

Rio Abierto - Mapa del Agua

Procedimiento para publicación

El presente documento busca recoger parte del trabajo sobre los datos de permisos de agua y el desarrollo del mapa realizado en base a...

[Geolocalización de permisos](#)

[Localización por padrones](#)

[Limpieza de datos](#)

[Cruces con datos de Catastro](#)

[Añadiendo valor: cruzando datos](#)

[Cruce con datos de DINAMA](#)

[Cruce con datos del INIA](#)

[Conclusiones](#)

[Instalación del Mapa del Agua \(Ojo Público\)](#)

[Adaptaciones de datos](#)

[Instalación](#)

[Normalización de datos](#)

[Otros trabajos no incluidos](#)

[Granularización de cuencas](#)

[Anexos](#)

[Anexo 1: Bash scripting para procesamiento de datos](#)

[Anexo 2: Código PHP para procesamiento de datos](#)

[Anexo 3: Mail a DINAMA](#)

Geolocalización de permisos

La planilla de partida:

<https://docs.google.com/spreadsheets/d/1YGOy1SeH1ORrIVzATEfJihZ8nSdLtluni703dDDUE4/>

Esta planilla es fruto del trabajo del equipo de La Diaria involucrado, la misma cuenta con los permisos de agua...

Dado que no se contaba con una localización confiable de los permisos de agua, la forma de geolocalizar los permisos que figuran en la planilla fue a través del número de padrón y el departamento. Para ello se generó una clave conformada por el padrón seguida del

departamento. Éste es uno de los puntos a verificar, dado que para algunos casos puede ser que los resultados arrojen más de un valor.

El procedimiento que deviene en la localización de los Permisos se detalla a continuación.

Localización por padrones

Limpieza de datos

Lo primero que se realizó fue un filtro con la herramienta de Open Refine, de manera de incluir únicamente Permisos o Concesiones vigentes. Es decir

1. Se filtró el campo "Tipo de Permiso" por Concesiones y Permiso
2. Se filtraron permisos con "Fecha Vencim." menor a Enero de 2019.

Mediante esta operación se generó la planilla

<https://drive.google.com/open?id=1GXyDI3laxA9HEzxaJws90k8UvTUEnqCZ>

Se observó que los campos de Padrón en la planilla se encontraban agregados y muchas veces con saltos de línea y otros caracteres que no hacían posible el análisis, por lo que se procedió a su limpieza. Para ello se realizó un script en bash, el cual se detalla en el [Anexo 1](#). Esto generó la planilla

https://drive.google.com/open?id=12RAIV7N8gtRe3DZU0B_V8T_GyTlclv4r, la cual contiene una línea por padrón y un nuevo campo con la clave PADRÓNDEPTO.

Cruces con datos de Catastro

El primer cruce se realizó con la capa de datos de padrones realizadas por catastro y que es encuentran en el catálogo nacional de datos:

<https://catalogodatos.gub.uy/dataset/shapes-del-parcelario-rural-y-urbano/resource/2073596d-f122-4030-9eb4-eaca1cdf1e9c>

El cruce se realizó con la herramienta QGIS y consistió, someramente, en los siguientes pasos:

1. Se genera clave única para indexación por PADRON + DEPTO = LKEY en las capas de parcelario: (concat("PADRON","DEPTO"))
2. Se genera una nueva capa con los datos geográficos que devienen de la unión de La Planilla con la capa de parcelarios a través del código LKEY generado para ambas instancias.

Consideraciones: Se detectan varios padrones en 0, se estima son padrones urbanos, por lo que se considera la misma operación sobre dichos padrones.

Los resultados muestran que faltan 11 registros en la planilla final del procesamiento de datos de padrones, según las conciliaciones. Hay 3077 palabras Padrón en La Planilla, pero hay 3066 filas en la Planilla Geolocalizada. Aún no se han trabajado estos 11 registros.

Se excluyen 111 registros de los cuales no se pueden matchear los padrones

Se generan 11.664 registros ya que se identifica que los padrones no son únicos. En los datos de padrones aparece además la localidad o cuadrícula, que sería al parecer la forma de diferenciarlos, pero que no tenemos en La Planilla.

Ésta metodología debería ser revisada teniendo en cuenta el comentario anterior. Se toma la solución más excluyente que es considerar únicos los registros, con lo cual se perderán más valores en los cruces con DINAMA.

Añadiendo valor: cruzando datos

DATOS INTERESANTES

<https://www.mvotma.gub.uy/buscar-inteligente>

No aparecen las resoluciones de agua, se probaron unas cuantas y no se encontraron resultados

Cruce con datos de DINAMA

Si bien se especificó que los datos podrían no ser precisos se decidió cruzar los datos de La Planilla Geolocalizada ya que esto nos agregaba un nivel más de detalle en el uso de los Permisos. En los puntos de la DINAMA se agrega el destino, el cual es un nivel de detalle mayor que el uso. A modo de ejemplo, dónde originalmente teníamos como uso Riego, con el detalle del destino en los puntos se transformó a Riego->Forrajes/Pradera/Semilla o Riego->Frutales.

Se filtraron los puntos de DINAMA contra los Números de expediente y resultaron en 96 registros, por lo que claramente estábamos perdiendo de un montón de expedientes que por los datos que da la capa no sabemos a quiénes pertenecen y otros datos que sí están en nuestra planilla. Se decide continuar con la metodología de los padrones.

Lamentablemente no podemos verificar que los datos de la DINAMA sean los mismos, ya que se encuentran en un portal que no arroja metadata de los mismos. Se contactó a las autoridades responsables pero no se obtuvo respuesta, puede verse el correo en el [Anexo 3: Mail a DINAMA](#). Por lo que se ingresan los links actuales a las diferentes capas, las cuales parecen tener datos a partir de 2004.

Para ello se descargaron los datos de las fuentes de agua, que en su momento arrojaron los siguientes resultados:

- Tomas: Se listan 717,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c392&outputFormat=SHAPE-ZIP>
- Pozos: Se listan 3879,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c394&outputFormat=SHAPE-ZIP>
- Reservorios: Se listan 31,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c393&outputFormat=SHAPE-ZIP>
- Represas: 1094,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c391&outputFormat=SHAPE-ZIP>

- Tajamares: 418,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c389&outputFormat=SHAPE-ZIP>
- Tanques: 769,
<https://www.dinama.gub.uy/geoserver/u19600217/ows?service=WFS&version=1.0.0&request=GetFeature&typeName=u19600217:c390&outputFormat=SHAPE-ZIP>

Se genera una única capa con los datos antes especificados y se cruza con la herramienta QGIS con La Planilla Geolocalizada. El cruzamiento se hace por intersección entre los polígonos devueltos por el cruzamiento con los datos de Catastro y los puntos generados con la planilla anterior.

Se cruza con capa de puntos de DINAMA:

- Se generan 1996 registros contra 2689 que tiene la tabla original. Por lo que menos los registros no encontrados por padrón nos da unos 582 sin machear, los cuales habría que revisar individualmente.
- Ésto debe ser revisado por contraposición con una metodología más incluyente.

Cruce con datos del INIA

Para la visualización de los datos se decide mostrar los niveles hídricos sobre los que están localizados los permisos.

Consideraciones: el índice de balance hídrico se elabora mensualmente, por lo que se decide tomar un mes como referencia (ej: el mes más "seco") y aclarar que se hace contra eso. En su sistema hay una valoración de 0 a 180 mm con 15 mm de paso (lo que hacen 12 valores en la escala, ej:0-15, 15-30...), por lo que se agregan las escalas para determinar un rango de 6 valores, lo que hace un paso de 30mm.

El procedimiento efectuado se asemeja al siguiente:

1. Se descargan capas de Balance hídrico de suelos: <http://sig.inia.org.uy/sigras/>
2. Se descarga capa de percentil 90 para enero (mes de mayor sequía en rango más bajo) y capa de percentil 10 para junio (mes de menor sequía en rango más alto)
3. Se cambia proyección de capa de puntos-padr para trabajar con EPSG:4326
4. Se generan datos con valores de los percentiles. Éstos datos se agregan a La Planilla Geolocalizada
5. Se estudia tabla generada
 - a. Hay 6 registros fuera del mapa de balances
 - b. Hay 5 registros cuya diferencia de un período a otro es mayor a 2 clases (la clase está asociada al rango de balance)

Conclusiones

De esta forma se concluye el procesamiento de datos, sabiendo que debido al propio procesamiento y falta de datos, los datos expuestos son sólo una aproximación y pueden incluir muchos errores e inconsistencias. La versión actual de La Planilla Geolocalizada:

Instalación del Mapa del Agua (Ojo Público)

Adaptaciones de datos

1. En QGIS se exporta La Planilla Geolocalizada de shape a GeoJson
2. Se transforma a topojson con nodejs geo2topo del paquete topojson:
 - `sudo npm install -g topojson`
 - `sudo ln -s /usr/bin/nodejs /usr/bin/node`
 - `geo2topo Parcelas_filtradas_ejemplo.geojson -o Parcelas_filtradas_ejemplo.topojson` (el parámetro -q comprime los datos, ej: -q.25)
 - `topoquantize 1e4 deptos_uy.topojson -o deptos_uyQ.topojson`
- Se cambian atributos `nombrec1` por `NOMBRE` y `NIVEL2` (codcuena) por `CODIGO`, que son los campos que usa la herramienta.
- Se cambia `cuencas_uruguay` por `UH`

Instalación

INSTALACIÓN Debian 9:

- `curl -sL https://deb.nodesource.com/setup_6.x | sudo -E bash -`
- `sudo apt-get install -y nodejs`
- `npm install`
- `sudo npm install gulp -g`
- `gulp scripts`
- `gulp styles`
- `gulp images`
- `npm install backbone`

Se levanta instancia de Ojo Público

Normalización de datos

- Tipos de uso, se agregan:
 - Consumo Humano
 - Otros Usos -> Otros usos
 - Otros Usos Agropecuarios
 - Riego
 - Usos No Consuntivos
 - No definido (cuando no hay datos)

- Clase de fuente
 - Se aplica Subterráneo para todas las tomas de pozos y superficial al resto, según lo acordado con el equipo de Rio Abierto
- Resolución:
 - Se normalizan hay typos

Otros trabajos no incluidos

Granularización de cuencas

Se detecta tras análisis de herramienta que la localización es por cuenca, y ya están relacionados a las cuencas. Por lo que se trabaja sobre mapa de cuencas:

<https://www.dinama.gub.uy/geoservicios/>

<https://www.dinama.gub.uy/geoserver/u19600217/wms?service=WMS&version=1.1.0&request=GetMap&layers=u19600217:c097>

Se detecta que la estructura utilizada por OP Tiene en el archivo de cuencas varios niveles anidados con el nombre de:

UH_nivel1, etc.

Habría que ver cómo se enlazan las cuencas, ya que parece que los archivos que se extraen de las cuencas no tienen las cuencas padre, por lo que habría que hacerlo por geolocalización. Lo mismo para la planilla de licencias, ya que en la planilla tienen seteado el código de licencia de nivel 1, cuando en realidad debiera de ser el de nivel 5.

Se procesan capas de cuencas de niveles 1 a 5

Se detecta que la intersección de las capas se duplica para cuencas en zonas de límites de capas. Se asume que es porque comparten el límite

<https://gis.stackexchange.com/questions/38868/calculate-area-of-intersecting-polygons-in-qgis>

Por alguna razón se pierden los datos al hacer varios joins por localización, por lo que se hace de a pares

Se filtran padrones rurales por PADRDEPTO y se cruza capa de N5

Se obtienen los siguientes resultados:

En la capa quedan aproximadamente 2500 registros de padrones

Se crea script php para validación de datos, con los siguientes resultados:

INICIALES: 3942

Repetidos: 913

TOTALES distintos: 3029

Por lo que se estima que se pierden padrones REVISAR (puede ser porque sean urbanos u otro, hay que hacer validación contra la capa)

Se decide continuar con la integración y revisar más adelante.

<http://tilemill-project.github.io/tilemill/docs/guides/joining-data/>

Anexos

Anexo 1: Bash scripting para procesamiento de datos

Primer versión de bash scripting utilizado para el procesamiento de padrones en La Planilla. (El código es aproximado, puede haber sufrido modificaciones en su aplicación, pero su fondo es el mismo).

```
rm Datos/planilla_ordenada.csv
while read line;
do
  depto=$(echo -e "$line" | cut -f9 -d',' | awk '{print toupper($0)}' | awk '{$1=$1};1');
  echo -e -n "$line," >> Datos/planilla_ordenada.csv;
  padrones=$(echo -e "$line" | cut -f25 -d',' | tr -d '\n' | grep -Po '(?<=Padrón\s)[^\s]*' | cut -f1 -d, |
echo $(cat $1));
#echo $padrones;
oldIFS=$IFS;
IFS=' ';
for padron in $padrones; do
  #padron=$(echo $padron | awk '{$1=$1};1' );
# if [[ $padron =~ ^-[0-9]+$ ]];
# then
  echo -n "$padron$depto," >> Datos/planilla_ordenada.csv;
# fi
done;
IFS=$oldIFS
echo "" >> Datos/planilla_ordenada.csv;
done < Datos/Permisos-de-agua-2011-2018-OP-refined.csv
```

Anexo 2: Código PHP para procesamiento de datos

Código PHP utilizado para el procesamiento de datos

```
<?php
/*
```

//DIFERENCIAS DE PADRONES

```
$pad_str = file_get_contents('./PadrDepto-inline');
$pad_arr = explode(',', $pad_str);
$pad_arr_ord = array();
$repeated = 0;
print "\nINICIALES: ".sizeof($pad_arr);
foreach ($pad_arr as $pad) {
    if ( array_key_exists($pad, $pad_arr_ord) ){
        $repeated++;
    }
    else {
        $pad_arr_ord[$pad] = 1;
    }
}
print "\nRepetidos: ".$repeated;
print "\n TOTALES: ".sizeof($pad_arr_ord)."\n";
*/
/*
```

//ANIDACIÓN DE NIVELES

```
//NIVEL 1-2
$result = array();
$handle = fopen('./Cuencas-Emparejado-Niveles-n1-n2.csv', 'r');
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    $cod = $row['0'];
    $result[$cod] = array(
        "NOMB_UH_N1" => $row['6'],
        "NOMB_UH_N2" => $row['2'],
        "NIVEL1" => $row['3'],
        "NIVEL2" => $cod,
    );
}
fclose($handle);
//NIVEL 2-3
$tmp_res = array();
$handle = fopen('./Cuencas-Emparejado-Niveles-n2-n3.csv', 'r');
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    $code = $row['0'];
    $last_code = $row['4'];
    $tmp_res[$code] = $result[$last_code];
    $tmp_res[$code]["NOMB_UH_N3"] = $row['3'];
    $tmp_res[$code]["NIVEL3"] = $row['0'];
}
$result = $tmp_res;
```



```

fclose($handle);
//NIVEL 3-4
$tmp_res = array();
$handle = fopen('./Cuencas-Emparejado-Niveles-n3-n4.csv', 'r');
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    $code = $row['0'];
    $last_code = $row['4'];
    $tmp_res[$code] = $result[$last_code];
    $tmp_res[$code]["NOMB_UH_N4"] = $row['3'];
    $tmp_res[$code]["NIVEL4"] = $row['0'];
}
$result = $tmp_res;
fclose($handle);
//NIVEL 2-3
$tmp_res = array();
$handle = fopen('./Cuencas-Emparejado-Niveles-n4-n5.csv', 'r');
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    $code = $row['0'];
    $last_code = $row['4'];
    $tmp_res[$code] = $result[$last_code];
    $tmp_res[$code]["NOMB_UH_N5"] = $row['3'];
    $tmp_res[$code]["NIVEL5"] = $row['0'];
}
$result = $tmp_res;
fclose($handle);
//Print to file
$output = fopen("./cuencas_anidadas.csv",'w') or die("./cuencas_anidadas.csv");
fputcsv($output,
array("NOMB_UH_N1","NOMB_UH_N2","NIVEL1","NIVEL2","NOMB_UH_N3","NIVEL3","NOM
B_UH_N4","NIVEL4","NOMB_UH_N5","NIVEL5"));
foreach($result as $cuenca) {
    fputcsv($output, $cuenca);
}
fclose($output) or die("Can't close php://output");*/

// ASIGNACIÓN DE CÓDIGOS DE CUENCA
//Get padrdept - cuenca_id codes
/*$handle = fopen('./MAPas/padrdepto-en-mapa.csv', 'r');
$codes = array();
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    $code = $row['0'];
    $codes[$code] = $code;
}

```

```

print "\n Codes handed \n";
$tmp_res = array();
$handle = fopen('./planilla_ordenada.csv', 'r');
$i=0;
$output_err = fopen("./permisos_UY_No-encontrado.csv",'w') or
die("./permisos_UY_No-encontrado.csv");
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    if ($i==0){
        $i++;
        continue;
    }
    $pdrdeptos = explode(",",$row['24']);
    $changed = 0;
    foreach ($pdrdeptos as $pdrdepto) {
        if ( array_key_exists($pdrdepto, $codes) ){
            $row['24'] = $codes[$pdrdepto];
            $tmp_res[] = $row;
            $changed = 1;
        }
        else {
            fputcsv($output_err, array($pdrdepto));
            print "\n Padrones no encontrados:";
            print_r($pdrdepto);
        }
    }
}
}
$output = fopen("./permisos_UY_single.csv",'w') or die("./permisos_UY.csv");
//fputcsv($output,
array("NOMB_UH_N1","NOMB_UH_N2","NIVEL1","NIVEL2","NOMB_UH_N3","NIVEL3","NOMB_UH_N4","NIVEL4","NOMB_UH_N5","NIVEL5"));
foreach($tmp_res as $cuencia) {
    fputcsv($output, $cuencia);
}
fclose($output) or die("Can't close php://output");
*/
//UNA LINEA POR PADRON
$handle = fopen('Datos/Permisos-de-agua-2011-2018-OP-refined.csv', 'r');
$i=0;
$unwanted_array = array('á'=>'A', 'é'=>'E', 'í'=>'I', 'ó'=>'O', 'ú'=>'U');
while (($row = fgetcsv($handle, 1000, ',')) !== FALSE){
    if ($i==0){
        $i++;
        $tmp_res[] = $row;
    }
}

```

```
        continue;
    }
    $depto = strtr( strtoupper( $row['8'] ), $unwanted_array );
    //print_r($depto);
    $pdrdeptos = explode(" ", $row['24']);
    print "\n";
    foreach ( $pdrdeptos as $pdrdepto ) {
        $pdrdepto = explode(" ", $pdrdepto);
        foreach ( $pdrdepto as $wkey => $word ) {
            if ( $word == 'Padrón' ){
                print_r($pdrdepto[$wkey+1].$depto);
                $row['24'] = $pdrdepto[$wkey+1].$depto;
                $tmp_res[] = $row;
            }
        }
    }
}
}
}
$output = fopen("Datos/permisos_UY_single.csv", 'w') or die("./permisos_UY.csv");
//fputcsv($output,
array("NOMB_UH_N1", "NOMB_UH_N2", "NIVEL1", "NIVEL2", "NOMB_UH_N3", "NIVEL3", "NOM
B_UH_N4", "NIVEL4", "NOMB_UH_N5", "NIVEL5"));
foreach($tmp_res as $cuenca) {
    fputcsv($output, $cuenca);
}
fclose($output) or die("Can't close php://output");
?>
```

Anexo 3: Mail a DINAMA

Subject: Geoservicios DINAMA

To: dia@mvtoma.gub.uy

From: Fernando Uval <fernando@datauy.org>

Estimadas/os,

Mi nombre es Fernando Uval e integro la organización DATA. Estamos trabajando con datos ambientales para diversos proyectos y tengo algunas consultas y observaciones que me gustaría hacerles llegar.

En primer lugar felicitarlos por el sitio y las mejoras que vienen implementando, me parece que está muy bueno, sobretodo la integración de diversos conjuntos de datos con diversas fuentes. Es algo que personalmente creo agrega valor y ayuda a integrar las diversas cosmovisiones de algo tan complejo como el medio ambiente.

En segundo lugar les quería realizar una consulta respecto a las capas de cuencas hidrográficas. Veo que los diversos niveles no incluyen el "padre" o nivel superior de la cuenca.

Es decir, las cuencas de nivel 2 no están asociadas a las de nivel 1 directamente (sí, claro está geográficamente). Pero para el procesamiento de datos, contener la cuenca a la que pertenece de nivel inmediatamente superior, simplificaría el trabajo. ¿Es esto posible?

En tercer lugar sugerirles que los recursos sean enlazados en el catálogo nacional de datos (<http://catalogodatos.gub.uy/>), para de esta manera hacerlos más accesibles al público.

Por último quisiera dejarles algunas observaciones que surgen del uso del portal y los datos:

- La página de geoservicios enlaza correctamente a los contenidos (visualizador y descarga), sin embargo las páginas del catálogo parecen haber quedado mal enlazadas:
 - No funcionan los links para los conjuntos de datos, ej:
<https://www.dinama.gub.uy/geonetwork/srv/spa/catalog.search#/metadata/7b72fb44-9b83-42b5-92d5-5426c696f3c2>
 - Los conjuntos de datos solo se acceden a través de la búsqueda, lo que no permite tener URLs descriptivas (ej la anterior)
 - Existe dentro del conjunto un enlace a la página de geoservicios, sin embargo todas apuntan a: <https://www.dinama.gub.uy/geoservicios/>, lo que nos lleva a un problema de URLs en el acceso a recursos en la página de geoservicios.
- En la página de geoservicios no se puede enlazar a un conjunto de datos específico, teniendo que navegar "a mano" hasta el recurso
 - Esto podría arreglarse (al menos para enlazar, no para indexar) de manera sencilla, simplemente agregando un id para cada h4, ej:
 - `<h4 class="panel-title">Áreas de reproducción de peces nectónicos<i class="fa fa-plus-circle fa-lg"></i></h4>`
 - `<h4 class="panel-title" id="areas-de-reproduccion-de-peces-nectonicos">Áreas de reproducción de peces nectónicos<i class="fa fa-plus-circle fa-lg"></i></h4>`
 - <https://www.dinama.gub.uy/geoservicios/#areas-de-reproduccion-de-peces-nectonicos>
 - Viendo el sitio deberían de agregar JS para poder expandir el recurso cuando se accede por dicha URL

Perdón por el largo del correo, quedo a las órdenes y a la espera de su respuesta.

Un cordial saludo

--

Fernando Uval
Equipo de DATA
contacto@datauy