In [1]:
```
import sys
import cv2
import numpy as np
from matplotlib import pyplot as plt
from matplotlib.patches import Rectangle
sys.path.append('mask_bracket')
import params

%matplotlib inline
```

```
Using TensorFlow backend.
```

In [2]:
```
plt.rcParams["axes.edgecolor"] = "black"
plt.rcParams["axes.linewidth"] = 1

def show_image(image, title=None):
    channels = image.shape[2] if len(image.shape) == 3 else 1
    if channels == 3:
        image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
    oldfigsize = plt.rcParams["figure.figsize"]
    plt.rcParams["figure.figsize"] = (5,4)
    plt.autoscale(enable=True, axis='both', tight=False)
    plt.axis('off')
    plt.imshow(image, aspect='equal', interpolation='gaussian', cmap='gray'
    plt.show()
    plt.rcParams["figure.figsize"] = oldfigsize
```

Using our trained UNET model to highlight brackets. Wrapped it in highlight_brackets function

Train script is located at mask_bracket/train.py

In [3]:
```
model = params.model_factory()
model.load_weights('mask_bracket/baseline.h5')
```

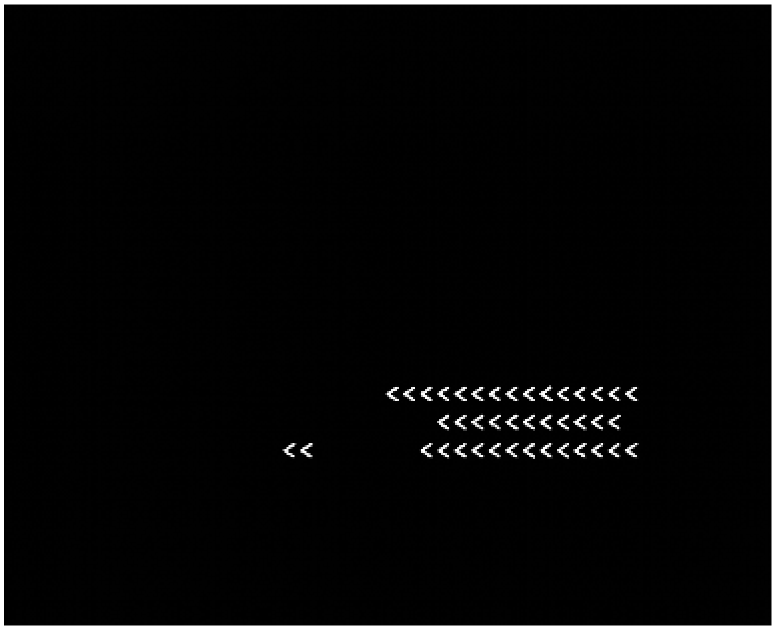
In [4]:
```
def highlight_brackets(img):
  input_size = params.input_size
  orig_height, orig_width = img.shape
  img = cv2.resize(img, (input_size, input_size))
  img = np.array(img, np.float32) / 255
  preds = model.predict(img.reshape(1,input_size,input_size,1))
  preds = np.squeeze(preds, axis=3)
  pred = preds[0]
  pred = cv2.resize(pred, (orig_width, orig_height))
  res = (pred >= 0.5).astype(int)*255
  return res.astype("uint8")
```

In [5]:
```
path = "sample6.jpg"
```

In [6]:
```
original_image = cv2.imread(path)
original_image = cv2.cvtColor(original_image, cv2.COLOR_BGR2RGB)
image = cv2.cvtColor(original_image, cv2.COLOR_BGR2GRAY)
image = params.simplify_image(image)
```

Getting masks for brackets

In [7]:
```
highlighted = highlight_brackets(image)
show_image(highlighted)
```



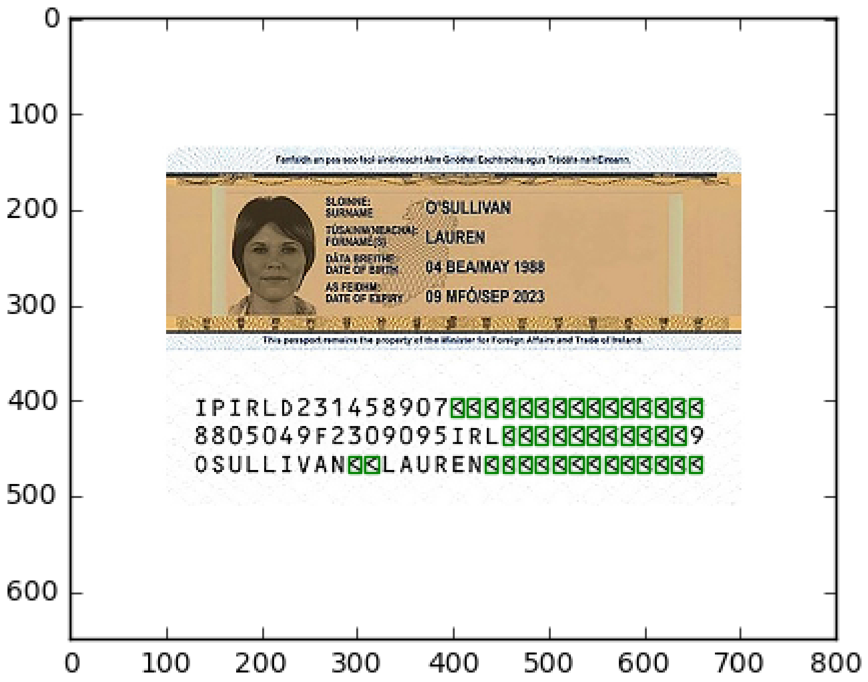
Then, it will be easy to find boundaries with cv2.findContours

In [8]:
```
def get_boundaries(imgray):
    _, contours, _  = cv2.findContours(imgray, cv2.RETR_LIST, cv2.CHAIN_APP
    boxes = []
    for cords in contours:
        cords = np.squeeze(cords)
        if len(cords.shape)>1:
            x = [a[0] for a in cords]
            y = [a[1] for a in cords]
            minx, maxx = np.min(x), np.max(x)
            miny, maxy = np.min(y), np.max(y)
            boxes.append((minx,miny,maxx,maxy))
    return boxes

def draw_boundaries(image, boxes):
    fig,ax = plt.subplots(1)
    ax.imshow(image)
    for (minx,miny,maxx,maxy) in boxes:
        rect = Rectangle((minx-1,miny-3),maxx-minx+1,maxy-miny+3,linewidth
        ax.add_patch(rect)
```

Let's see the result for this step

In [9]:
```
boxes = get_boundaries(highlighted)
draw_boundaries(original_image, boxes)
```



Fanfaidh an pas seo faoi úinéireacht Aire Gnóthaí Eachtracha agus Trádála na hÉireann.

SLOINNE: SURNAME O'SULLIVAN
TÚSAINM/NEACHA: FORNAME(S) LAUREN
DÁTA BREITHE: DATE OF BIRTH 04 BEA/MAY 1988
AS FEIDHM: DATE OF EXPIRY 09 MFÓ/SEP 2023

This passport remains the property of the Minister for Foreign Affairs and Trade of Ireland.

```
IPIRLD231458907<<<<<<<<<<<<<<<
8805049F2309095IRL<<<<<<<<<<<9
OSULLIVAN<<LAUREN<<<<<<<<<<<<<
```

Converting image to united lines

In [10]:
```
def get_lines(img):
    gradX = cv2.morphologyEx(img, cv2.MORPH_CLOSE, cv2.getStructuringElemen
    return cv2.threshold(gradX, 0, 255, cv2.THRESH_BINARY | cv2.THRESH_OTSU
```

In [12]:
```
lines = get_lines(image)
show_image(lines)
```

Finally searching for rows with brackets. Let's fine all rectangles, and ensure it contains at least one bracket

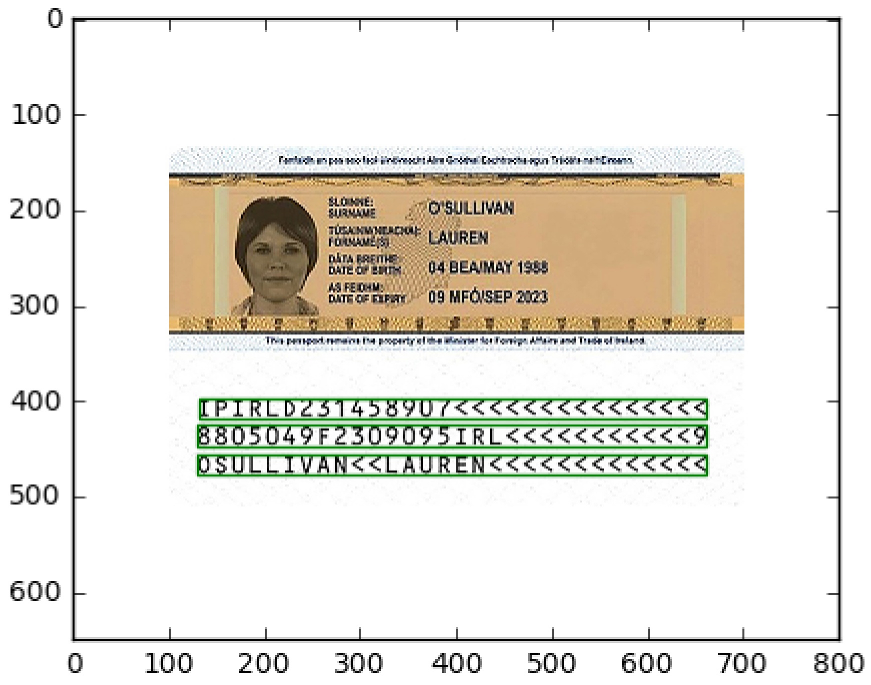
In [13]:
```
def get_bracket_lines(lines, boxes):

    def region_includes(region, box):
        #return region[0]<=box[0] and region[2]>=box[2] and region[1]<=box[
        return region[1]<=box[1] and region[3]>=box[3]

    _, line_contours, _  = cv2.findContours(lines, cv2.RETR_LIST, cv2.CHAIN
    brackets_lines = []
    for cords in line_contours:
        cords = np.squeeze(cords)
        if len(cords.shape) > 1:
            x = [a[0] for a in cords]
            y = [a[1] for a in cords]
            minx, maxx = np.min(x), np.max(x)
            miny, maxy = np.min(y), np.max(y)
            if (maxx - minx) > 100:
                region = (minx,miny,maxx,maxy)
                for box in boxes:
                    if region_includes(region, box):
                        brackets_lines.append(region)
                        break
    return brackets_lines
```

Showing the final result

In [14]:
```
brackets_lines = get_bracket_lines(lines, boxes)
draw_boundaries(original_image, brackets_lines)
```


0
100
200
300
400
500
600
0
100
200
300
400
500
600
700
800

Fanfaidh an pas seo faoi úinéireacht Aire Gnóthaí Eachtracha agus Trádála na hÉireann.

SLOINNE:
SURNAME
O'SULLIVAN

TÚSAINM/NEACHA:
FORNAME(S)
LAUREN

DÁTA BREITHE:
DATE OF BIRTH
04 BEA/MAY 1988

AS FEIDHM:
DATE OF EXPIRY
09 MFÓ/SEP 2023

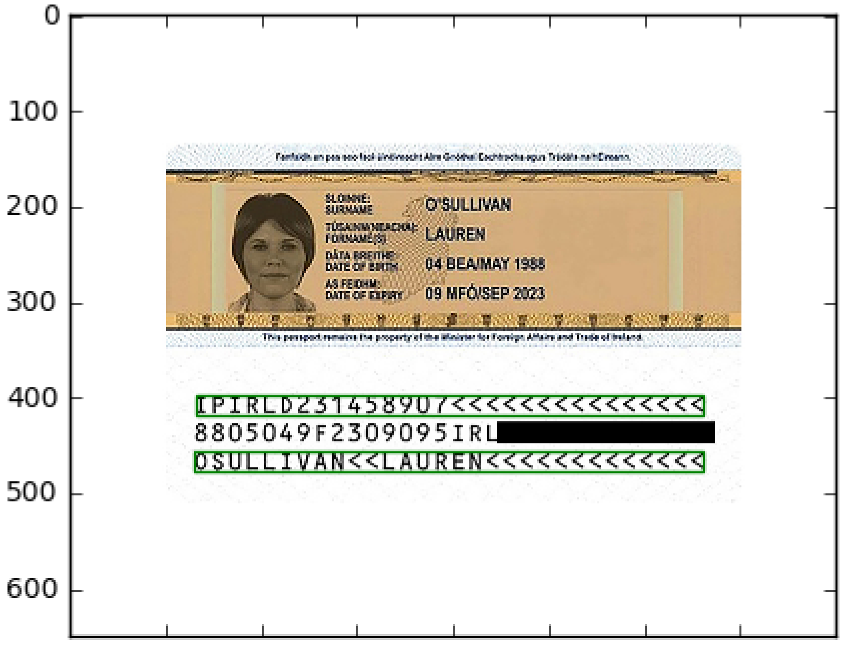
This passport remains the property of the Minister for Foreign Affairs and Trade of Ireland.

```
IPIRLD231458907<<<<<<<<<<<<<<<
8805049F2309095IRL<<<<<<<<<<<9
OSULLIVAN<<LAUREN<<<<<<<<<<<<<
```

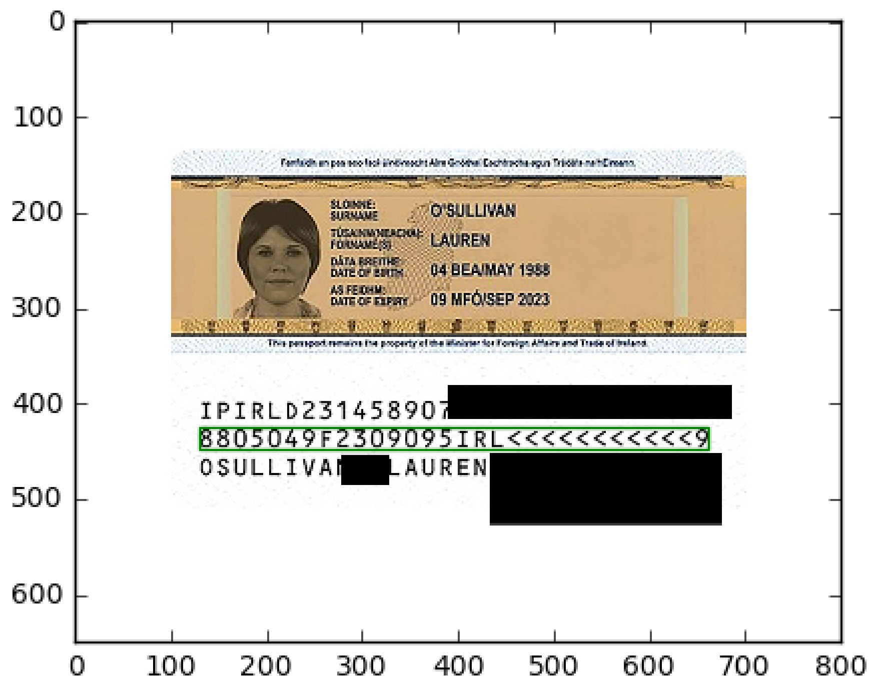
Let's test on something else...

In [15]:
```
def draw_mrz_lines(path):
    original_image = cv2.imread(path)
    original_image = cv2.cvtColor(original_image, cv2.COLOR_BGR2RGB)
    image = cv2.cvtColor(original_image, cv2.COLOR_BGR2GRAY)
    image = params.simplify_image(image)
    highlighted = highlight_brackets(image)
    boxes = get_boundaries(highlighted)
    lines = get_lines(image)
    brackets_lines = get_bracket_lines(lines, boxes)
    draw_boundaries(original_image, brackets_lines)
```

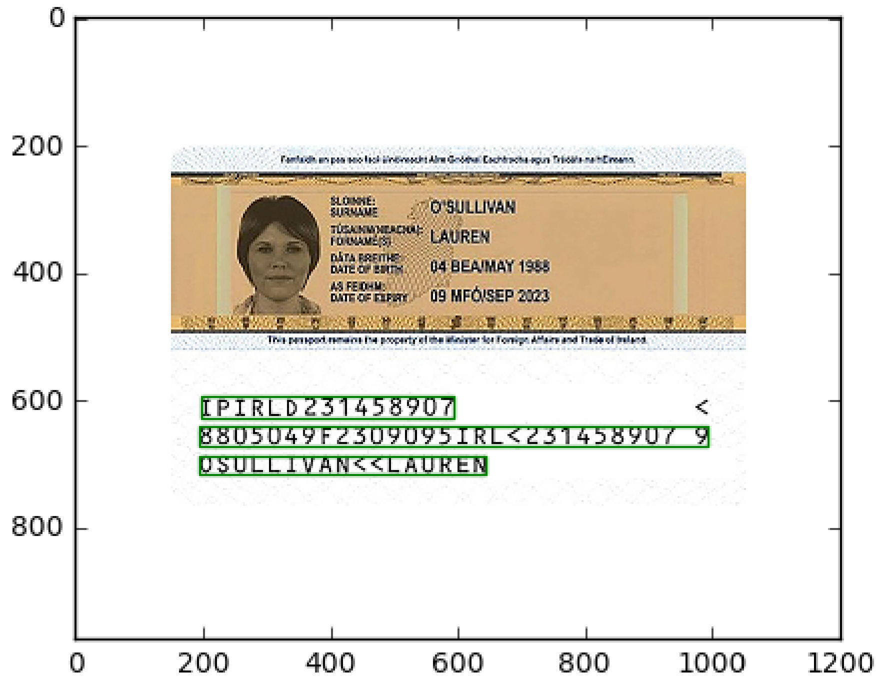
In [16]:
```
draw_mrz_lines("sample62lines.png")
```



In [17]:
```
draw_mrz_lines("sample6oneline.png")
```

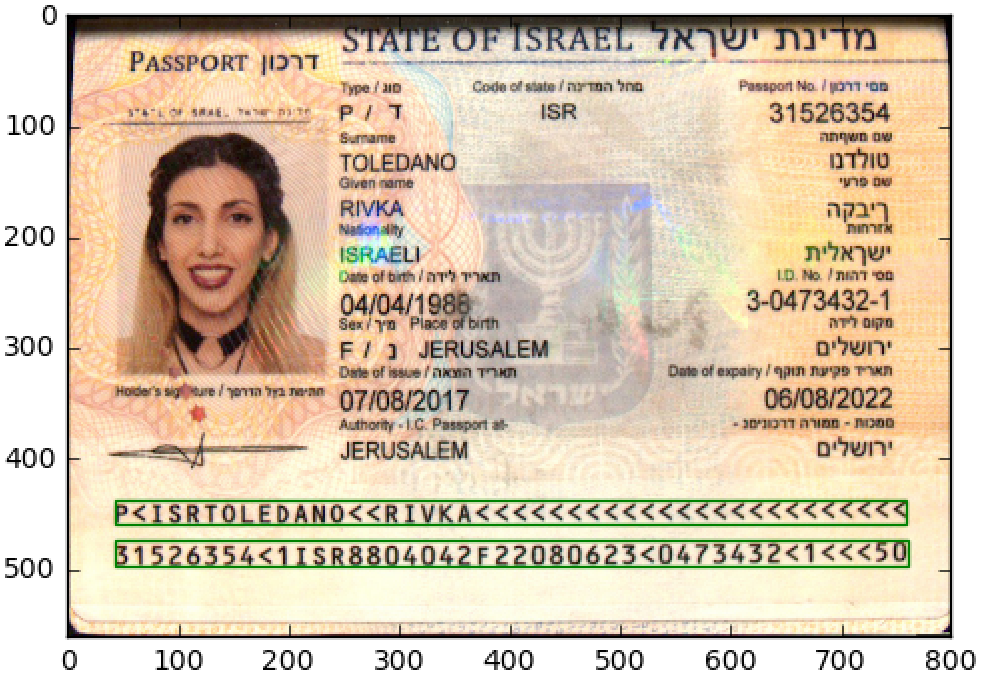


In [18]:
```
draw_mrz_lines("sample6fewbrackets.png")
```



Let's try another passport

In [19]:
```
draw_mrz_lines("israel.jpg")
```



STATE OF ISRAEL מדינת ישראל

PASSPORT דרכון

Type / סוג: P / ד

Code of state / קוד המדינה: ISR

Passport No. / מס׳ דרכון: 31526354

Surname / שם משפחה: TOLEDANO טולדנו

Given name / שם פרטי: RIVKA רבקה

Nationality / אזרחות: ISRAELI ישראלית

Date of birth / תאריך לידה: 04/04/1988

I.D. No. / מס׳ זהות: 3-0473432-1

Sex / מין: F / נ

Place of birth / מקום לידה: JERUSALEM ירושלים

Date of issue / תאריך הנפקה: 07/08/2017

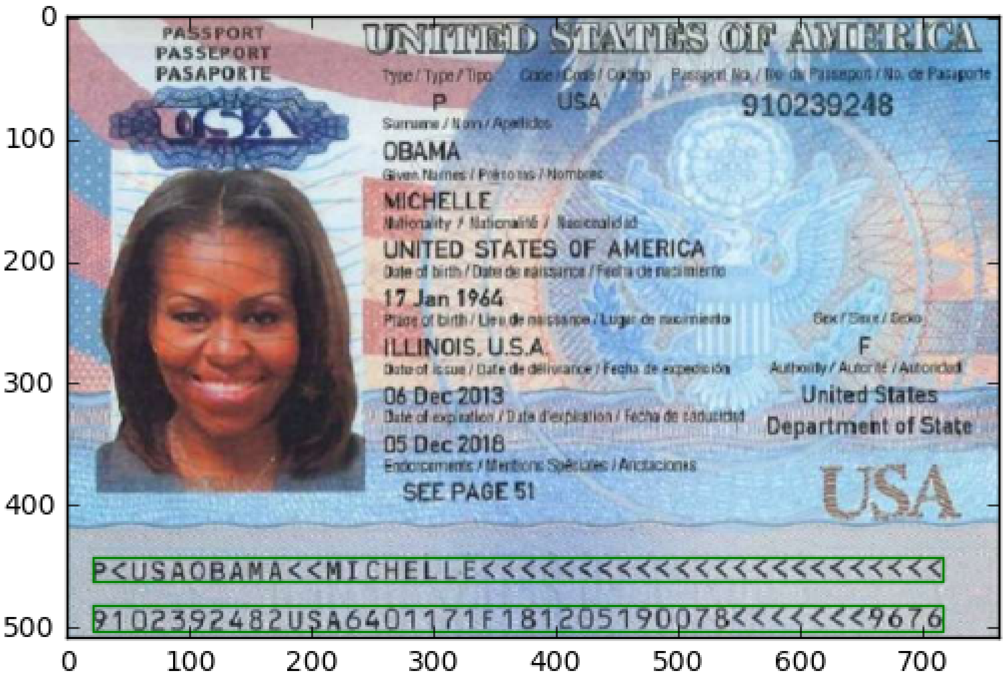
Date of expairy / תאריך פקיעת תוקף: 06/08/2022

Authority - I.C. Passport at- / מפנות - משרד הדרכונים -: JERUSALEM ירושלים

Holder's signature / חתימת בעל הדרכון

```
P<ISRTOLEDANO<<RIVKA<<<<<<<<<<<<<<<<<<<<<<<<
31526354<1ISR8804042F22080623<0473432<1<<<50
```

In [20]:
```
draw_mrz_lines("usa.jpg")
```



PASSPORT PASSEPORT PASAPORTE

UNITED STATES OF AMERICA

USA

Type / Type / Tipo: P

Code / Code / Código: USA

Passport No. / No. du Passeport / No. de Pasaporte: 910239248

Surname / Nom / Apellidos: OBAMA

Given Names / Prénoms / Nombres: MICHELLE

Nationality / Nationalité / Nacionalidad: UNITED STATES OF AMERICA

Date of birth / Date de naissance / Fecha de nacimiento: 17 Jan 1964

Place of birth / Lieu de naissance / Lugar de nacimiento: ILLINOIS, U.S.A.

Sex / Sexe / Sexo: F

Date of issue / Date de délivrance / Fecha de expedición: 06 Dec 2013

Authority / Autorité / Autoridad: United States Department of State

Date of expiration / Date d'expiration / Fecha de caducidad: 05 Dec 2018

Endorsements / Mentions Spéciales / Anotaciones: SEE PAGE 51

USA

```
P<USAOBAMA<<MICHELLE<<<<<<<<<<<<<<<<<<<<<<<<
9102392482USA6401171F181205190078<<<<<<<<9676
```