

TI@SOFT

Producción de software a medida

Consultas Sunat

2 MAYO, 2016 DE WILBERTVIL

CONSULTA DIRECTA DE RUC, SIN CAPTCHA Y SIN ESTAR INGRESANDO A LA WEB DE LA SUNAT

Después de tantos intentos pude solucionar los problemas del captcha, hay que ser perseverantes y no caer en el intento, la innovación es importante para los programadores, solo es investigación, dedicación y mucho tiempo, así construimos nuevos conocimientos y que sirva de base para los que están empezando.

Las consultas de ruc vía Sunat, se complico desde que habilitaron el “famoso” captcha, antes la consulta solo era utilizando el siguiente url:

<http://www.sunat.gob.pe/w/wapS01Alias?ruc=xRuc> (<http://www.sunat.gob.pe/w/wapS01Alias?ruc=xRuc>)

Donde xRuc es el número de ruc a consultar, se recibía el parámetro vía post, y se obtenía el archivo XML, a partir de este archivo, se jalaba todos los datos del ruc. **Ahora la cosa cambio, el url actual es:**

<http://www.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias> (<http://www.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias>)

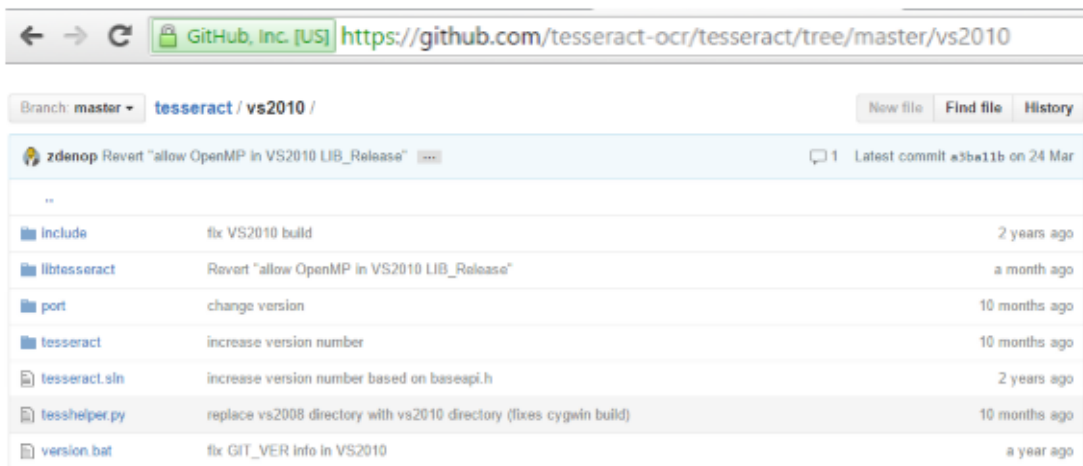


(<http://www.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias>)

Aquí les dejo la secuencia para evitar este paso.

1.- Primero utiliza un software OCR, que convierta la imagen del captcha a un texto, yo utilizo el Tesseract open source de mayor precisión, mejorando su algoritmo leptónico, la traducción es buena. Busca en la web el Tesseract-OCR. los descargas y lo instalas.

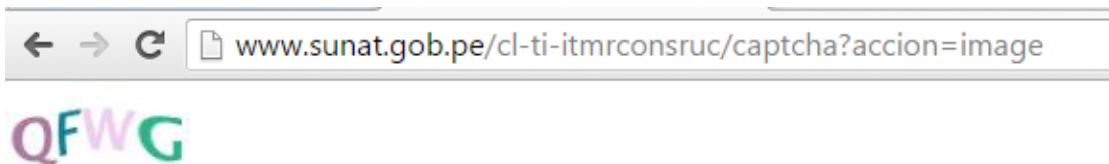
Pero si quieres mejorar el algoritmo descarga desde:



Encuentras para varios lenguajes el que utilizo es de visual studio

2.- Ahora como bajo la imagen del captcha desde la página de la Sunat.

Escribe exactamente igual al siguiente URL



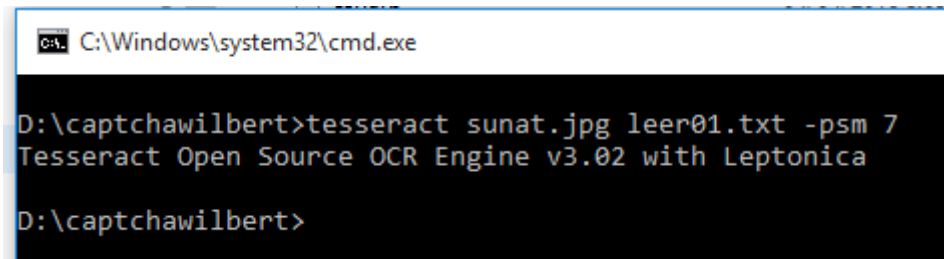
3.- Descarga la imagen, dependiendo del lenguaje de programación que utilizas usa el comando que descarga archivos desde la web. En csharp utilizo el DownloadFile.

Nota. Los tiempos de proceso influyen en la consulta, has que tu código sea lo más rápido posible.

Para que el proceso se mas rápido yo utilizo APIS, ej. En csharp

```
[DllImport("urlmon.dll", CharSet = CharSet.Auto, SetLastError = true)]
static extern Int32 URLDownloadToFile(.....
```

4.- Una vez que descargas la imagen, **lo conviertes a texto, cómo?**, ejecutando el tesseract, desde la línea de comandos, suponga que la imagen lo descargaste con sunat.jpg

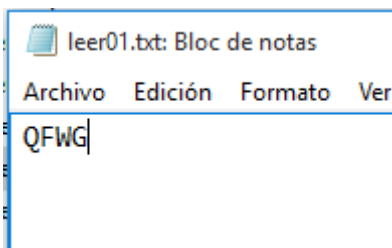


```
C:\Windows\system32\cmd.exe

D:\captchawilbert>tesseract sunat.jpg leer01.txt -psm 7
Tesseract Open Source OCR Engine v3.02 with Leptonica

D:\captchawilbert>
```

La ejecución del tesseract crea el archivo leer01.txt, con el siguiente contenido



No olvides trabajar en la ruta donde están tus archivos. Esta instrucción crea el archivo leer01.txt, archivo plano con el contenido QFWG, el captcha traducido en texto. Para hacer este proceso supongo que en tu programa no estarás pensando salir a la consola del cmd para ejecutar el tesseract y....., se pierde tiempo, yo utilizo:

- **En csharp**

```
Process proc = new Process();
```

- **En excel**

```
eje.Run "d:\excelruc\tesseract.exe d:\excelruc\sunat.jpg d:\excelruc\captcha -psm 7", 0, True
```

- **En visual Basic**

```
Dim ejecutar As Object
```

```
ejecutar = Microsoft.VisualBasic.CreateObject("WScript.Shell")
```

```
ejecutar.Run("d:\busruc\tesseract.exe d:\busruc\sunat.jpg d:\busruc\captcha -psm 7", 0, True)
```

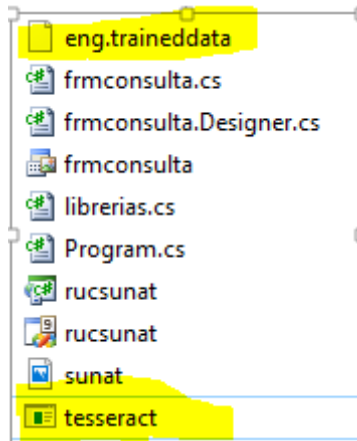
- **En visual fox**

```
baja = CreateObject("WScript.Shell")
```

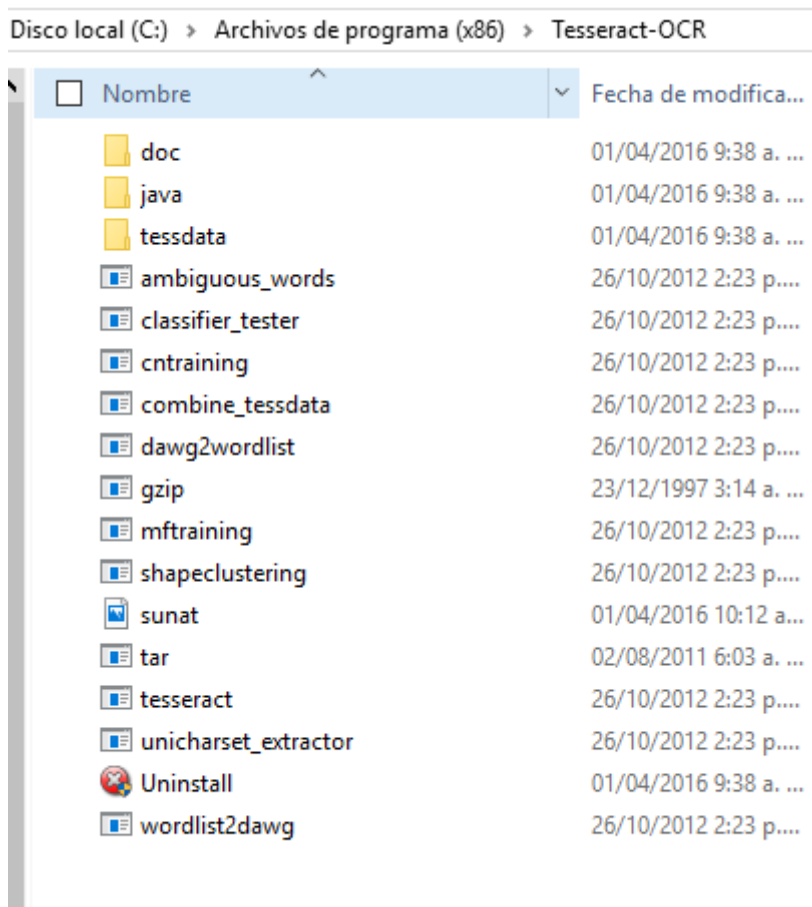
```
baja.Run("e:\ocr\tesseract.exe e:\ocr\captcha.gif e:\ocr\captcha -psm 7", 2, .T.)
```

```
DECLARE LONG URLDownloadToFile IN "urlmon";
```

Nota: solo se requieren dos archivos para ejecutar el tesseract, estos archivos deben estar en la ruta de tu sistema.



Los archivos son el eng.traineddata y el tesseract, que lo puedes copiar de la carpeta donde instalaste el software OCR.



5.- Ahora el texto traducido **lo obtienes en una variable.**

```
string dame = "";
dame = leecaptcha();
```

La función leecaptcha, recoge el contenido del archivo leer01.txt, utiliza una instrucción que lea un archivo texto, y lea el contenido de la línea, en csharp lo trabajo así

```
FileStream stream = new FileStream(@"d:\captchawilbert\leer01.txt", FileMode.Open, FileAccess.Read);
StreamReader reader1 = new StreamReader(stream);
```

Y lo capturo con la siguiente instrucción.

```
string resultado = "";
resultado = reader1.ReadLine().Trim();
```

La función devuelve la variable resultado

6.- Ahora para ejecutar el URL de la consulta utiliza.

[http://www.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias?accion=consPorRuc&nroRuc=xruc
&codigo=dame &tipdoc=1](http://www.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias?accion=consPorRuc&nroRuc=xruc&codigo=dame &tipdoc=1)

Donde xruc es el número de ruc a consultar y dame es el captcha traducido.

7.- Consume la web, descargando el URL en un archivo, yo utilizo el DownloadFile

```
Ejecuta.DownloadFile(URI, @"d:\captchawilbert\archiwil.txt");
```

8.- Bingo, ya tienes el XML en archivo plano, recorre con un bucle el txt y captura en variables los datos del número de ruc.

Ej. Para capturar la razón social, código en csharp

```
linea = reader.ReadLine();

while (linea != null) // mientras exista contenido
{
    .....
    .....
    linea = reader.ReadLine();
    ini = linea.Contains("N&uacute;mero de RUC:");
    if (ini == true)
    {
        linea = reader.ReadLine().TrimEnd().TrimStart();
        int pi = "<td class='bg' colspan=3>".Length + 14;
        int pf = linea.Trim().Length - pi;
        rs = linea.Substring(pi, pf - 5);
    }
}
```

Que tal, “el famoso captcha” ya es vulnerable, para los programadores no hay nada difícil, para todos los que desarrollamos software esta forma de consumir web nos ayudara a muchos, para registrar o consultar clientes, proveedores y otros procesos.

El software lo tengo desarrollado en:

- Excel (vba)

- Visual fox
- Visual Basic
- Csharp

Y bueno con esto, para los que también desarrollamos sobre web, php, aspx, java y otros lenguajes, se hace más sencillo, pero aquí está la base.

Acá les muestro un fragmento de código, de algunos procedimientos en visual basic, para la muestra un botón.

```
Option Explicit On
Imports System.IO
Imports System.Threading
Imports Microsoft.VisualBasic
Imports System.Text.RegularExpressions

Public Class Form1
    Dim strURL As String
    Dim LocalFilePath As String
    Dim DownloadStatus As Long
    Dim linea As String
    Dim no As String
    Dim xcodigo As String, bruc As String, xres As String
    Dim rs As String, fecha As String, tipo As String, est As String, cond As String, dir As String, tel As String

    Private Declare Function URLDownloadToFile Lib "urlmon" Alias "URLDownloadToFileA" (ByVal pCaller As Long, ByVal szURL

Sub bajaimagen()
    System.IO.File.Delete("d:\busruc\captcha.txt")
    Dim eje As Object
    eje = Microsoft.VisualBasic.CreateObject("WScript.Shell")
    Dim waitOnReturn As Boolean : waitOnReturn = True
    Dim windowStyle As Integer : windowStyle = 1
    strURL = "http://www.sunat.gob.pe/cl-ti-itmrconstruc/captcha?accion=imagen"
    DownloadStatus = URLDownloadToFile(0, "http://www.sunat.gob.pe/cl-ti-itmrconstruc/captcha?accion=imagen", '
    *****

Dim r As String
r = valida(xcodigo)
If (r = "error") Then
    bajaimagen() ' recursivo, si hay error repite la busqueda
Else
    strURL = "http://www.sunat.gob.pe/cl-ti-itmrconstruc/icrS00Alias?accion=consPorRuc&nroRuc="
    LocalFilePath = "d:\busruc\archiwil.txt"
    DownloadStatus = URLDownloadToFile(0, "http://www.sunat.gob.pe/cl-ti-itmrconstruc/icrS00Ali
    'DownloadStatus = URLDownloadToFile(0, strURL, LocalFilePath, 0, 0)
    muestradatos()
    *****

Sub muestradatos()
    Dim fn As String = "d:\busruc\archiwil.txt"
    Dim reader As New System.IO.StreamReader(fn)
    Dim f As Integer
    f = File.ReadAllLines("d:\busruc\archiwil.txt").Length.ToString()
    If (f <= 124) Then
        MsgBox("Error, no existe")
    Else
        linea = reader.ReadLine
        f = 0
        Do While (f <= 300)
            Dim ini As Boolean
            Dim pi As Integer, pf As Integer
            linea = reader.ReadLine()
            *****
        End While
    End If
End Sub
```

Y acá les dejo consulta ruc en vídeo.

Consulta ruc sin captcha



Consulta en Excel con VBA

E6 : [X] [✓] [fx] 20527882266

	A	B	C	D	E	F	G	H
1								
2								
3								
4								
5								
6								
7								
8								
9								
10								
11								
12								
13								
14								
15								
16								
17								
18								
19								
20								

Sunat

Consulta de Ruc

Ingrese Ruc: 20527882266

Razon Social: TABLACENTRO E.I.R.L.

Dirección: AV. PROL. AV. DE LA CULTURA NRO. 3044 INT. B (4TO Y 5TO PARADERO COSTADO INKA INVERSIÓN) CUSCO

Teléfono: / 974231465 / 984354138

Condición: HABIDO

Estado: ACTIVO

Fecha Inicio: 06/01/2007

Tipo Contribuyente: EMPRESA INDIVIDUAL DE RESP. LTDA

Copyright© - Todos los derechos reservados
Lic. Wilbert Vilcahuaman Mamani. - Analista programador de sistemas
Email: wilbert.vil@gmail.com
Tlf: 984-716936

Consulta en Visual Fox

Microsoft Visual FoxPro

Consulta Ruc

Consulta RUC:

Ingrese RUC: **20510069251** **Consulta** **Nuevo**

RUC: **20510069251**

Razón Social: **DH EMPRESAS PERU S.A.**

Dirección: **AV. MANUEL OLGUIN NRO. 211 INT. 504 LIMA**

Tipo Contribuyente: **SOCIEDAD ANONIMA**

Condición: **HABIDO**

Estado: **ACTIVO**

Teléfono: **7129200 / 993503441**

Todos los derechos reservados (C) Copyright Lic. Wilbert Vilcahuaman Mamani
Analista programador de sistemas
Email: wilbert.vil@gmail.com
Tlf: 984-716936 084-274048

Consulta en Visual Basic

Consulta Sunat

Búsqueda Sunat

Consulta

Ruc: **10239382385** **Buscar** **Nuevo**

Razon Social: **VILCAHUAMAN MAMANI WILBERT**

Dirección: **JR. LIMA NRO. K INT. 26 URB. AMADEO REPETO CUSCO**

Teléfono: **274048 / 984716936**

Fecha Ini. Act.: **21/03/1995**

Tipo Contribuyente: **PERSONA NATURAL SIN NEGOCIO**

Estado: **ACTIVO**

Condición: **HABIDO**

Copyright() - Todos los derechos reservados
Lic. Wilbert Vilcahuaman Mamani
Analista programador de sistemas
Tlf. 984-716936
Email: wilbert.vil@gmail.com

Cerrar

Consulta con Csharp

The screenshot shows a web application titled "Búsqueda Sunat". It features a search bar with the RUC "20493020618" and buttons for "Buscar" and "Nuevo". Below the search bar, there are several fields displaying company information:

Razon Social:	TIENDAS PERUANAS SA
Dirección:	AV. AVIACION NRO. 2405 URB. SAN BORJA LIMA
Teléfono:	6170800 / 7001500 / 993765359 /
Fecha Inicio Act.:	02/07/2008
Tipo Contribuyente:	SOCIEDAD ANONIMA
Estado:	ACTIVO
Condición:	HABIDO

On the right side of the form, there is a copyright notice and contact information for Lic. Wilbert Vilcahuaman Mamani, including his title as a system analyst, email (wilbert.vil@gmail.com), and phone number (984-716936). A "Cerrar" button is located at the bottom right of the form.

Bueno a pedido de los seguidores, estas son las nuevas direcciones, que cambio la Sunat

Consulta RUC

<http://e-consultaruc.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias> (<http://e-consultaruc.sunat.gob.pe/cl-ti-itmrconsruc/jcrS00Alias>)

Dirección para bajar la imagen del Captcha

<http://e-consultaruc.sunat.gob.pe/cl-ti-itmrconsruc/captcha?accion=image> (<http://e-consultaruc.sunat.gob.pe/cl-ti-itmrconsruc/captcha?accion=image>)

Sigue la secuencia de los pasos, con estas nuevas direcciones

Soy analista programador con más de 15 años de experiencia, desarrollo software a medida, tengo una mala experiencia de compartir código en la web, porque los usuarios que descargan se los hacen como si fuera su propiedad intelectual y hasta lo venden.

Consultas.

Wilbert Vilcahuaman Mamani
Analista programador de sistemas
Cel. 984716936
email: wilbert.vil@gmail.com



Publicado en [Consultas Sunat](#), [Manuales](#) | [10 comentarios](#)

[Crea un blog o un sitio web gratuitos con WordPress.com.](#)