

Hacking Archaeology

Building an open API for archaeological data

@archaeogeek / Jo Cook / Astun
Technology

This talk is on github so follow along at
<http://archaeogeek.github.io/EAA2015/index.html>

I'm an ex-archaeologist
and GeoGeek, Passionate
about open source
software, open data and
open standards, (and still
very keen on archaeology)

Beyond the
Castle and
HackLancaster:
our story begins

At the Castle: 29th–30th November

All
Events
Free



Public Archaeology Exhibition: What Have We Found?

Heritage Lottery funding has enabled the first comprehensive archaeological survey on the site of Castle Hill and the site of Lancaster's Roman Fort. This marks the beginning of an exciting journey of discovery about the early origins of our city.

29
NOV 1pm – 5pm

30
NOV 10am – 3pm

A-wing, Lancaster Castle

HACK LANCASTER

A-wing
Lancaster
Castle



29th–30th
November
2014

A 24hr hackathon in celebration of local data in Lancaster.

Explore, interpret and make using newly released local archaeological, historical and social data.

Register to attend:
beyondthecastle.org/hacklancaster

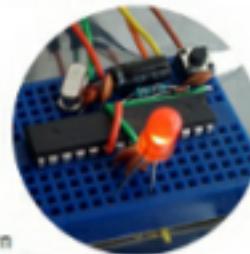
@ShrimpingIt Family Workshops

Morecambe makers @ShrimpingIt will teach novices how to cheaply and easily invent and deploy their own digital devices using electronics prototyping materials.

All Ages 7+
Children accompanied by an adult.

30
NOV 11am – 3pm
A-wing, Lancaster Castle

Register to attend:
beyondthecastle.org



The Brief: Open up
the data in the
Lancaster UAD for
use in a 24hr
hackathon

...in a prison

HACK LANCASTER

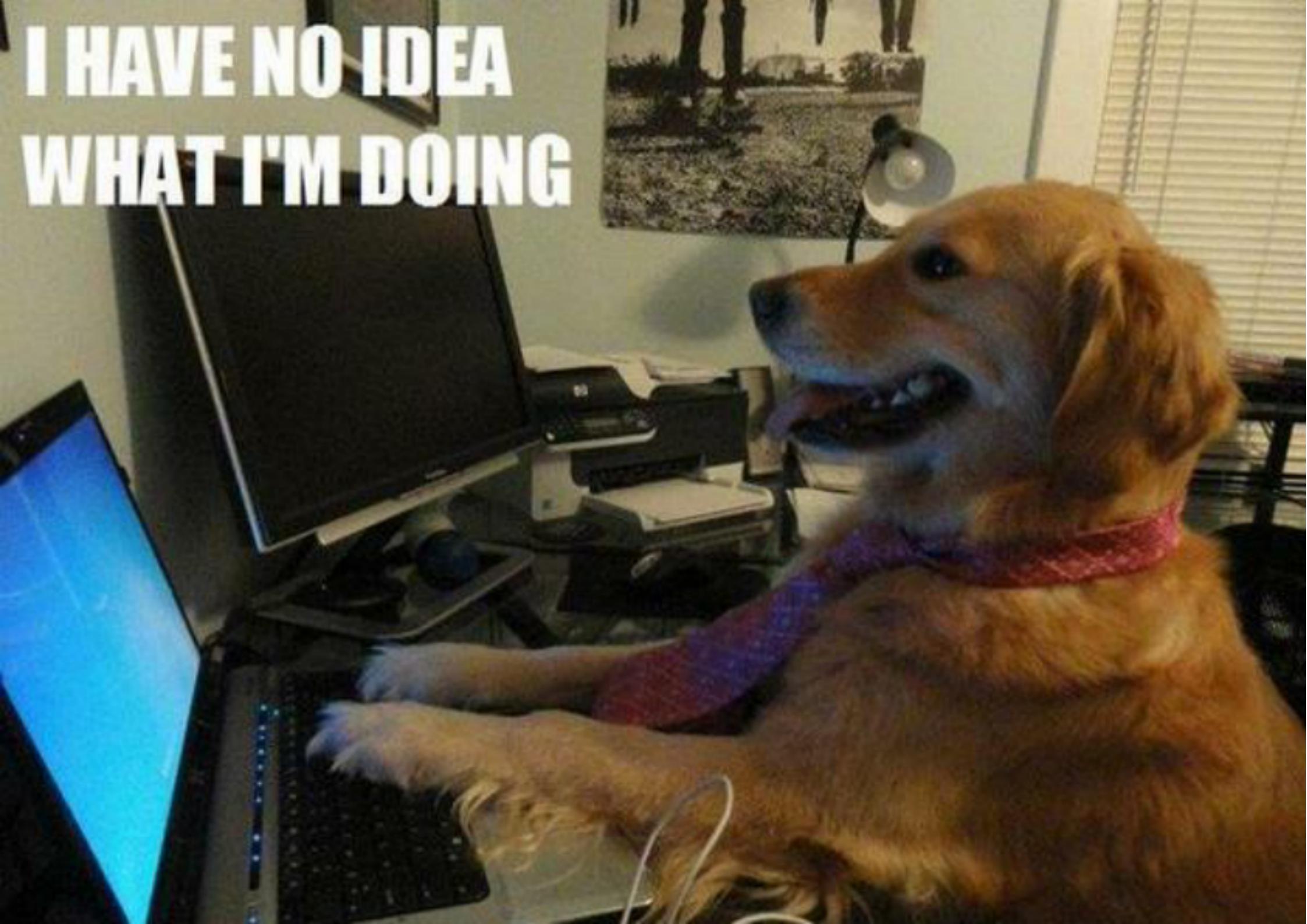
A-wing
Lancaster Castle



29th–30th
November 2014

Question: How
to achieve this
with not a lot of
time (or money)?

I HAVE NO IDEA
WHAT I'M DOING



UADs are great resources, but often locked away and never really used. The best-case scenario is absorption into the SMR/HER.

...useful?



Myscientificcheckonthis



archaeogeek @archaeogeek · Aug 21

Archaeology-twitter-types, is anyone doing fun open-data(ish) things with UAD or HER/SMR data? asking for a friend, of course

RETWEETS

2

FAVORITES

2



2:05 PM - 21 Aug 2015 · Details



...

Reply to @archaeogeek

The
Result.

...

(nothing)

Beyond the hackathon,
whatever we did had to be
standards-compliant,
updateable, reusable and
sustainable in the long term

The data I had to work with:
A number of spatial datasets (shape
files),
An access 97 database,
Lots of georeferenced scans of historic
maps,
Lots of historic images such as
paintings or photos

Task #1: Get
the data out of
Microsoft Access
97!

PostgreSQL was a no-brainer due to its spatial capabilities, and there are plenty of open source, OS-agnostic libraries for working with MS Access.

I used the following and combined them into some python scripts for future reuse:

- PyODBC
- CSV
- mdbtools

(see github.com/archaeogeek/mdb2postgres)

Task #2: Make
the spatial data
available via a
web interface

Again, looking for open source, sustainable, re-usable options, I choose Geoserver for the web-based gui



GeoServer

This also allowed me
to create seamless
base mapping layers
from the historic
maps



Task #3: Make
the non-spatial
data available

RESTful approaches seemed best, so I found ArrestDB. This allowed us to use a simple tablename/attribute protocol to get at data in a json format, eg first 5 entries in the bibliography table, ordered by title

Task #4:
Allow some
spatial analysis

REST again!

Dirt-Simple POSTGIS

HTTP API

This provided an interface for accessing PostGIS functions such as buffers, intersects, and reprojections, again returning results in json format

For example the centroids of a specific monument

Along with the images,
all of this was hosted on
a linux virtual machine
running a simple web
server

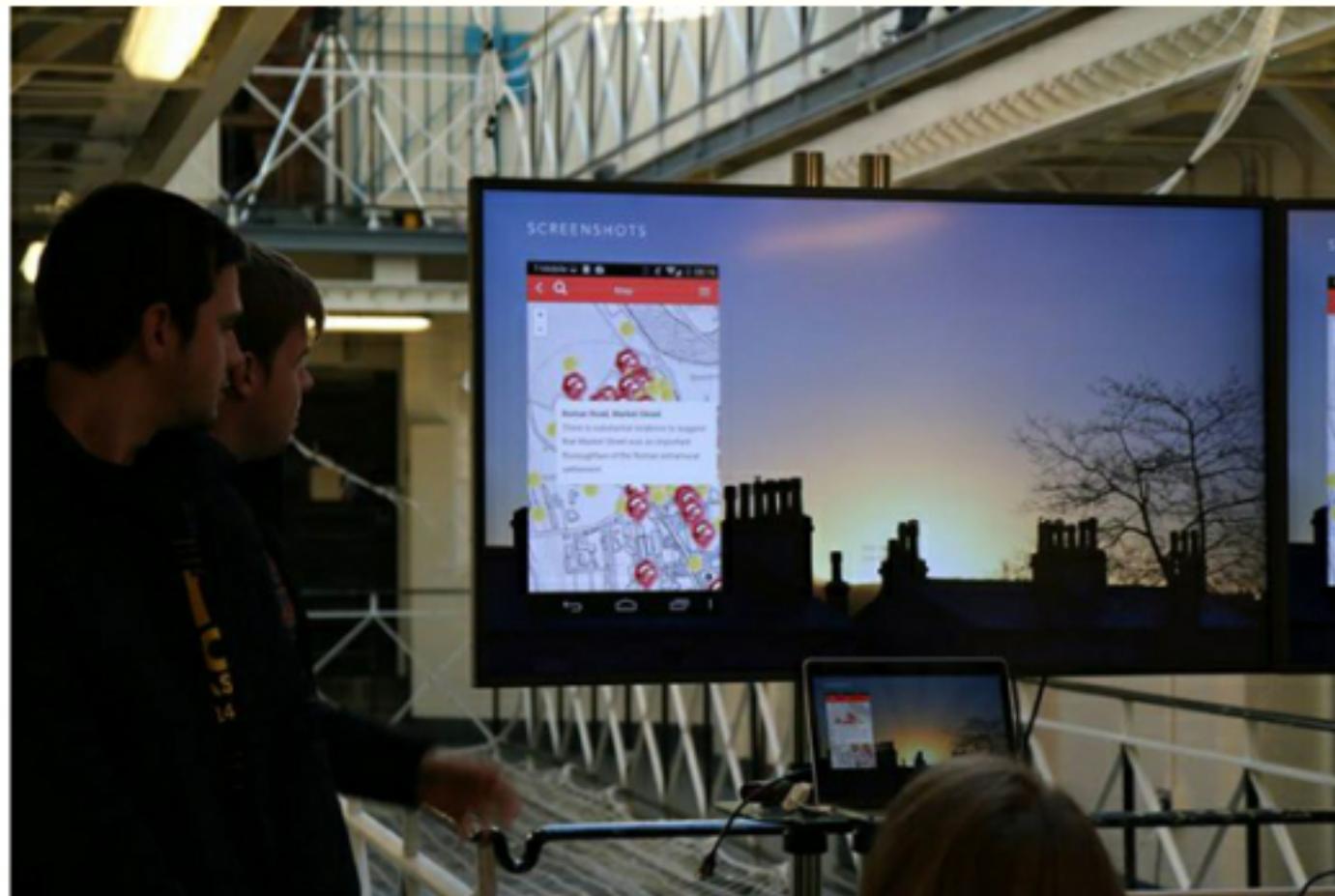
Documentation was produced using [GitBook](#). So it was available online and could be updated throughout the hackathon as needed.

So, how
did it go?

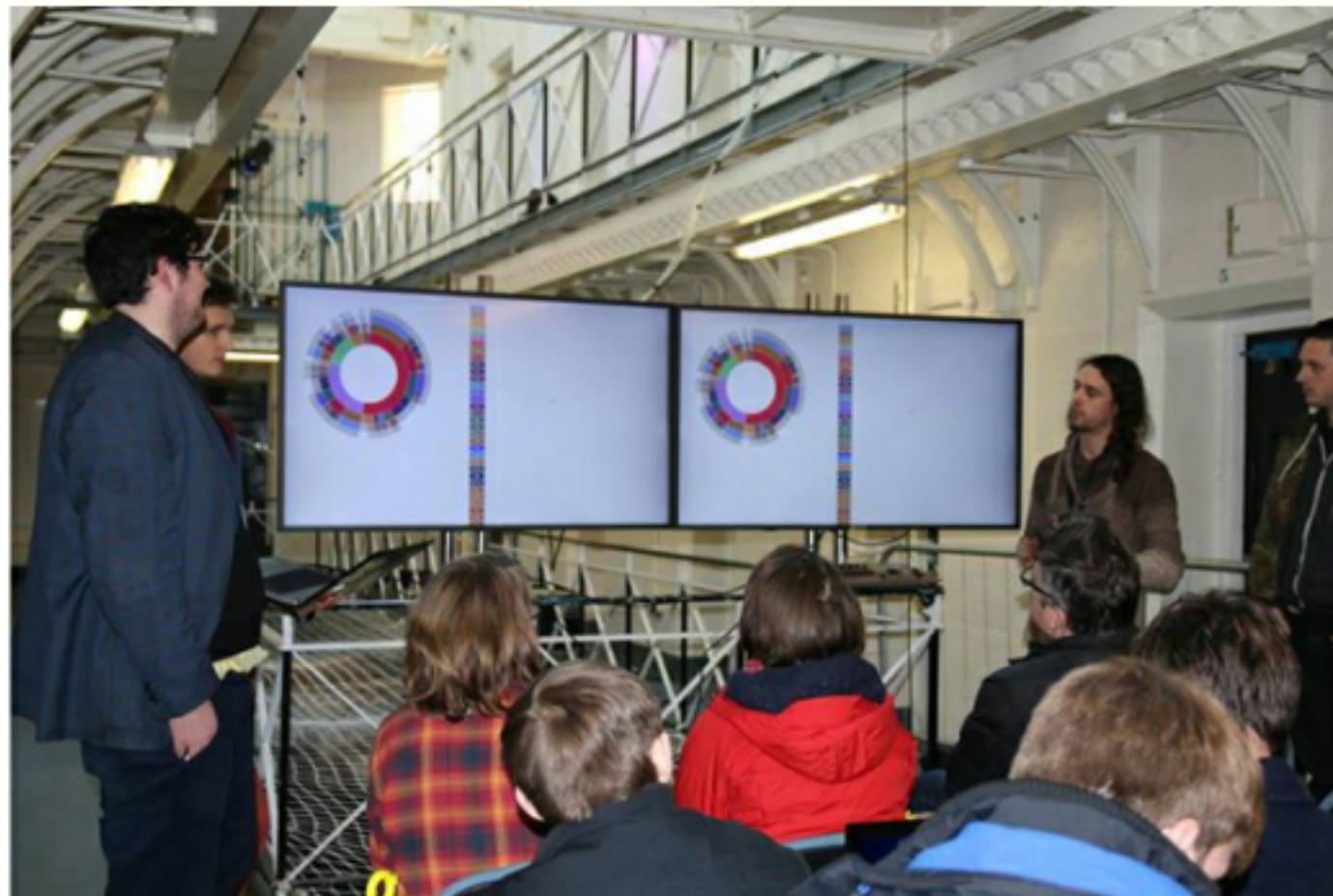
Two teams
of hackers
used the data



The first created a mobile app, allowing people to follow trails around the City "checking in" to monuments



The second created
a tool visualising the
distribution of finds
by period and type



Observations
(NOT
criticisms)

1: UAD data is complicated!

The heavily relational structure is hard to navigate

The controlled language is opaque and can be difficult to understand.

See the [RCHME Archaeological Periods List](#) for an example

Simple questions become complex queries

2: GIS knowledge is specialised

The historic base maps and PostGIS functions were not used

We can't assume everyone has geospatial or analytical expertise

3: People may jump to tenuous conclusions

Correlation does not equal causation!

“(non-professional) interpretations
may... be regarded by the archaeologist
as an exercise in myth or superstition”

Suzie Thomas, the University of Helsinki

In conclusion, non-archaeologists may use the data in unexpected ways...
BUT using genuine archaeological data is a step forward in public engagement with the discipline.

Moving
forward...

We have a framework that could be quickly and easily applied to any UAD to make it available online and for wider use/analysis by both archaeologists and the general public

All the components
are open source, and
don't require a huge
amount of expertise to
assemble or configure

Processes can be automated and repeated, so the data can be updated whenever necessary

BUT ...

We need a "translation-layer" to convert UAD terminologies and structures into something easier for non-specialists to understand, without losing the academic rigour

Deeply relational structures
could be flattened into single
tables
but this may bring additional
complications
for updating, querying,
displaying etc...

However the flexibility of this approach means that multiple front-ends can be constructed to serve different purposes and engage with different demographics

So you want to
produce a Tinder
for Monuments?

Go ahead!



You can also
take advantage
of the latest
technology

Min Stad

Göteborgs stad
Stadsbyggnadskontoret

Klara E want to Cycling here, cykelbana!

12:15 July

AGENCY 9

Citizen suggestions

GÖTEBORG 2021

URBAN DEVELOPMENT

CITIZEN SUGGESTIONS

CREATE A SUGGESTION

TOP CITIZEN SUGGESTIONS

Adina Lundgren May 12, 2013, 08:29

Vasaglassen är ett kaos! Bilar, spårvagnar, cyklister, fotgängare och utryckningsfordon trängs på en liten yta och vid rusningstrafik känns det ottryggt att ta sig över vägarna. Dessutom ligger det en grundskola alldeles vid

12 Shares 1 / 10

SUGGESTIONS DATE | HOT

Göran Svensson August 26, 13:40

Bevara gärna Barken Viking, finfin boende!

Henrik Dahlberg August 18, 13:29

Använd stationshuset och stationshusets tomt för att tillförva värde till samhället. Har kan man till exempel

Maximilian ehm August 17, 17:20 from Älvammet

ett museum med fiskar

veronica alfredsson August 15, 16:18 from Älvammet

Utomhusgym

ulrika åberg August 15, 13:29 from Älvammet

1 / 16 of 910 >>

The risk of **not**
engaging is that real
harm is done through
lack of knowledge



El Huffington Post @ElHuffPost

Follow

La historia del merendero construido sobre un yacimiento arqueológico de más de 6.000 años

huff.to/1hYnGCI

8:55 AM - 25 Aug 2015

100 17

The most
important thing is
to make your data
available because...

“The coolest thing to do
with your data will be
thought of by someone else”

Rufus Pollock and Jo Walsh, the Open Knowledge Foundation

and...

“The set of useful things one can do with a given informational resource is always larger than can be done (or even thought of) by one individual or group”

Rufus Pollock and Jo Walsh, the Open Knowledge Foundation

Thanks for listening!

@archaeogeek / Jo

Cook / Astun

Technology