

Answer Sheet Scanning Requirements

Multiple answer sheets will be provided that must be processed. Following are general requirements that will apply to all answer sheets:

1. Program must accept multi-page TIFF images. It would be good if program can accept multipage PDF images also but not mandatory.
2. Multi-page document may include answer sheets of different types in the same file.
3. First step of scanning process must be identification of answer sheet template. If image cannot be identified it must be sent to separate folder for user to manually assign sheet using another workstation.
4. System must scale images based on registration marks (Rectangles in corners).
5. System must accept images that are skewed or misaligned.
6. System must have configuration file for modifying conditions for minimum fill % to detect a filled bubble. If no bubbles in the zone meet minimum fill %, then output is blank.
7. System must have configuration file for modifying condition for comparing two or more filled bubbles and selecting correct bubble. For example, user may define that selected bubble must be 10% more filled than other bubbles or it considers this as multiple bubbles. Multiple bubbles have output of *.
8. Each answer sheet will have text and bar-code information that must be processed. In all cases, data is both Bar code and text. System should read both and if they match, use the date. If barcode and text do not match, select the one that system is most confident in. System must provide method for sending sheets that have missing or questionable data in these zones to separate queue so that user on another machine can process these without stopping primary machine. Secondary process must allow for easy viewing of questionable areas and ability for user to manually input data. For example, if student ID is not read, screen should show that area of the answer sheet and provide area for user to type in student ID.
9. Final file output must be usable by eDoctrina. At completion of scanning, image must be converted to PDF with filename based on information in the image (District ID, Student ID and Test ID).

Some example TIFFs

[LaMond Qtr 3 2013 001.tif](#)

[Scanned from a Xerox multifunction device001.tif](#)

[Dewey SLO Qtr 3 #2 Redo 001.tif](#)

[Scanned from a Xerox multifunction device002.tif](#)

[Scanned from a Xerox multifunction device003.tif](#)

[Dewey SLO Qtr 3 Redo 001.tif](#)

[Scanned from a Xerox multifunction device001.tiff](#)

[Scanned from a Xerox multifunction device004.tif](#)

Config file format

```
{ name: 'FASTeST Small'
, darknessPercent: ''
, darknessDifferenceLevel: ''
, regions: { { name: "district_id", outputPosition: "0" type: "barCode&Text"
              , areas: { { type: "barCode", left: "10", top: "10", width:
"100", height: "10" }
                      , { type: "numbersText", left: "10", top: "5", width:
"100", height: "5" }
                    }
              , active: true|false
            }
      , { name: "test_id", outputPosition: "1", type: "barCode&Text"
        , areas: { { type: "barCode", left: "10", top: "30", width:
"100", height: "10"} }
        {"type":"numbersText","left":1307,"top":201,"width":26,"height":41 }
        , active: true|false
      }
      , { name: "sheetIdentifier"
        , type: "barCode&Text"
        , areas: { { type: "barCode", left: "10", top: "30", width:
"100", height: "10" }
                  , { type: "numbersText", left: "10", top: "5", width:
"100", height: "5" }
                }
      }
      , { name: "index_of_first_question"
        , type: "barCode&Text"
        , areas: { { type: "barCode", left: "10", top: "30", width:
"100", height: "10" }
                  , { type: "numbersText", left: "10", top: "5", width:
"100", height: "5" } }
        , active: true|false
      }
      , { name: "amout_of_questions"
        , type: "barCode&Text"
        , areas: { { type: "barCode", left: "10", top: "30", width:
```

```

"100", height: "10" }
    , { type: "numbersText", left: "10", top: "5", width:
"100", height: "5" } }
    , active: true|false
  }
  , { "name":"leftTopBlackBox",
    "type":"marker"
    , areas: { left: "10", top: "30", width: "100", height: "10" }
    , active: true|false
  }
  , { "name":"rightTopBlackBox"
    "type":"marker"
    , areas: { left: "10", top: "30", width: "100", height: "10" }
    , active: true|false
  }
  , { "name":"rightBottomBlackBox"
    , "type":"marker"
    , areas: { left: "10", top: "30", width: "100", height: "10" }
    , active: true|false
  }
  , { "name":"leftBottomBlackBox"
    , "type":"marker"
    , areas: { left: "10", top: "30", width: "100", height: "10" }
    , active: true|false
  }
  , { name: "answers", outputPosition: "4", indexOutputPosition:
"3", type: "bubblesRegions"
    , areas: [{ left: "10", top: "40", width: "100", height: "100"
      , lineHeight: "10", bubblesPerLine: "10"
      , bubble: {"x":248,"y":752,"width":62,"height":38}
    }
    , { left: "10", top: "40", width: "100", height: "100"
      , lineHeight: "10", bubblesPerLine: "10"
      , bubble: {"x":248,"y":752,"width":62,"height":38}
    }
    ]
    , active: true|false
  }
}
, additionalOutputData: { { outputPosition: 10, value: "%fileName%" }
  , { outputPosition: 11, value: "la-la-la" } }
, outputFileNameFormat:
"{district_id}_{test_id}_{student_uid}_{index_of_first_question}_{amout_of_qu
estions}"
}

```

Field definitions

List of sheet attributes

Name	Required	Description
name	no	Sheet name
regions	yes	List of regions

Name	Required	Description
additionalOutputData	no	Some additional data which must go into output file
outputFileNameFormat	yes	Format for output file name. Here {region_name} must be replaced with value taken from regarding region. For example: {district_id}_{test_id}_{student_uid}_{firstQuestionIndex}_{numberOfQuestions}

regions

List of sheet regions

Name	Required	Description
name	no	Region name
outputPosition	no	Position in output file
type	yes	Type of region. See below.
areas	yes	List of region areas
value	no	Value found in region. Only for sheetIdentifier region.
active	no	Active by default but could be deactivated by adding active: false

region.type

Name	Description
singleValue	Single value - such as district id or teacher name for example
bubblesRegion	Bubbles region
firstQuestionIndex	Index of first question on the page
numberOfQuestions	Number of questions on the page
sheetIdentifier	Answer sheet identifier - we need to search for this on the page and compare
leftTopBlackBox, rightTopBlackBox, leftBottomBlackBox, rightBottomBlackBox	Black boxes positions

region.areas

List of region areas

Name	Required	Description
left, top	yes	Coordinates of left top position
width, height	yes	Dimensions of the area
type	no	Type of area data. See below.
lineHeight	no	Height of one bubbles line. Only for bubblesRegion region.
bubblesPerLine	no	Amount of bubble per line. Only for bubblesRegion region.

region.area.type

Name	Description
barcode	Area contain barcode
text	Area contain text
capitalText	Area contain text in capital letters
numbersText	Area contain text with number only

additionalOutputData

Set of additional hardcoded values in output.

Output Positions

All output position indexes starts from 1 (not 0)

Output file format

- one output CSV file per one image in input file
- no column headers
- everything goes to the same output folder
- file name must be generated based on answer sheet config + adding CSV or TIF extension. If such file already exists - you must add index (1, 2, 3, etc...). So file name will be like ORIGINALNAME-INDEX.CSV

Basically all you need is to use OutputPosition attribute from config file. In our example output file be like below:

```
district_id,test_id,,0,answer[0],,,,,,fileName,la-la-la
district_id,test_id,,1,answer[1],,,,,,fileName,la-la-la
...
district_id,test_id,,10,answer[10],,,,,,fileName,la-la-la
```

- there is no value in position 2 because we don't have any areas which has OutputPosition set to 2.
- there is big gap between 4 and 9 because there is no settings for them but exists settings for position 10
- answers occupy two positions in output file - number of question and answer value. So that's why there are two settings in config file outputPosition and indexOutputPosition
- answer could be multiple, in this case it must be represented as values separated by |, for example 1|5

```
68,5789,00142674,0,1,,,,,input01.tif,la-la-la
68,5789,00142674,6,1|2|3,,,,,input01.tif,la-la-la
```

- answer could be on multiple lines, in this case it must be represented as values selected on first line separated by ~ with values from second line, for example 1|2~5|6. ~ must always present for such answer sheets.

```
68,5789,00142674,0,1|2|3~0|7|8|9,,,,,input01.tif,la-la-la
68,5789,00142674,6,0~0,,,,,input01.tif,la-la-la
```

Answer sheet explained

Student: Alexis Akulis
710080586

Teacher: LaMond Gail
31046



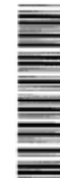
USE PENCIL ONLY. MARK EACH BUBBLE FULLY. CLEARLY ERASE ALL MISTAKES. Poor ☒ ☒ ☒ Good ☐

Health Gr 6 Post Test

--

- 1 ☒ ☐ ☐ ☐
- 2 ☐ ☐ ☐ ☒
- 3 ☐ ☒ ☐ ☐
- 4 ☐ ☒ ☐ ☐
- 5 ☐ ☐ ☒ ☐
- 6 ☐ ☐ ☒ ☐
- 7 ☐ ☐ ☒ ☐
- 8 ☒ ☐ ☐ ☐
- 9 ☐ ☒ ☐ ☐
- 10 ☐ ☐ ☒ ☐
- 11 ☐ ☐ ☒ ☐
- 12 ☐ ☒ ☐ ☐
- 13 ☐ ☒ ☐ ☐
- 14 ☐ ☐ ☒ ☐
- 15 ☒ ☐ ☐ ☐
- 16 ☐ ☒ ☐ ☐
- 17 ☐ ☐ ☒ ☐
- 18 ☒ ☐ ☐ ☐
- 19 ☐ ☐ ☐ ☒
- 20 ☐ ☐ ☒ ☐
- | | |
|---|-----------------------|
| 21 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> | <input type="radio"/> |
| 22 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> | <input type="radio"/> |
| 23 <input type="radio"/> <input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> | <input type="radio"/> |
| 24 <input type="radio"/> <input type="radio"/> <input checked="" type="radio"/> <input type="radio"/> | <input type="radio"/> |

SMALL



372



24



1



Test Files Will Be In FTP Folder: 72.45.148.242

kwcsdadmin / safe2pas