Background

The CAD Reader add-on to DotImage can give you all of the geometry of a CAD Drawing. You must provide a class that will receive the pieces and process them. One easy thing to do is to just extract all of the text for indexing purposes. Here is the code you need to do that:

First, we declare a class that will hold a piece of extracted text

```
lass ExtractedText { private int _page; private string _text; private Point _pos; public int Page { get { return _page; } } public string Text { get { return _text; } } public Point Pos { get { return _pos; } } public ExtractedText(int page, string text, Point pos) { _page = page; _text = text; _pos = pos; } public override string ToString() { return Text + " (Page: " + Page + " / PosX: " + Pos.X + " / PosY: " + Pos.Y + ")"; } }
```

Next, create a geometry handler which will be called with each piece of the CAD document. In it, if we get text, we raise an event.

```
lass TextExtractorGeometryHandler : GeometryHandler { public delegate void OnTextFoundHandler(ExtractedText txt); public event OnTextFoundHandler OnTextFound; private int _page; public TextExtractorGeometryHandler() { _page = 0; } // this gets called inside of DwgDecoder.Distill() // when a new page is starting public override bool OnBeginPage(int page, SizeF pageSizeInInches) { // remember the page _page = page; return base.OnBeginPage(page, pageSizeInInches); } // this gets called inside of DwgDecoder.Distill() when text is found public override DwgReduction OnBeginText( int shapeId, double[] position, double[] directionVector, double[] upVector, string msg, bool raw, DwgTextStyle textStyle, double[] extrusionVector) { // if there is a hander for the event if (OnTextFound != null) { // tell the event handler about the text // (The page, text and position of the text) OnTextFound(new ExtractedText( _page, msg, new Point((int)position[0], (int)position[1])) ); } // don't turn the text into vectors return DwgReduction.DontReduce; } }
```

To use these classes, create a TextExtractorGeometryHandler, handle the text found event and then call Distill on a DwgDecoder with your CAD File. Here is the code:

```
rivate List<ExtractedText> _textList = new List<ExtractedText>(); private void ExtractText(string file) { if (file == null) return; // Clear list _textList.Clear(); // Create a text handler TextExtractorGeometryHandler txtHandler = new TextExtractorGeometryHandler(); // handle the event when text is found txtHandler.OnTextFound += new TextExtractorGeometryHandler.OnTextFoundHandler( txtHandler_OnTextFound ); // create a CAD reader (reads DWG and DXF) DwgDecoder dwg = new DwgDecoder(); // Open the file using (Stream s = File.OpenRead(file)) { // read all of the geometry from the file and call the handler // the txtHandler only cares about text // and will call the event (OnTextFound) when // it finds text dwg.Distill(s, txtHandler, null); } // _textList has a list of ExtractedText
```

```
objects // with the text that was found } // the event handler that gets called for each
piece of text found void txtHandler_OnTextFound(ExtractedText txt) { _textList.Add(txt); }
```

Implementation Example

There is a sample implementation of this code as a console application in the attachment: CadExtractTextConsole.zip

Original Article

Q10305 - HOWTO: How to extract text from a CAD Drawing

Last Update

2026-11-04 - TD - (Fixed code fences and headings)

Atalasoft Knowledge Base
https://www.atalasoft.com/kb2/KB/50171/HOWTO-How-to-extract-text-from-a-CAD...