

Scannable Form Design Best Practices

Verity White Paper



Scannable Form Design Best Practices

April 2005

Table of Contents:

Use Verity TeleForm Designer to Create Forms Quickly, Easily	3
Preparing for Form Design	3
The following steps should be taken before beginning any new form.	4
Key Questions	4
Form Type Selection	4
Traditional Forms	4
VersiForms.....	4
Field Type Selection	5
Field Type Optimization	5
Configuring Constrained Print Fields.....	5
Configuring Image Zone Fields	6
Configuring Choice Fields.....	6
Configuring Entry Fields	7
Configuring Capture Zones	7
Edit Boxes and Choice Lists	8
Configuring Tab Order within a Capture Zone.....	9
General Capture Zone Properties	9
Point-and-Click OCR	9
General Field Properties	9
Settings within the Validation Tab	9
Required Field	9
Always Review	9
Numeric Ranges	9
Database Lookups	9
Additional Settings within the Processing Tab	9
Confidence Levels	9
Preprocessing.....	10
Dictionary Files	10
Additional Settings within the Post-Processing Tab	10
Default Values	10
Job Maintenance	10
Optimizing a Job for Form Identification	10
Image Quality Control (Image QC)	11
Data Review	11
Form Optimization Recommendations	11
Optimizing OCR / ICR Fields	11
Capture Zones	12
Existing Forms	12
Multi-Page Forms	13

Publisher's Note: Information contained in this document is intended for guideline purposes only. Verity product documentation supersedes information contained in this document. The situations described in this document are offered as examples; actual configurations and results will vary from system to system.

Use Verity TeleForm Designer to Create Forms Quickly, Easily

Verity TeleForm Designer is a point-and-click application for easily implementing automated forms processing solutions. Designer includes complete layout tools to create new forms and to automate your existing forms. It also has configuration options that use pre-made business rules, recognition settings to maximize throughput, and Connect Agents to tie your forms and documents to back-end systems.

The TeleForm Designer interface is similar to most other form design applications on the market. And, the user interface is very intuitive. However, should you desire, form design training is available from Verity Education at your request.

How forms are designed will have a significant impact on their readability once they are returned to the system for processing. Three types of forms can be produced within TeleForm Designer: Traditional Forms, VersiForms, and Existing Forms. Review the Form Types section below for additional assistance in choosing the appropriate form for your needs.

There are also numerous field types available within the Designer module. The types of fields that are used and how these fields are configured are the most important aspects of form design. Each field type is available within the Form Design Toolbar as depicted above. This document contains specific information on which fields are to be used and how they are best configured.

This Guide also details ways to organize and streamline the process of form design. For example, shape libraries and expandable form trees will help when organizing forms and objects. For more information on this topic, consult the Form Design Organization section that follows.

The aesthetics and usability of a form is as important as its readability. The Form Design Guide also contains information on how to make sure your forms are user-friendly documents that are targeted for their intended audiences.

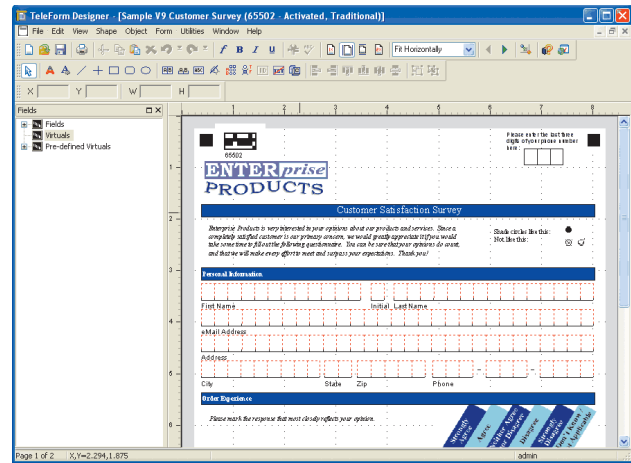


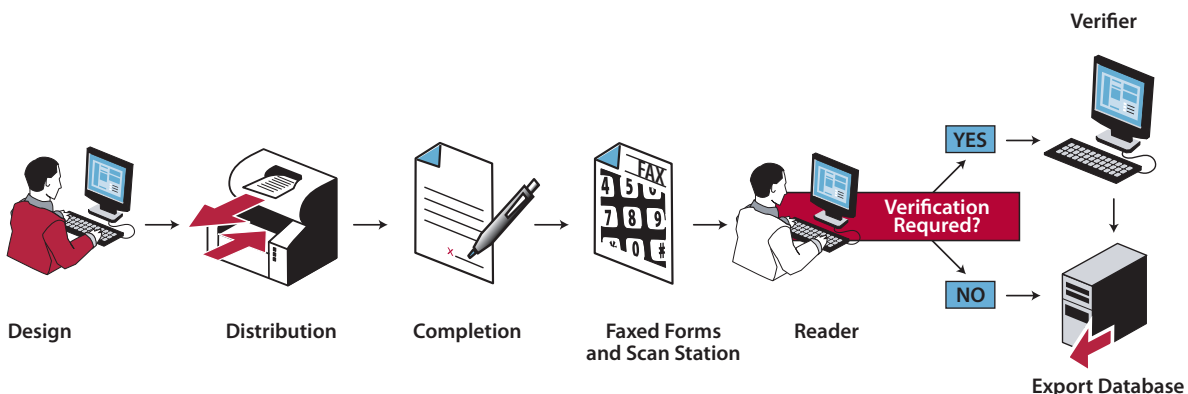
FIGURE 1. The TeleForm Designer Interface

Preparing for Form Design

When designing new forms, try to maintain a balance between aesthetics and readability. Every form is unique and should be approached with a cost-benefit analysis. The goal of form design is not to eliminate all data entry but to create cost-efficient forms. Keep in mind that there will be at least a small degree of Verification required for any group of forms. However, the elimination of 80%—95% of manual entry is a realistic goal. Even when some manual data entry is required, processing is significantly more efficient when using the Verifier module.

Before designing any new form, it is important to take time to plan out how the form should be designed. The form types and field types required for the layout should be identified prior to development. The specific purpose of each form will play a significant role in how the forms are to be developed.

Due to physical requirements, some forms will yield a higher degree of recognition. Readability is very important when designing any form, however aesthetics is often your primary concern. Always keep in mind your intended audience and your form's intended purpose. For example, aesthetics may be of significant concern when developing marketing materials or order forms. Specific individual fields or forms may also require a higher certainty of confidence.



The following steps should be taken before beginning any new form.

- 1) Identify and meet with the form owner and determine whether the form can be redesigned.
- 2) Obtain copies of any existing versions of this form. Work with the form owner to identify any required changes.
- 3) Identify the goal or intended purpose of your form and your form's audience.
- 4) Working with a DBA, technical liaison, or business owner, obtain a list of all required fields and field properties. Field properties should include maximum field lengths, required fields, and field types.
- 5) Identify which form type will be used to recreate your new template. The following section details the three types of TeleForm forms.

The list of additional questions below may be helpful when preparing to design a new form. Keep these thoughts in mind when designing any new form. Always ensure that any unique requirements are taken into consideration before proceeding with development.

Key Questions

- 1) Who is my audience? Does my audience require special assistance? (For example: larger fonts, extra space)
- 2) Are aesthetics an especially important feature of this form?
- 3) What is the form's intended purpose or goal?
- 4) Approximately how many fields are on this form? Are there any areas on my form which require free form entry?
- 5) Does a previous version of this form exist? How can the layout of the original form assist with the redesign of the new template?
- 6) Who owns the existing form? Since forms with external owners cannot be redesigned, is redesign a possibility?
- 7) Have the existing forms already been distributed? Will we want to capture the data from these forms until new forms can be distributed?
- 8) What is the volume of forms that will be returned? Will the forms be returned all at once or periodically? Do these forms require peak processing periods?
- 9) Can a packet of forms be broken into multiple templates? If so, does each form have a unique identifier to link the documents together after processing?
- 10) Is my form a multi-page document? Do we need to ensure that the form pages are properly linked together?
- 11) Does this form require any special considerations or development?
- 12) Will this form be printed internally or printed using an outside vendor?
- 13) Do we know the field requirements for every field on the form? (Form example: field length, required fields, field types) Where can this information be obtained?

Form Type Selection

Before you begin development of any form, you must identify which type of form will best suit your needs. When building a form in Designer, you can choose from three styles: Traditional, VersiForms, and Existing forms.

Traditional Forms

Traditional Forms are designed completely within TeleForm Designer. Visually, they have four square "cornerstone" Reference Marks and a fixed Form ID block. You cannot change the appearance of the Reference Marks or the Form ID

block, but you can move the Form ID block between the upper left and lower right corner of the form. You can also place the Form ID block in both locations. TeleForm uses these marks and blocks to identify what form is being read and where to find the data entry fields on that form.

Traditional forms are by far the most reliable of the form types within TeleForm. It is recommended that all forms be redesigned as Traditional forms when possible. Traditional forms may not be the best choice if:

- 1) Aesthetics are of the utmost importance.
- 2) The form cannot be redesigned or already exists in the field.
- 3) The cornerstone blocks and form ID cannot appear on the form.

VersiForms

VersiForms are created completely within TeleForm Designer. Visually, they contain Reference Marks and a Form ID feature, like Traditional forms. Also like Traditional forms, these marks and Form IDs are used to identify the form template and locate the data entry fields on it. However, VersiForms allow you more flexibility in choosing how your form will look.

You can edit both the size and appearance of the Reference Marks on VersiForms. You can create square, circle, triangle, and angle Reference Marks. Although you cannot delete or obstruct the Reference Marks, you can move them to different places on the page. For multi-page forms, each page can use different Reference Mark styles and locations. You can choose between a text or bar code Form ID, and place it in any of the four corners of the form. VersiForms are very reliable, however they are not quite as reliable as Traditional Forms. VersiForm should be used when:

- 1) Aesthetics are of the utmost importance
- 2) The cornerstone blocks and form ID cannot appear on the form.

The majority of forms you design within TeleForm will use Traditional Form templates. The case for using VersiForms and Existing Forms should be justified for each required form.

Below are two examples of Traditional Form IDs. Form IDs are used by TeleForm to identify each incoming form. The ID on the right has been activated, while the ID on the left has not.



Existing Forms can be quite difficult to configure and require significant testing. Please review the Form Optimization sections for additional information on configuring existing forms. Existing Forms should only be used when:

- 1) The form cannot be redesigned. Examples include insurance forms or forms controlled by government agencies.
- 2) You are planning to redesign a form, but you wish to capture all incoming data from the existing forms in the field.

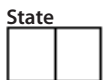
There may be additional scenarios where Existing Forms may be used, however keep in mind that the quality of recognition for an Existing Form is rarely as accurate as for a form redesigned within TeleForm.

Traditional Form	VersiForm	Existing Form
Must have the fixed cornerstone reference marks and the traditional form ID.	Has flexible reference marks and a form ID that can be text or bar code.	Relies on red identification lines and blue reference blocks in the template for recognition.
Create from scratch	Create from scratch	Pre-created and distributed form
Great for faxing	Good for faxing	Bad for faxing
Great for scanning	Great for scanning	Great for scanning

Field Type Selection

The six standard types of fields used within TeleForm Designer are described below. Each of these field types is highly customizable and configurable and has been detailed within both this section and within the Form Optimization section.

Constrained print or comb-style fields will yield the highest degree of character recognition. Where possible, these field types should be used for capturing data within your forms. Image zones can be used to capture and recognize free-form text, however image zones require significant testing and will never obtain the levels of recognition received with constrained print fields. Image zones should be used infrequently and most likely will only be used extensively on existing forms. If you need to capture multiple lines of free-form text, Verity recommends using Capture Zones. Capture Zones allow verifiers to manually enter large blocks of free form text.



- Constrained Print Fields consist of sequential boxes. One character can be written in each box.



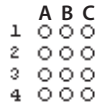
- Comb-Style Fields consist of short lines between each character. One character can be entered between each pair of lines.



- Image Zones define an area on the form where text, bar codes, and encrypted text can be interpreted, or where TeleForm stores signatures and drawings.

Gender
 Male
 Female

- Choice Fields are fields that have a single column (or row) consisting of a list of options. There are various styles of choice fields.



- Entry Fields are characterized by an array of bubbles and are similar to the fields used in educational standardized tests.

- Capture Zones are used for capturing large blocks of data that are best entered manually. Verifier operators can read the data from the form image and type it in. During verification, a tool called Point-and-Click OCR may be leveraged to OCR data printed on the Verification screen.

Field Type Optimization Configuring Constrained Print Fields

Comb-style and date fields are simply preconfigured constrained print fields. These field types will receive nearly the same quality of recognition as all other print fields. Print fields are very reliable and will by far be the most successful field types. In addition to obtaining higher levels of recognition, constrained print fields are much faster and easier to correct within Verifier.

Prior to developing your form, it may be necessary to obtain a copy of the database schema for your backend system. When designing constrained print fields, you will need to specify field lengths, data types, and required fields. It is important that these properties are consistent with those in your backend environment. Field lengths on your form can be shorter than those identified within your backend system, however they should never be longer.

It is always best to limit the set of expected characters as much as possible. If a field is almost always entered using hand-written

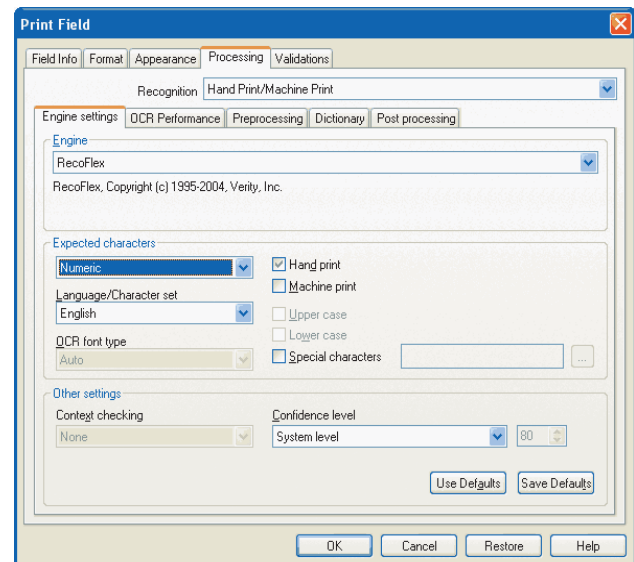


FIGURE 2. Print Fields obtain higher levels of recognition.

font, do not select the machine-print option within the processing window. Limiting the expected character set will increase the level of recognition achieved for an individual field.

Within all field types you can also raise and lower the confidence levels for a given field. A confidence level measures how certain TeleForm must be before bypassing verification. It is recommended that confidence levels are adjusted sparingly. The default confidence levels set by the application are very reliable. Only in rare situations should these levels be modified. In some instances you will need to raise a confidence level for a field that is very critical. There are very few circumstances by which you would want to lower these levels. You will want to conduct additional testing anytime confidence levels have been modified on a given form.

Additional properties are available for constrained print fields. Please review the General Field Properties section within this document for additional information on these settings.

Configuring Image Zone Fields

Image zones are an extremely flexible tool and can be extremely useful when designing forms. These fields are never as reliable as constrained print fields for OCR and ICR. Image zone can be used to read text within Existing Forms and can be useful when the constrained print fields are not an option due to space limitations. Image zones can be used for any of the following purposes:

- Reading hand and machine print.
- Reading or printing barcodes.
- Calculating the percentage of an area filled.
- Saving an area of a form as an image or BLOB.

You can identify an image zone's function within the Recognition dropdown in the Processing tab. The options available for hand- and machine-print are the same as those available within a constrained print field. An additional option titled Context Checking is also available to image zones. Context Checking improves recognition by relying on nearby character types to identify the

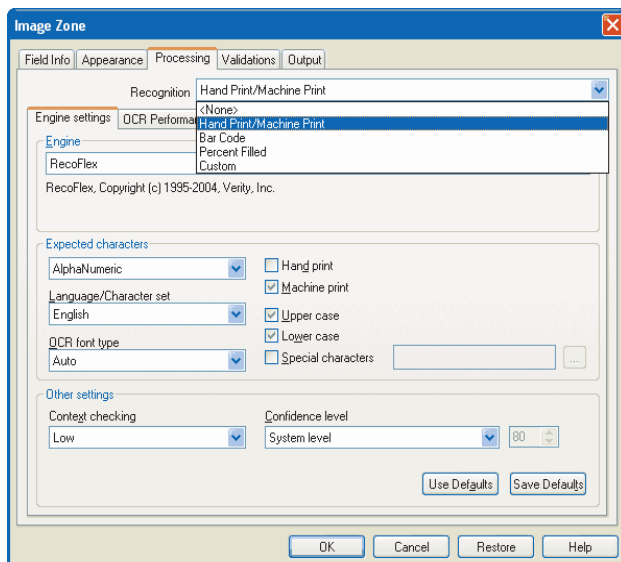


FIGURE 3. Image zones are an extremely flexible tool.

type of a questionable character. If for example an address field is alphanumeric, Context Checking will be helpful for improving recognition of the street number and street name. The “1” in 412 Holland Lane may be interpreted as an “i” if Context Checking were disabled. Within image zones you can also specify OCR font type. You can improve recognition when reading machine print by specifying the font type within this field.

Additional properties are available for image zone fields. Please review the General Field Properties section within this document for additional information on these settings.

As with all field types, you will want to ensure that values captured within TeleForm are no longer than the field lengths for your backend database. To adjust TeleForm field lengths modify the Length/Dec fields within the Output tab. To set length requirements for a given field modify the Length Restrictions properties within the Validations dropdown. Within this section you can specify an exact length or a maximum length for capture.

Configuring Choice Fields

Choice Fields consist of a list of options. The recipient of the form chooses from these options by selecting the appropriate response mark. The available types of response marks include bubbles, boxes, brackets, lines, and the name of the option. You can choose the type of response mark when you add the Choice Field to your form. Bubble fields will always yield the highest degree of accuracy. It is recommended that you use Choice Fields where applicable. Checkboxes are also very effective and are often more appealing aesthetically. The other types of response marks should be used sparingly. Keep in mind that these options will not yield the same degree of accuracy as bubble fields and checkboxes.

Although the choices shown on the form may be simple, such as “blue,” you can configure the field to export detailed values based on what option is selected. For example, if your company uses part numbers to indicate the paint selection for a product,

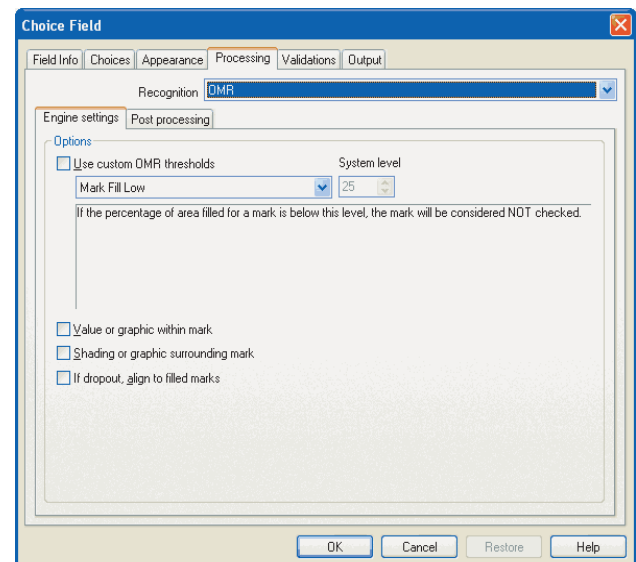


FIGURE 4. Configure Choice Fields with a variety of response marks.

you could set up a Choice Field with simple display values, such as “blue,” “red,” and “green,” but configure it to export the corresponding part number so the manufacturing department knows what color to paint the product.

When initially creating a choice field, always select Manual Entry as opposed to Assisted Entry. Configuring a choice field is very simple and selecting Assisted Entry only limits your options. The Choices tab allows you to configure how your individual choices are displayed and how they are stored within the database. Always confirm with a DBA or technical liaison how data should be stored within your database.

The Output tab within the properties window allows you to determine whether more than one choice is permitted. Within this window you can also define how multiple choices are exported.

After your choice field has been configured, you can move around or reposition the individual choices by holding down the ALT key and clicking on the choice.

Additional properties are available for Choice Fields. Please review the General Field Properties section within this document for additional information on these settings.

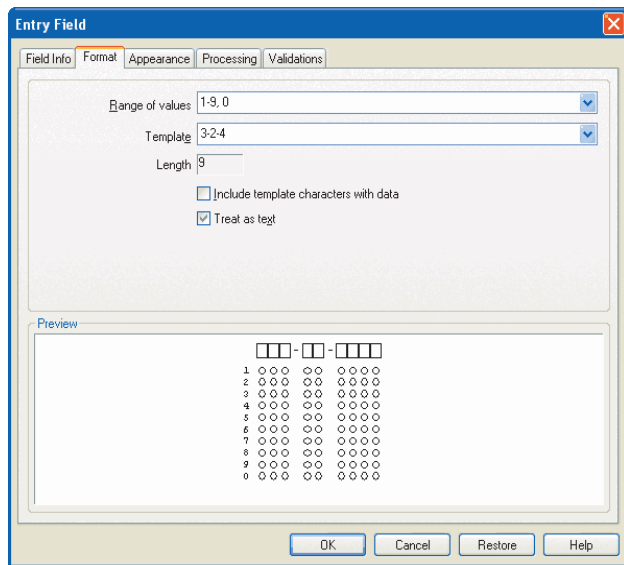


FIGURE 5. Entry Fields are commonly used in surveys or tests.

Configuring Entry Fields

Entry Fields contain an array of response marks, also called “bubbles.” Each response mark is tied to a specific value. You can customize the range of values, such as A-Z; 1-5; A-E; 1-9. Entry Fields are often used in multiple-choice tests, surveys, and standardized tests. Entry Fields are also used when you want to be completely certain that a field has been correctly recognized by TeleForm. Entry fields are extremely accurate, however they may be inefficient for entering alphanumeric and alphabetic text.

Within the Processing tab, you may adjust the OMR thresholds for bubbles within the Entry Field. These OMR thresholds should be the same as those detailed within the Choice Fields section above. These OMR thresholds will rarely ever be adjusted. The thresholds assigned by default are optimal for the TeleForm recognition engines.

Additional properties are available for Entry Fields. Please review the General Field Properties section within this document for additional information on these settings.

Configuring Capture Zones

Capture Zones combine the reliability and indexing ability of manual data entry with TeleForm’s powerful automated recognition and data validation features. Capture Zones allow you to draw a band around an area, which can be displayed for verification. A GUI interface will be displayed with the banded image to allow Verifiers to enter information pertaining to a specific form. Within Designer, you can control the look of this interface. Data validations and TeleForm’s Point-and-Click OCR functionality can be incorporated into the functionality of a Capture Zone.

Capture Zones are recommended any time large amounts of free-form text need to be captured. Capture Zones should also be used any time recognition of an area or set of fields is either difficult or inefficient to recognize. In other scenarios, you may want the Verifier to provide subjective input regarding information on an individual form. For example, a Verifier may subjectively rate the quality of an essay that appears on a college application form.

There are a number of fields and settings available within a banded Capture Zone. The two main field selections are the edit box and choice list. An edit box allows you to type text into a field, while the choice list allows you to select between a list of available options. After drawing one of these fields on to a form, a properties box will appear.

Within the edit box properties tab you must name the field and identify the expected input. The options within the properties box will change depending upon which field type is selected. As with other fields, you can identify whether the text entered is alpha, numeric, or alphanumeric. In addition, you also have the option to specify a unique template, currency or date.

When you are working with Capture Zones, the Designer workspace consists of four components: (See figure 6 on page 8.)

1. **The Capture Zone Menu:** This menu is added to the regular Designer menu bar when you open the SKFI Design Window. It provides all of the commands available for a SKFI Zone and the objects within these Zones.
2. **The Basic Designer Window:** This window shows you the part of the “actual” form on which the SKFI \ Capture Zone is located. You can’t work in the Basic Designer window when the SKFI Design Window is open.

- The Capture Zone Design Window:** This window is similar to the regular Designer window. It serves as an electronic pasteboard on which you can add and arrange objects in the SKFI \ Capture Zone. The Status Bar at the bottom of the SKFI Design Window shows you the position and size of the object that is currently selected in the SKFI Design Window.
- The Capture Zone Toolbar:** The toolbar buttons provide shortcuts to commands that are also available from the Capture Zone Design menu.

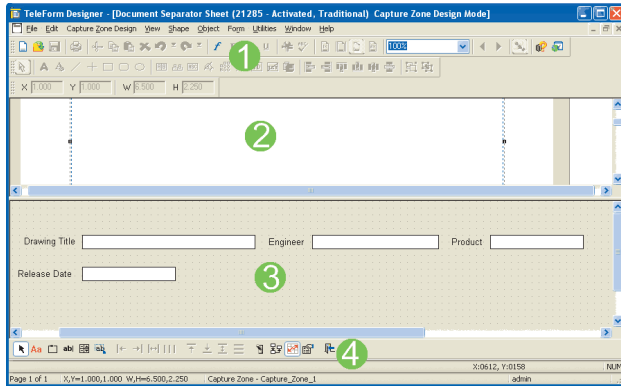


FIGURE 6.

Edit Boxes and Choice Lists

Edit boxes and Choice Lists are the two field types available within Capture Zones. Edit boxes have standard field properties similar to those available for other fields. Any form edit box can be configured to accept the following input:

- Alphanumeric characters
- Alpha characters
- Numeric characters
- Template
- Currency

Template fields can be used to force a Verifier to adhere to a specific format of text. Many of the standard properties are

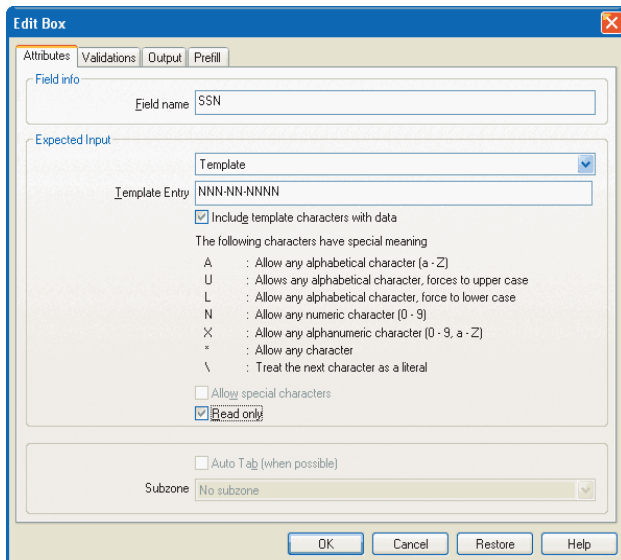


FIGURE 7. In this edit box, a Social Security Number Field has been created.

similar to those for other field types. Please review the General Field Properties section for additional information.

Within the Prefill tab within your Edit Box properties window, you can configure what text appears within the edit box as the corresponding Capture Zone is opened. You have the option of either setting a constant value, prefilling the field with the entry from the last Verified form, or prefilling the field with text from another field. The prefill options available can be very useful for eliminating unnecessary data entry. This feature should be enabled whenever possible.

Choice lists can also be configured within your Capture Zone. These fields allow the Verifier to select between a set of available options. Only one selection can be made per a choice list.

Within the main portion of the properties tab, you can identify whether to enter a list of available choices or to configure the field to pull the list from an available database.

Any time the list of available options is subject to change, it is recommended that the choices are stored within a database. Please consult with a DBA or other appropriate contact to create a list within a database. These database lists should be reused where possible to reduce maintenance.

The Use drop-down list style option available on the form should be used to consolidate space. These drop-down styles are similar to those you find in many websites.

To minimize manual effort, the Prefill with last option may be very useful. This option should be used when forms with similar selections are likely to be Verified in order. Additional field properties are available for Choice Lists. Many of these standard properties are similar to those for other field types. Please review the General Field Properties section for additional information.

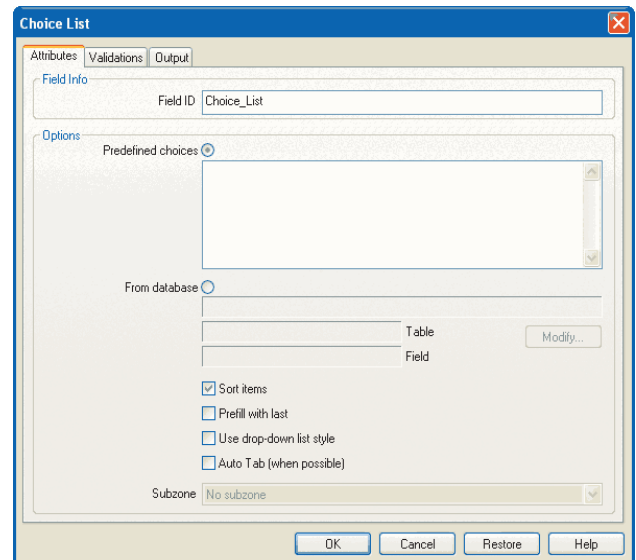


FIGURE 8. Choice lists allow the Verifier to select between a set of options.

Configuring Tab Order within a Capture Zone

Anytime you are creating a Capture Zone always ensure that you configure the tab order for the underlying fields. Verifiers will not want to use the mouse during Verification. Configuring the tab order will ensure that Verifiers can navigate between available fields using the Tab and arrow keys. To configure tab order, click the tab order button and then click the blue box next to each field in the order that you wish configured for the zone.

General Capture Zone Properties

To configure the general properties for a Capture Zone, click on the properties button on the zone toolbar. Within this window you can configure the position of the Image and Data windows. In addition, you can identify what image will be displayed to the Verifier. By default, TeleForm will display the banded zone on your form, however the zone can also be configured to display either the first or last page or an attachment.

Under the Zone Density heading you can configure TeleForm to skip a particular Capture Zone, if no data appears within the corresponding banded area. This is very useful if your Capture Zone is only used when a corresponding area has been completed.

Point-and-Click OCR

The Point-and-Click OCR functionality available within every Capture Zone is extraordinarily powerful, as it will assist with the capture of variable form data. This feature enables Verifiers to drag a box around a zone on the form specifying an area, which contains text you wish to capture. Verifiers will have the ability to manually enter field data in the event recognizable text is not available on the form. This feature is very easy to configure and is quite powerful. Any field using Point and Click OCR should undergo testing using actual forms. To create a field that leverages this functionality, click the Point and Click OCR button located on the Capture Zone toolbar.

Point-and-click OCR functionality is also very effective at capturing variable data that may change location on a form. For example, assume a date was to be rubber-stamped on every incoming form. Capturing this date may be difficult if the location of the stamp is variable and can be arbitrarily placed. Point and click OCR can be enabled to allow the Verifier to identify the location of this banded area for OCR.

General Field Properties

The following section details various field properties, which are generic across most or all field types. These properties have not been detailed within the field specific sections of this document.

Settings within the Validation Tab

Required Field

There are instances where certain fields on your form must be included on every received form. Fields within your backend application may also be required. You will want to ensure that any such fields are identified within the corresponding fields properties. Within the validation tab you can identify whether

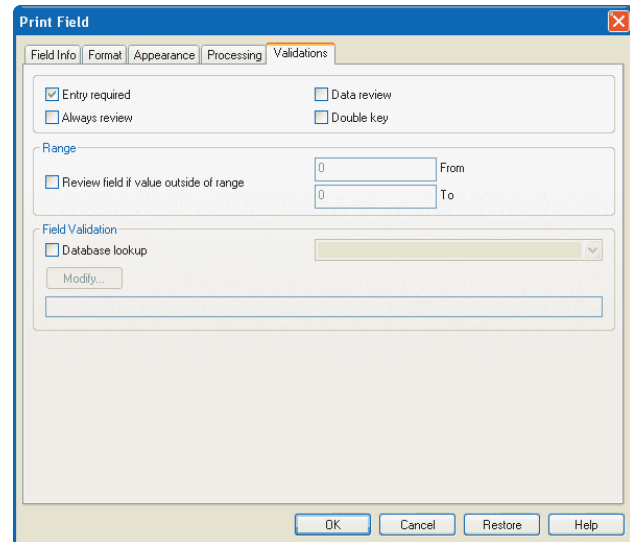


FIGURE 9. Check the Entry Required box to identify a Field requirement.

a given field is required. To identify a field as being required check the Entry Required box.

Always Review

Within the Validation tab you also have the ability to configure fields so that they are always displayed to the user for verification. This feature can be used as an additional validation to ensure that data has been correctly recognized by the application. To enable this feature simply check the Always Review checkbox.

Numeric Ranges

The range section as depicted to the left enables you to set a numeric range for a given field. If any data is entered that is outside of this numeric range, a dialog box will be displayed to the user along with a display message. This property is only available for numeric fields. To enable this feature simply check the box within the Range section of the dialog and specify the range values within the fields provided.

Database Lookups

Database lookups can also be configured within the Validations tab. They allow you to validate a field against a list of values within a database table. You can enable a database lookup by checking the Database lookup checkbox. An ODBC connection must be enabled for the corresponding database with which you wish to connect. To select a database table click the Modify button and configure your database link. If a given field is not found within a database table during verification, a dialog box will be displayed and focus will be set to the corresponding field.

Additional Settings within the Processing Tab

Confidence Levels

Many of the critical properties and features available within the Processing Tab have already been previously detailed within the field types above. One of the principal features within the

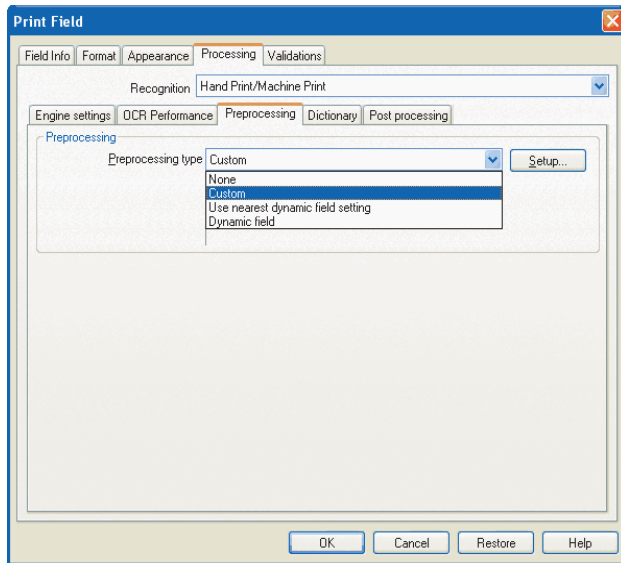


FIGURE 10. The custom option allows you to configure how incoming forms are optimized.

Engine settings sub-tab is the ability to change field Confidence Levels. Every character that is evaluated by TeleForm’s RecoFlex recognition engine is assigned a “confidence” level to indicate its certainty that the character it selected was correct. Numbers close to “0” indicate low confidence while numbers close to “100” indicate high confidence. The System level confidence of 80 is applied by default to all fields.

You can change the confidence level for Constrained Print Fields and Image Zones on the form by selecting the Use field confidence checkbox and typing a value in the Field Level box. By raising the confidence level the number of fields displayed for verification will increase, however likelihood that an invalid character will be accepted by the application will decrease. It is not recommend that you raise the confidence level higher than 90% unless you wish to review the majority of received fields during verification.

The default confidence levels have been optimized for the application and should only be adjusted sparingly. Fields such as credit card numbers or social security numbers may require a greater level of confidence if they are very critical to the overall application.

Preprocessing

The field preprocessing controls can be found within the Preprocessing sub-tab. This feature is not commonly used, but can however be quite useful. In the event that forms are being returned in poor quality, the preprocessing features available to TeleForm may be useful for cleaning up the form images. By selecting the custom option from the Preprocessing type drop-down and clicking Setup, you can configure how incoming form images are optimized. This option should only be used when images are consistently experiencing the similar problems.

The most efficient way to configure the preprocessing settings is to select a Dynamic field from each corner or area of your form. All other form fields should be configured to use the nearest dynamic field setting. With these features enabled, TeleForm will automatically determine how to optimize each field. Form fields will be automatically adjusted for contrast, speckling, and rotation. Characters and lines will also be softened or darkened for the best possible recognition.

The dynamic preprocessing settings are commonly enabled with forms that will be retrieved via fax or for forms that are likely to be photocopied.

Dictionary Files

Dictionary files are very useful in aiding TeleForm with recognition. Within a dictionary file you can list some or all of the expected values that may be appear within a corresponding field. Within a state dictionary file, for example, you may want to list the abbreviations and full names of every state. This dictionary could be reused within every state field you create. This dictionary file will assist with recognition of the underlying field; in addition you can restrict a field to the list of underlying values. In the event that a word entered cannot be found in the list it can be displayed for further review. To enable the review option, check the Review if word is forced, or ambiguous, or not found checkbox. To allow TeleForm to automatically force the best match click the Attempt to force best match check box.

Dictionary files are useful for streamlining the verification process. This option should be enabled as much as possible. Dictionary files are reusable and can be applied to as many forms as needed.

Additional Settings within the Post-Processing Tab

Default Values

There are instances where you will want to automatically populate a field in the event that it is not completed on your form. You can activate this feature by checking the Use this if field not filled option within the Post processing sub tab. This feature is very useful and should be used for fields that are required for export.

Within this subtab, you can also choose to capitalize all received data. To capitalize the content of the field, simply check the Convert to uppercase checkbox.

Job Maintenance

Optimizing a Job for Form Identification

For each incoming batch of images scanned or picked up from a directory, TeleForm can apply a unique set of rules for processing. These options are explored in greater detail within the TeleForm Administrator’s Guide. Some of the options are available at the job level as detailed ahead.

Within the Form ID/Attachments tab of the Job Maintenance window, users have the ability to either optimize the job for processing a form or set of forms, or force all received images

to be evaluated as a specific form type. This feature can be very useful when processing existing forms which are difficult to identify or can be easily confused with other templates.

A recognition set can be configured to a set of templates, which will be grouped together during processing. Recognition sets can be configured from the File dropdown with the Designer module.

Within the Form ID/Attachments tab a job can also be configured to link non-forms or attachments to the preceding form template. These attachments will be available for display during Verification. Capture Zones can be configured to display these attachments for manual entry. Point-and-Click-OCR may also be leveraged to capture data received on incoming attachments.

Non-forms are forms that can be identified using the set of templates available so that they are forcibly identified as a given form type. This feature can be configured within the Job Maintenance window.

Image Quality Control (Image QC)

Within the Review tab of the Job Maintenance window, you can enable Job QC for a given job. This feature gives the Verifier the ability to ensure images have been correctly scanned and identified. Job QC mode allows you to check the following aspects of a batch:

- Each image belongs in the batch
- The images are of sufficient quality
- Each image is correctly identified
- Each image is correctly oriented
- Attachments are properly linked to forms

As depicted within the image below, the Verifier will be given the ability to identify an image as another form type, identify attachments, and delete images. Image QC is an efficient tool for ensuring the quality of images and templates identified by TeleForm.

Job QC will automatically be triggered prior to form verification within the Verifier window. Images that are identified as another template type within Job QC will automatically be re-evaluated within the Reader module. After being re-evaluated within Reader these images will be available for verification.

Job QC should be used when additional certainty is needed to ensure that forms have been identified correctly. It is also very effective for use when managing received attachments.

Data Review

There are instances where you may want to ensure the accuracy of data that is being processed by TeleForm. In these cases, the Data Review feature within TeleForm can be activated. This feature can be enabled within the Review tab of the Job Maintenance window. Only fields, which have been flagged for Data Review or Double Key within the field properties window, will be displayed for review.

When all of the forms in a batch have been corrected and have the status “Evaluated OK,” the batch is ready for Data Review. Data Review allows a Verifier operator to quickly review field

values before the batch is committed. Data Review also includes the re-entry of fields that were marked for Double Key review.

Depending on how your forms are set up, certain fields on each form will be visited during Data Review. Keep in mind that a field must have the Data Review or Double Key constraint enabled within the field properties window, in order to be visited during Data Review.

In order to use Data Review in Verifier, you must log onto a Verifier workstation as a user who can perform the Data Review operation. You will only be able to perform Data Review if there is at least one batch in the system that requires this operation.

Since a Data Review QC mode includes both Data Review and any date entry fields that were configured to use the Double Key validation, you may see the Double Key Inconsistency dialog as you perform Data Review.

Accept: Clicking this button will overwrite the original value with the value typed during Data Review.

Cancel: Clicking this button will erase the values typed during Data Review and send you back to the data entry field to perform Double Key again.

The Data Review feature is very useful for ensuring the accuracy of data Verified by TeleForm, however this feature should only be used for testing purposes. The Double Key validation feature is useful for double-checking received data and may be used for highly sensitive and critical fields.

Form Optimization Recommendations

How a form is configured will have a direct impact on the quality of recognition when the form is processed within TeleForm. The following tips will be very useful for creating efficient forms. This list should be reviewed before finalizing or approving a form template.

Optimizing OCR / ICR Fields

TIP 1: Always limit the list of acceptable characters for a given field as much as possible.

Within the Processing tab of the field properties window, you can configure the list of expected characters for a given field. The list of expected characters should always be as narrow as possible. Expected characters can be alpha, numeric, or alphanumeric. In addition, you should specify whether received text will include hand- or machine-print. For special characters you can even identify which special characters you anticipate to receive.

For example, a telephone number will never include alphabetical characters and should be limited to numeric. You may also want to identify the dash within the list of expected characters if a template field was not used.

TIP 2: Use dictionary files where applicable to raise the readability for an individual field.

Dictionary files can be configured for both image zones and constrained print fields. When configured, TeleForm will use this list of values within a dictionary to aid in OCR / ICR of a given field. Taking the time to compile a list of expected values upfront will save a great deal of processing time down stream. Within the field

properties processing tab, the dictionary file settings can be found within the Dictionary tab.

Dictionary fields can be easily reused as they are imbedded within the application. State and month fields should always be configured to leverage this functionality.

TIP 3: Context-checking should be used where applicable to improve recognition of alphanumeric fields.

When the context-checking feature is enabled for a given field, TeleForm will utilize the surrounding characters to assist with the differentiation of character types for each individual character. This assists TeleForm when distinguishing between O's and 0's and I's and 1's.

For example, context checking should be utilized for street address fields since characters will often be grouped by character type. This will ensure that the street numbers are identified as numeric characters and the street address characters are currently identified if they are numeric characters.

TIP 4: Database lookup validations should be enabled where possible to raise the reliability of captured and verified data.

Within the Validations tab of field properties window, you have the option to enable database look-up validations. These validations will ensure that the captured values within any recognized field are available within a configured database table. If the value cannot be found within the corresponding database field, the field will be flagged and presented for Verification. This feature should only be enabled when all possible field values exist somewhere within an existing database table. If a table consists of some, but not all of the records, the validation should not be configured.

TIP 5: The Verifier tab order should be configured for every TeleForm form.

During verification, Verifiers must have the ability to easily tab between fields. The tab order should be configured for every form designed within TeleForm. To adjust tab order, select the Fields option from the Form dropdown and click on the Tab Order tab. Only the fields available within the Included window will be accessible to the Verifier. The tab order may be automatically configured by field location using the Sort drop down menu at the bottom of the Fields window.

The field tab order should be adjusted in a logical format. Form designers may want to work with the verifier group to help determine the order in which the fields should be configured.

Capture Zones

TIP 6: Capture Zones should be configured so that they are bypassed when the corresponding zone density is below a certain threshold.

When the underlying fields within a Capture Zone are optional, you should configure TeleForm to bypass the Capture Zone if corresponding zone density is below a certain threshold. Within the properties window for a given Capture Zone, you can calculate the zone density for the area on your form covered by

your Capture Zone. If the bypass property is enabled, the Capture Zone will not be displayed for Verification if the corresponding area is not filled. This will ensure that Verifiers do not have to manually bypass this window.

TIP 7: The tab order within every Capture Zone should be configured for Verification.

During verification, Verifiers must have the ability to easily tab between fields. The tab order should be configured for every form designed within TeleForm. To configure tab order click the tab order button and then click the blue box next to each field in the correct Verifier tab order. Configuring the tab order will ensure that Verifiers are able to easily navigate between fields.

TIP 8: Like fields should be group together using group boxes.

The interface designed for use within your Capture Zone should be user-friendly to the Verifiers. If there are many fields within your interface, group them together logically using group boxes. This will make the fields easier for the Verifiers. To draw a group box click on the Group Box button within the Capture Zone toolbar and draw a box around the corresponding area.

TIP 9: Leverage the Point-and-click OCR functionality availability within Capture Zones where possible.

The point-and-click OCR feature may be very useful at eliminating data entry within a Capture Zone. Please review the Capture Zone section above for additional information. When this feature is activated, the Verifier will have the ability to drag a box around an area for verification on the fly. Once the data within the zone is captured it will be populated within the corresponding field. Always test this feature using realistic test cases.

Existing Forms

TIP 10: Always ensure that you conduct significant testing of Existing Forms.

Any time you design a template for an Existing Form, always ensure that the form is tested significantly. In order to optimize the form, you may need to adjust Explicit IDs, ID blocks, and other form properties. You may want also to confirm that there are no additional outstanding forms in the field, which closely resemble the form template for your current form.

Ensure that your test environment is identical to your production environment. Similar-looking forms can pose a problem if your forms are not designed correctly. You could risk encountering problems later on if the form is not tested against all similar forms.

TIP 11: If you plan on building templates for many similar-looking forms or if TeleForm is having trouble recognizing your form template, try adjusting your template's form IDs and Explicit IDs.

TeleForm uses ID Objects and Explicit IDs to aid in recognition of existing forms. There should not be any enabled ID Objects in areas where there will be a great deal of writing. To disable an ID Object, ensure that objects are visible by selecting Show ID Objects from the View dropdown. Next double-click on the ID block and choose to disable the block.

If too many blocks are disabled, an Explicit ID will be required to properly identify the form template.

In the event that TeleForm is either having difficulty recognizing your form or TeleForm is having trouble differentiating between form templates, you will want to use an Explicit ID. Explicit IDs can use a label or block of text on your form to help to aid in form identification. This feature is commonly used and is often required for large volumes of similar looking forms. A form number or distinct titles are very useful as Explicit IDs.

TeleForm is very effective at distinguishing between nearly identical looking forms. Keep in mind however that you may need to try different Explicit IDs before you find one that works correctly. The key to finding an explicit ID that is consistently accurate is through testing. Existing forms always require significantly more testing than Traditional and VersiForms.

TIP 12: *If TeleForm is having difficulty recognizing a single form or a set of existing forms, a job can be optimized to enhance identification.*

Jobs can be customized for processing sets of forms that are scanned using Scan Station or picked up from an individual directory. Any job can be optimized for identification of an individual form or for a set of forms. Optimizing or restricting a job to a set of forms will assist TeleForm with form identification. Review the Job Maintenance section above for additional information regarding job configuration.

Multi-Page Forms

TIP 13: *Use page links, when designing forms that are longer than a single page front and back.*

Many of the forms that you will design will be more than a single page. It's recommended however that you redesign all multi-page forms within the TeleForm application. Page links should be used to ensure that individual form pages are properly linked. Page links are pre-filled fields that appear in the same location on every page. Each field on every page will contain the same set of alpha or numeric characters. When your forms are scanned in and

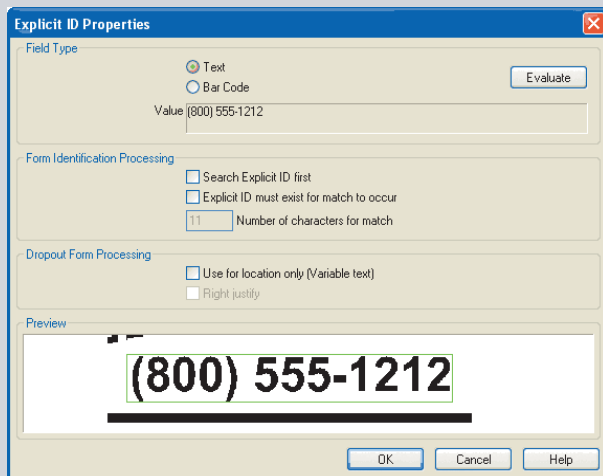


FIGURE 11. The block of text shown helps with form identification.

processed, TeleForm can use these fields to keep pages from the same form together within the application.

To identify a field as a page link, click on the field then select Page Link from the Object dropdown menu. The selected field will then bleed through on to every page within the form. Double-sided two-page forms often may not require page links, however significant testing is recommended.

Optional pages or sections and form packets should be also divided into individual templates. Each template section must contain a unique identifier that can be linked to information contained on the related pages. It is important to keep in mind that all received documents may be linked together on the backend. Work with engineers for the corresponding back-end systems to ensure that sufficient identifying information has been included on each section of your packet.

TIP 14: *Avoid creating Multi-page Existing Forms longer than two full pages front and back.*

Avoid creating Existing Form templates for forms that are longer than two pages front and back. When possible, these forms should be recreated within TeleForm. In the event that the redesign of these forms is not a possibility, additional care must be taken when scanning and processing. The Scan QC and Job QC functions allow the Verifiers to confirm that forms have been scanned correctly in order. These features should be leveraged to ensure that pages are not accidentally mismatched. Additional precautions can be built into your template using form scripting and Visual Basic. Work with a technical contact or liaison to find out what additional steps can be taken to avoid the risk of mismatching incoming forms.

TIP 15: *Leverage the Job QC functionality when processing either multi-page forms without page links or forms including attachments.*

Job QC allows verifiers to confirm that pages have been scanned in the correct order and that they have been correctly identified. This feature may be leveraged anytime order or identification is of special concern to a set of incoming forms. The Job QC section included within this document contains additional information on this feature. Once Job QC has been configured images can be reordered or manually identified.

TIP 16: *Existing Forms should be redesigned if you plan to receive the forms via electronic fax.*

Form images are often reformatted or modified when faxed. For this reason, Verity recommends that Existing Forms should not be received via electronic fax. Where possible these forms should be redesigned as Traditional or VersiForms.

When redesign of an Existing Form is not an option, the form should undergo significant testing using an array of fax types. If an individual fax number can be dedicated to an existing form type, a job should be configured to optimize or force identification of the form set or form template.

Verity and Intellectual Capital Management

Verity provides software that enables organizations to maximize the return on their intellectual capital investment. The company's intellectual capital management (ICM) solutions provide integrated search, classification, recommendation, monitoring and analytics across the real-time flow of enterprise information, along with self-service Q & A. In addition, Verity's business process management and content capture solutions activate information and accelerate its flow from person to person and between systems. Verity technology also serves as a core component of more than 260 applications from leading independent software vendors.

Around the world, more than 11,500 organizations of all sizes and types rely on Verity to manage their intellectual capital. In the private sector, these include ABB, AT&T, AXA, Bristol-Myers Squibb, Capgemini, Ernst & Young, Capital One, Cisco, Deloitte Consulting, EDGAR Online, Ford, GMAC, Hewlett-Packard, Kaiser Permanente, KPMG, Mayo Foundation, PricewaterhouseCoopers, SAP, UBS and Verizon. In the public sector, customers come from all levels and branches of government, including the U.S. General Services Administration, the U.S. Air Force and Army, the U.S. Department of Defense, including the armed forces units, and the U.S. Departments of Energy and Justice. Independent software vendors that integrate Verity include EMC Documentum, FileNet, Lotus, Oracle, Stellent and TIBCO.

Verity, Inc.

894 Ross Drive
Sunnyvale, CA 94089
t. 408.541.1500
f. 408.541.1600

Verity San Diego

3220 Executive Ridge Drive
Vista, CA 92081
t. 800.659.8755
f. 760.936.4800

Verity New York

230 W. 41st Street
12th Floor
New York, NY 10036
t. 646.366.9500
f. 646.366.9540

Verity Washington, DC

2941 Fairview Park Drive
8th Floor
Falls Church, VA 22042
t. 703.289.8800
f. 703.289.8840

Verity Chicago

200 South Wacker Drive
31st Floor
Chicago, IL 60606
t. 312.674.4525
f. 312.674.4571

Verity Boston

275 Grove Street, 4th Floor
Newton, MA 02466
t. 617.663.5716
f. 617.663.5382

Verity Canada

400 – 4th Avenue SW
Suite 2500
Calgary, AB T2P 0J4
Canada
t. 403.750.4000
f. 403.750.4100

Verity Australia

115 North Oatlands Road
Yarrambat, VIC 3091
Australia
t. 61.3.9436.2024
f. 61.3.9436.2078

Verity Benelux

Coltbaan 31
3439 NG Nieuwegein
The Netherlands
t. 31.30.669.2120
f. 31.30.662.2094

Verity China

18F, China Merchants Tower
No. 118 Jian Guo Road
Chao Yang District, Beijing 100022
P. R. China
t. +86-10-65675899
ext. 6091

Verity France

Les Collines de l'Arche
Immeuble Madeleine
D76 route de la Demi-Lune
92057 Paris La Défense Cedex
France
t. 33.1.41.49.0450
f. 33.1.40.89.0981

Verity Germany

Robert-Bosch-Strasse 9 „Europahaus“
D-64293 Darmstadt
Germany
t. 49.6151.86082.0
f. 49.6151.86082.20

Verity Italia

S.R.L.c/o Ernst & Young Via Cornaggia,
n°1020123 Milano
Italia
t. +39 050 770 157z
f. +39 050 770 157z

Verity Japan

15th floor, Cerulean Tower,
26-1 Sakuragaoka-cho, Shibuya-ku, Tokyo,
150-8512, Japan
t: 81.3.5456.5792
f: 81.3.5456.5733

Verity Mexico

Camelia 253-6
Colonia Florida
Alvaro Obregon
Mexico D.F.
01030 Mexico
t. 52.5.661.7126
f. 51.5.598.9526

Verity United Kingdom

Verity House
2 Heath Road
Weybridge
KT13 8AP
United Kingdom
t. 44.1932.844200
f. 44.1932.843823

www.verity.com

© 2005 Verity, Inc. All rights reserved. Verity®, the Verity logo, Ultraseek®, TOPIC®, KeyView®, TeleForm®, and Knowledge Organizer® are registered trademarks. Verity® Portal One™, Verity® LiquidOffice™, Verity® LiquidCapture™, Verity Profiler™ are trademarks of Verity, Inc., in the United States and numerous other countries. All other trademarks or symbols are those of their respective owners.

**Sales and Product Information**

software-sales@verity.com or 888.328.7335

Partnerships

partners@verity.com

Technical Support for Existing Customers

tech-support@verity.com

Verity, Inc.

894 Ross Drive, Sunnyvale, CA 94089 | t. 408.541.1500 | f. 408.541.1600 | www.verity.com