento riconosciuto direttamente a Microsoft Word, il layout di pagina o il layout di stampa viene attivato automaticamente.



In breve, Readiris non solo riconosce i testi ma è anche in grado di formattarli per l'utente. OCR non è più solo riconoscimento di testo, ma significa anche sempre più **riconoscimento di pagina** o di documento!

## COLONNE, NON CORNICE!

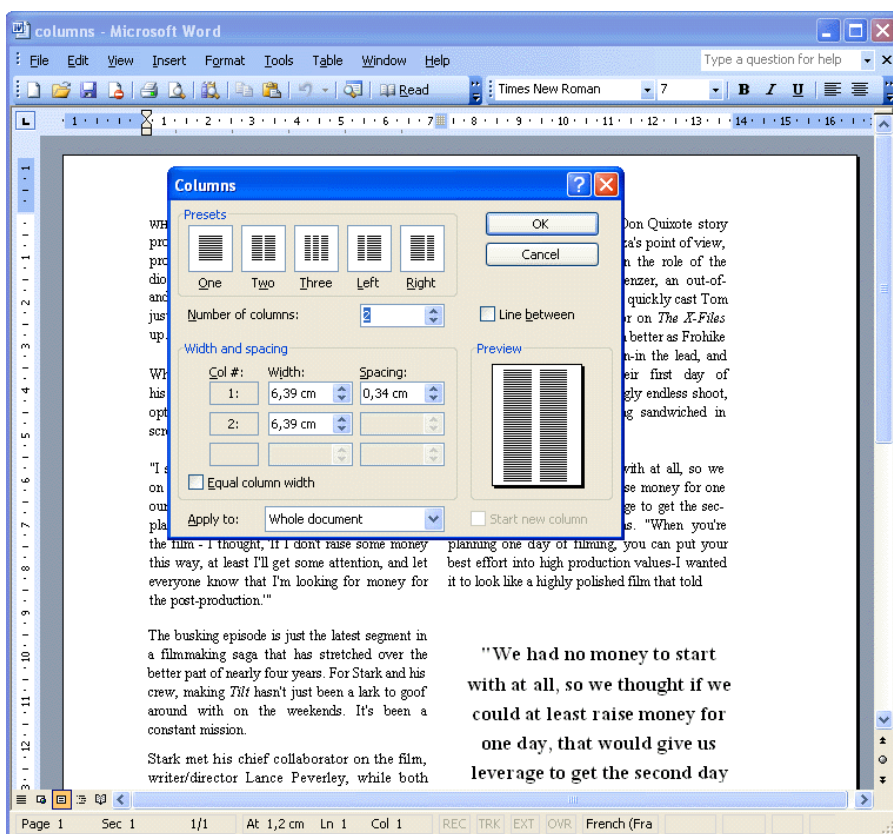
L'opzione di formattazione "Utilizza Colonne e Non le Cornice" stabilisce *come* viene eseguita l'autoformattazione: i blocchi di testo, le tabelle e i grafici possono essere memorizzati in riquadri o in **colonne** modificabili.



Le “cornice” sono contenitori separati di testi utilizzati per posizionare diversi blocchi di testo, grafici 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.

Supponiamo ora che vi siano colonne vere nel documento scandito: se il sistema non è in grado di individuare colonne nel documento sorgente, questo modo di formattazione utilizza comunque cornice come una posizione di “riserva”.

È possibile utilizzare l'immagine COLUMNS.TIF nella cartella Readiris se si vuole provarla.



L'opzione "Inserisci le Delimitazione Colonne" rifinisce la ricreazione di colonne: stabilisce se inserire o meno interruzioni di colonna "rigide" alla fine di ogni colonna.



Con le interruzioni di colonna, il testo che editate, aggiungete o spostate rimane all'interno della colonna. Nessun testo si sposta automaticamente con un'interruzione di colonna. Tutto il testo dopo un'interruzione di colonna viene spostato nella parte superiore della colonna successiva.

Attivate questa opzione se desiderate mantenere le delimitazioni di colonna dove queste sono state rilevate nel documento riconosciuto, indipendentemente dall'editing del testo dopo l'OCR. Nei quotidiani e nelle riviste, le diverse colonne di una pagina corrispondono spesso a diversi "pezzi" di un articolo. Se il testo si spostasse da una colonna a quella successiva, potrebbe non essere una buona idea.

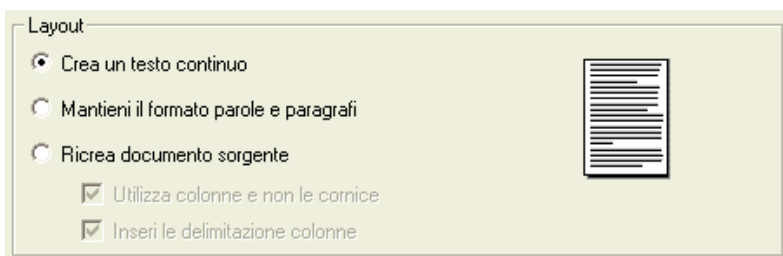
Disattivate questa opzione quando si avete un testo in colonna: in questo modo si garantirà lo spostamento naturale del testo da una colonna a quella successiva.

## 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 sul lato destro, creare un **corpo del testo** significa produrre un testo non formattato, "continuo". Il testo verrà catturato, ma ingorando completamente la sua formattazione. Utilizzare questa opzione quando si desidera ricattare solo un testo, ma non il suo layout.



Il testo del corpo è quello che si ottiene quando si esegue rapidamente il riconoscimento di una zona di testo cliccando con il lato destro del mouse e

selezionando il comando "Copia come Testo". Una volta eseguito il riconoscimento, si incollerà il testo del corpo nell'applicazione testo.

L'opzione "Mantieni il Formato Parole e Paragrafi" costituisce una via intermedia: durante il riconoscimento vengono mantenute la **formattazione della parola** - tipo di font, corpo e stile del carattere -, e la **formattazione del paragrafo** - 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.

## **ESPORTARE IL TESTO DIVERSE VOLTE**

È possibile esportare i risultati dell'OCR diverse volte senza dover ripetere il riconoscimento! Modificate il formato testo e le opzioni di formattazione sotto il tasto "Formato" e cliccate di nuovo il tasto "Riconoscimento". 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 testo e li invia all'applicazione target selezionata.



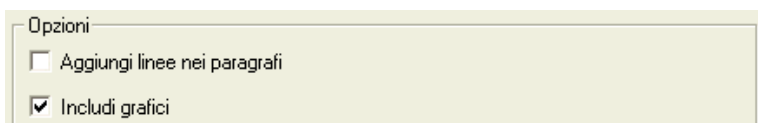
Lo stesso vale per qualsiasi altro elemento che modificate: quando aggiungete una pagina al lavoro OCR, solo quella pagina sarà riconosciuta. Quando create una nuova zona testo su una pagina, solo quella zona sarà riconosciuta prima dell'esportazione dei risultati.

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 DI GRAFICA**

Nell'esempio riportato, la grafica era inclusa nel testo riconosciuto; questo dipende dall'opzione di formattazione "Includi Grafici". Il salvataggio dei grafici all'interno del testo dipende ancora una volta dal modo d'output. I formati di testo "poveri" come (ANSI) ecc. non memorizzano grafici.



Con Readiris è possibile salvare i grafici senza eseguire il riconoscimento testo. Dal momento che Readiris crea immagini in bianco e nero, con scale di grigio e a colori, è possibile catturare senza problemi grafici al tratto e fotografie.

Come? Tracciare una zona di grafica intorno alle illustrazioni, ai disegni ecc. La creazione manuale di finestre di grafica avviene nello stesso modo del disegno di finestre di testo e tabelle, basta selezionare lo strumento "Finestra Grafica" sulla barra strumenti immagine

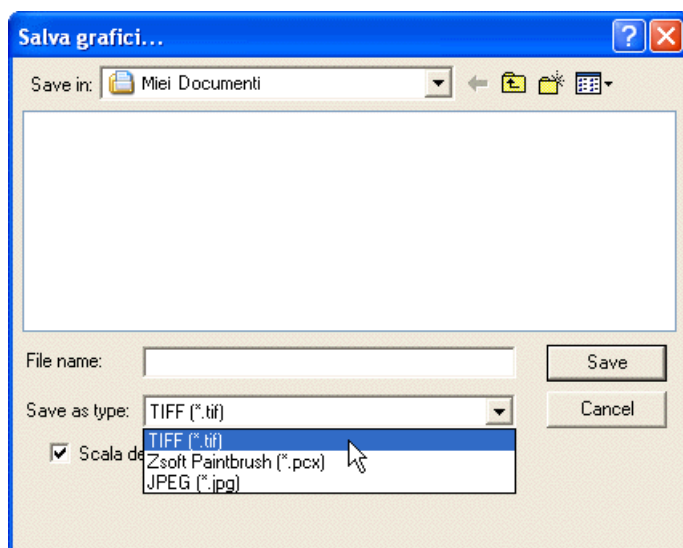


Quindi, scegliere il comando "Salva Grafici" del menu "File".

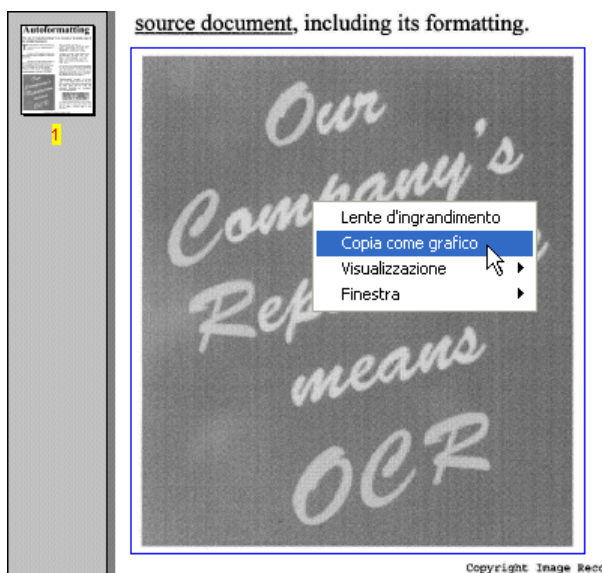


Il sistema chiederà di specificare un nome di file. Stabilisca il formato grafico che si intende utilizzare. Selezioni un formato sostenuto dal software di ritocco immagine. Sono disponibili i formati JPEG, TIFF e Paintbrush (PCX). Abiliti l'opzione "Scala dei Grigi/Colore" per salvare l'elemento grafico come grafico a colori o con scale di grigi.





Per inviare un grafico negli appunti invece di salvare un file immagine, cliccate con il lato destro del mouse sulla finestra e selezionate il comando "Copia come Grafico": la zona grafica sotto il mouse viene subito incollata!



## LETTURA DI FAX E RICONOSCIMENTO POSTICIPATO

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

Scandire semplicemente il documento. Selezionare il comando "Salva Pagina Completa come Immagine" nel menu "File" per salvare una pagina singola. Vi verrà chiesto di salvare l'intera pagina come file TIFF o Paintbrush (PCX).

Salva pagina completa come immagine...

Salva tutte pagine come immagine...

Selezionare il comando "Salva Tutte Pagine come Immagini" per salvare un documento multipagina. In questo caso è disponibili uno formato di file: TIFF multipagina.

A questo punto è possibile selezionare il disco come fonte di immagini ed aprire il file di immagini con il tasto "Apri" (o con il comando corrispondente nel menu "Azione"). (Qualora si utilizzi il comando "Apri" del menu "File", non è neppure necessario aggiornare la sorgente di immagini.)

Poiché immagini a colori, con scale di grigio e in bianco e nero sono supportate in eguale misura, Readiris apre documenti PDF Adobe Acrobat, immagini JPEG, immagini Paintbrush (PCX), immagini fax DCX (una versione multipagina del formato Paintbrush), immagini PNG, immagini TIFF (non compresse, compresse di LZW, PackBits, Gruppo 3 e Gruppo 4), immagini TIFF multipagina e bitmap di Windows (BMP).

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 - 100 x200 dpi ("qualità normale"), 200 x 200 dpi ("qualità minuta") e 200 x 400 dpi ("qualità superminuta") - e "pre-elabora" queste immagini automaticamente per garantire risultati OCR ottimali.

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

Per aprire immagini, non dimenticare che è possibile cliccare con il pulsante destro del mouse sulle immagini in Esplora Risorse di Windows e selezionare il comando "Riconoscimento" dal menu "Contesto" per aprire le immagini! In alternativa utilizzare la funzione "trascinare selezione": dopo aver trascinato file di immagini dal Esplora Risorse di Windows alla zona immagine o sull'icona di Readiris, è immediatamente possibile aprire questi file.

## **RICONOSCIMENTO DI TABELLE**

---

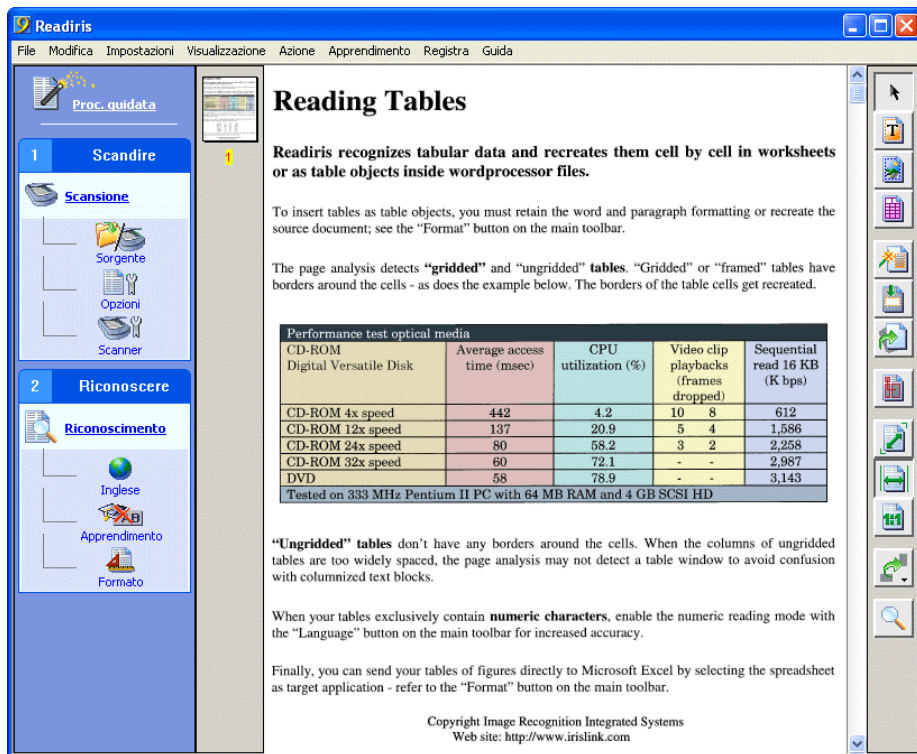
Fino ad ora sono stati spiegati il riconoscimento di testi e di fax e il salvataggio degli elementi di grafica. Si passerà ora all'elaborazione di una tabella. Prendere



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 e non a caso! L'analisi della pagina divide in zone l'intera tabella e Readiris è in grado di ricostruirla, ricreando le tabelle cella dopo cella nel foglio elettronico o inserendo un oggetto tabella all'interno del file del vostro sistema di videoscrittura.

Analizziamo le diverse soluzioni iniziando dalla tabella “con griglia” o “incorniciata”. È caratterizzata da bordi intorno alle celle.



Eseguire 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 il sistema di videoscrittura 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

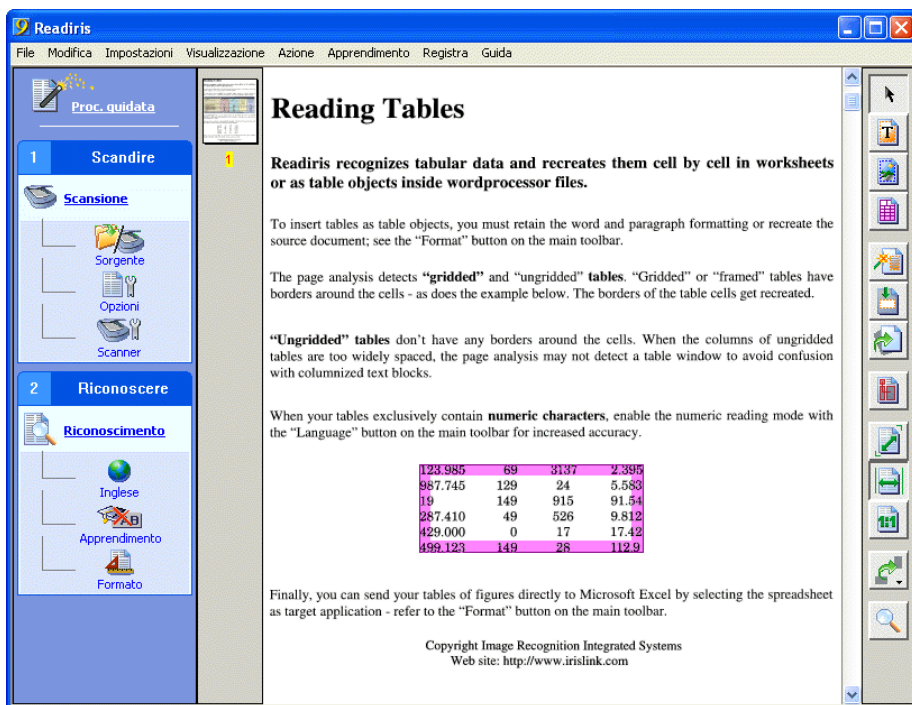
Type a question for help

Times New Roman 12 B I U

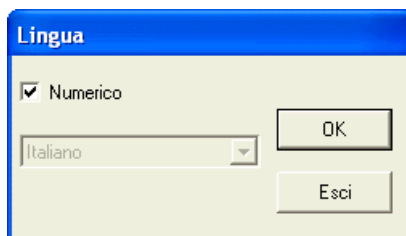
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

Performance test optical media					
CD-ROM	Average access	CPU	Video clip		Sequential
Digital Versatile Disk	time (msec)	utilization (%)	playbacks		read 16 KB
			(frames		(K bps)
			dropped)		
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					

Passiamo ora all'esempio "senza griglia". Non vi sono bordi intorno alle celle. L'analisi della pagina individua comunque la tabella.

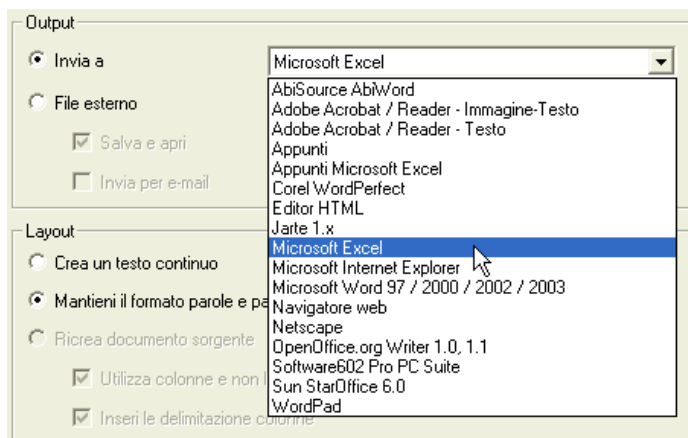


Per una precisione ottimale di OCR, limitare il riconoscimento ai **simboli numerici** mediante il pulsante "Lingua". (La modalità numerica non riguarda strettamente i numeri, ma include i simboli da 0 a 9, +, \*, /, %, , (virgola), . (punto), (, ), -, =, \$, £, ¥ e il simbolo €.)



Poiché questo procedimento è possibile solo quando la tabella non contiene simboli alfabetici - altrimenti non sarebbe possibile riconoscere correttamente le porzioni di testo -, ora è possibile attivare la modalità numerica, ma non per la prima tabella.

In questo caso, invieremo il risultato OCR direttamente al foglio elettronico Microsoft Excel. Selezionare quindi Excel come applicazione target con 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.

	A	B	C	D	E
1	123.985	69	313	2.390	
2	987.745	129	24	5.589	
3	19	149	915	91.549	
4	287.410	49	526	9.812	
5	429.000	0	17	17.429	
6	499.123	149	28	122.098	
7					
8					

Potreste avere tabelle “senza griglia” non individuate dall’analisi della pagina come zona tabella perché le colonne sono troppo spaziate. Readiris cerca di evitare confusione con i blocchi di testo organizzati in colonne. Per creare manualmente una finestra di tabella, cliccare "Finestra di Tabella" nella barra strumenti immagine, e procedere come d’abitudine; la casella strumenti indica di nuovo il numero di finestre di tabelle.

Disegna una finestra di tabella: 0



## USO DELLA GUIDA IN LINEA

Qui termina la panoramica di Readiris. Non è possibile includere nel presente manuale alcune informazioni dell’ultimo minuto. Quindi, per ulteriori informazioni su Readiris si consiglia di consultare il sistema di guida in linea.



Per fare ciò, passare al menu "Guida". Il comando "Argomenti della Guida" e il relativo tasto di scelta rapida F1 consentono di navigare attraverso i numerosi argomenti della guida.



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