El projecto

Motivació Introducció

nvestigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona

Estructura del sistem Eines emprades Metodologia

Problemes solucions

mmutabilitat PDFs

Anrenentatoes

Valoració del clien

Demostració

El Far

Un sistema de cerca d'informació per a Amnistia Internacional

Kenan Rhoton

Com funciona?

Estructura del sistema

Eines emprades

Metodologia

Desplegament

Problemes i solucions

Heroku i la immutabilitat

PDFs

Canvis de paradigmes

Aprenentatges

Valoració del client

Demostració

El projecte

Introducció Descripció del sistema

Investigació

Google

Versionista

Motivació

Visualping Google Docs

WebSite-Watcher

Distill Web Monitor

El Far

El projecte

El projecte

Motivació

Introducció

Descripció del sistema

Fer un sistema útil

El Far

Kenan Rhoton

El project

Motivació

Introducció

Investigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona

Estructura del sistem Eines emprades Metodologia

Problemes

Heroku i la immutabilitat PDFs

Aprenentatges

Valoració del clien



El Far

Kenan Rhoton

El projecto

Motivació

Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona

Estructura del sistema Eines emprades Metodologia

Problemes

immutabilitat
PDFs
Capvis de paradigme

Aprenentatges

Valoració del clien

Demostració

► Fer un sistema útil

▶ Fer un sistema diferent

El Far

Kenan Rhoton

El projec

Motivació

Introducció

Investigaci

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona?

Estructura del sistem Eines emprades Metodologia

Problemes solucions

immutabilitat
PDFs
Canvis de paradigme

Aprenentatges

Valoració del clier

Demostració

Fer un sistema útil

- ▶ Fer un sistema diferent
- ► Fer un projecte a més llarg termini

El Far

Kenan Rhoton

El projecte

Motivació

Introducció

Descripcio de

Investigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona?

Estructura del siste Eines emprades Metodologia

Problemes

Heroku i la immutabilitat PDFs

Aprenentatges

Valoració del clier

Demostració

Fer un sistema útil

- ► Fer un sistema diferent
- ► Fer un projecte a més llarg termini
- Aportar a la societat



▶ Un sistema per cercar informació

El Far

Kenan Rhoton

El project

IVIOTIVACIO

Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona

Estructura del sisten Eines emprades Metodologia

Problemes solucions

immutabilitat
PDFs

Aprenentatges

Valoració del clien



- Un sistema per cercar informació
- Monitoritzar diferents pàgines web



Kenan Rhoton

Introducció





- Un sistema per cercar informació
- Monitoritzar diferents pàgines web
- ► Cerca de paraules clau

El Far

Kenan Rhoton

El projecte

Motivació

Introducció

Descripció del siste

nvestigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Manito

Com funciona?

Estructura del sist Eines emprades Metodologia

Problemes

leroku i la nmutabilitat 'DFs

Aprenentatore

Aprenentatges

Valoració del clier





- Un sistema per cercar informació
- Monitoritzar diferents pàgines web
- Cerca de paraules clau
- Cerca en documents

El Far

Kenan Rhoton

Introducció



El projecte

Motivacio

Introducció

Investigació

Google Versionista

Versionista
Visualping
Google Docs
WebSite-Watcher

Distill Web Mo

Com funciona?

Com funciona?

Estructura del sist Eines emprades Metodologia

Probleme

Heroku i la immutabilitat PDFs

Aprenentatges

. . .

AMNISTÍA PINTERNACIONAL

- Un sistema per cercar informació
- ► Monitoritzar diferents pàgines web
- Cerca de paraules clau
- Cerca en documents
- Obtenir les notícies d'interés

Descripció del sistema

El Far

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació.
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents

El Far

Kenan Rhoton

Descripció del sistema



El projecto

Motivació Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Com funciona?

Estructura del sistem Eines emprades Metodologia Desplegament

Problemes solucions

Heroku i la immutabilitat PDFs Capuis de paradis

prenentatges

Valoració del client

Demostració

Com funcional

Estructura del sistema

Eines emprades

Metodologia

Desplegament

Problemes i solucions

Heroku i la immutabilitat

DFs

Canvis de paradigme

Aprenentatges

Valoració del clier

Demostracio

El project

Motivació

Descripció del sistema

Investigació

Google

Versionista

Visualping Google Docs

WebSite-Watcher

Distill Web Monitor

El Far

El projecte

Motivació Introducció

nvestigació

Google

Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona?

Estructura del sist Eines emprades Metodologia Desplegament

Problemes solucions

immutabilitat
PDFs
Capvis de paradigme

Aprenentatges

Valoració del client

Demostració

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents

Versionista

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents



Kenan Rhoton

El projecte

Introducció

Investigació

Contraction

Versionista

Visualping Google Docs WebSite-Watcher

Distill Web Mo El Far

Com funciona?

Estructura del siste Eines emprades Metodologia

Metodologia Desplegament

Probleme: solucions

immutabilitat PDFs Canvis de paradigr

Aprenentatges

Aprenentatges



Visualping

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents
- Falsos positius adicionals!



Kenan Rhoton

Visualping



El proj

Introducció

nvestigació

Google Versionist

Visualping Google Docs

WebSite-Watcher

Distill Web Monitor El Far

Com funciona?

Estructura del siste
Eines emprades
Metodologia
Desplerament

Probleme

Heroku i la immutabilitat PDFs

Aprenentatges

Demostració

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents
- Immens coneixement tècnic requerit per l'usuari

WebSite-Watcher

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents
- Cuota mensual

El Far

Kenan Rhoton

WehSite-Watcher



Distill Web Monitor

El Far

Kenan Rhoton

Distill Web Monitor

Capacitats

- Cerca d'Avisos automàtica
- Gestió de Fonts d'informació
- Gestió de Catàlegs de paraules clau
- Anàlisi de documents
- Increiblement senzill!

El pro

Motivació Introducció

Tourse also also

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher

2.

FI Far

Com funciona?

Estructura del sister Eines emprades Metodologia Desplegament

Problemes solucions

immutabilitat
PDFs
Capvis de paradigme

Aprenentatges

Valoració del clier

Demostració

Conclusió

Cap altra eina satisfà completament les característiques cercades.

El projecte

Motivació Introducció

Investigació

Investigació
Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona?

Estructura del sistema Eines emprades Metodologia Desplegament

Problemes

Heroku i la mmutabilitat PDFs

Apropoptators

Valoració del clien

Demostració

Com funciona?

Estructura del sistema Eines emprades Metodologia Desplegament

Problemes i solucions

Heroku i la immutabilitat PDFs

Canvis de paradigmes

Aprenentatges Valoració del client

El projecte

Motivació Introducció Descripció del sistem

Google

Versionista

Visualping

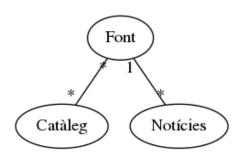
WebSite-Watcher

Distill Web Monitor

El Far

Estructura del sistema

Model de dades



El Far

Kenan Rhoton

Estructura del sistema

Eines emprades Django

Per què Django?

Python

El Far

Kenan Rhoton

Eines emprades

Django

Per què Django?

- Python
- ► Molt ben documentat

El Far

Kenan Rhoton

El pro

Motivació Introducció

Descripció de

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher

El Far

Lom Tunciona?

etructura dal cictan

Eines emprades Metodologia

Desplegament

Problemes solucions

Heroku i la immutabili PDFs

Canvis de paradigmes

Aprenentatges

Valoració del clien

Eines emprades Django

Per què Django?

- Python
- Molt ben documentat
- Modularitat

El Far

Kenan Rhoton

Eines emprades



Eines emprades Django

Per què Django?

- Python
- Molt ben documentat
- Modularitat
- ► Testabilitat

El Far

Kenan Rhoton

Eines emprades



Heroku

Per què Heroku?

Experiència prèvia

El Far

Kenan Rhoton

El pro

Motivació Introducció

Descripcio dei

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Com funciona?

om funciona:

Estructura del sisten

Eines emprades Metodologia

Metodologia Desplegament

Probleme solucions

mmutabilitat PDFs

Canvis de paradigmes

Aprenentatges

Valoració del clie

Heroku

Per què Heroku?

- Experiència prèvia
- ► Integració amb Django

El Far

Kenan Rhoton

El projec

Introducció

. ..

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona?

Caracarana dal atarana

Eines emprades

Metodologia Desplegament

Problemes

leroku i la nmutabilitat 'DFs

Canvis de paradigme

Aprenentatges

Valoració del clie

Heroku

Per què Heroku?

- Experiència prèvia
- Integració amb Django
- ► Gestió senzilla a través de CLI

El Far

Kenan Rhoton

El pro

Motivació Introducció

nvestigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona?

Estructura del sistem

Eines emprades

Metodologia Desplegament

Problemes

olucions Heroku i la

PDFs Canvis de paradigme

Aprenentatges

Valoració del clie

Altres

Mencions especials

Vim

El Far

Kenan Rhoton

Eines emprades

Altres

Mencions especials

- ▶ Vim
- ▶ Git

El Far

Kenan Rhoton

Eines emprades

Altres

Mencions especials

- Vim
- ▶ Git
- Redis

El Far

Kenan Rhoton

El proje

Motivació Introducció

...

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona?

Eines emprades

Metodologia Desplegament

Problemes

mmutabilitat PDFs

Aprenentatges

Demostració

Altres

Mencions especials

- Vim
- ▶ Git
- Redis
- Django File DB Storage

El Far

Kenan Rhoton

Eines emprades



-1

Motivació Introducció

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona?

Estructura del sistem

Eines emprades

Metodologia Desplegament

Problemes

immutabilitat

Aprenentatges

Mencions especials

- Vim
- ▶ Git

Altres

- ► Redis
- Django File DB Storage
- ► PyPDF2

Metodología

Àgil

- Sprints per funcionalitats
- Reunions amb el client.
- Adaptar el sistema als requisits canviants

El Far

Kenan Rhoton

Metodologia



Metodología

TDD

- ► Test first
- Develop later
- ▶ Un test per funcionalitat

El Far

Kenan Rhoton

Metodologia



Metodología

TDD

- Test first
- Develop later
- Un test per funcionalitat

Frrorl

No es va respectar des del principi

El Far

Kenan Rhoton

Metodologia



Desplegament

El Far

Kenan Rhoton

El pro

Motivació Introducció

Investigació

oogle ersionista isualning

Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Com funciona?

Eines emprades Metodologia

Desplegament

Problemes

eroku i la nmutabilitat DFs

Aprenentatges

Demostració

Heroku

- ► Codi localitzat a mighty-ridge-1447.herokuapp.com
- Desplegat mitjancant Git
- Contenidors immutables

El projecte

Motivació Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona?

Estructura del sisten Eines emprades Metodologia

Problemes i solucions

Heroku i la immutabilitat PDFs

Inrenentatoes

Valoració del client

Demostració

Com funciona?

Estructura del sistema

Eines emprades

Desplegament

Problemes i solucions

Heroku i la immutabilitat PDFs

Canvis de paradigmes

Aprenentatges
Valoració del client
Demostració

El project

Motivació

Introduccio

Descripció del sistema

nvestigació

Google

Versionista

Visualping

Google Docs

WebSite-Watcher

Distill vveb Monitor

El Fai

Heroku i la immutabilitat

El Far

Kenan Rhoton

El pro

Motivació Introducció

Investigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monito

Estructura del sistem Eines emprades Metodologia Desplegament

Problemes solucions

Heroku i la immutabilitat PDFs Canvis de paradigm

Aprenentatges

Problema

Els contenidors de Heroku són immutables i per tant qualsevol informació guardada al sistema de fitxers es perd.

Heroku i la immutabilitat

El Far

Kenan Rhoton

El pro

Motivació Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Estructura del sisten Eines emprades Metodologia Desplegament

Problemes i

Heroku i la immutabilitat PDFs Canvis de paradigmo

Aprenentatges

. . .

Problema

Els contenidors de Heroku són immutables i per tant qualsevol informació guardada al sistema de fitxers es perd.

Solució

Fer servir la eina Django DB File Storage per guardar els fitxers de referències de les Fonts a la Base de Dades.

El pro

Introducció

Investigació

Google Versionista Visualping Google Docs WebSite-Watcher

Com funciona

Estructura del sistem Eines emprades Metodologia Desplegament

Problemes solucions

immutabilitat PDFs

Canvis de paradigme

Aprenentatges

.

Damasakusai

Problema

Les poques eines que hi han per tractar PDFs en Python no fan el que volem (extreure el text) ni están ben documentades en general.

PDFs

Problema

Les poques eines que hi han per tractar PDFs en Python no fan el que volem (extreure el text) ni están ben documentades en general.

Solució

Per sort vam acabar trobant la eina PyPDF2, que tot i no estar massa documentada, permet extreure'n text pla de un arxiu PDF.

Canvis de paradigmes

Problema

La solució inicial considerant les Notícies com a seccions concretes de les pàgines web provocava problemes tals com:

- Necessitat de acció concreta (i complexa) de l'usuari per especificar-les
- Debilitat a canvis estructurals

El Far

Kenan Rhoton

El projec

Motivació Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

om funciona?

Estructura del sistem Eines emprades Metodologia Desplegament

Problemes

leroku i la nmutabilitat DFs

Canvis de paradigmes

prenentatges

Valanasić dal alia

)emostració



Canvis de paradigmes

Problema

La solució inicial considerant les Notícies com a seccions concretes de les pàgines web provocava problemes tals com:

- Necessitat de acció concreta (i complexa) de l'usuari per especificar-les
- Debilitat a canvis estructurals

Solució

Es va canviar a la estructura actual d'Avisos, on l'usuari només necessita indicar la Font.

Aprenentatges

Aprenentatges

4日 × 4周 × 4 至 × 4 至 × 至 。

Aprenentatges

- El Far Kenan Rhoton
- El projecte Motivació

Introducció

Investigació

Google Versionista Visualping Google Docs WebSite-Watcher Distill Web Monitor

Com funciona

Estructura del sisten Eines emprades Metodologia

Problemes

immutabilitat
PDFs

Aprenentatges

Valoració del client

)emostració

► TDD!

► Familiaritat no és una garantia



El proje

Motivació Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Com funciona?

Estructura del sistem Eines emprades Metodologia

Problemes solucions

Heroku i la immutabilitat PDFs Canvis de paradigm

Anrenentatoes

Valoració del client

Demostració

Com funciona

Estructura del sistema

Eines emprades

Metodologia

Desplegament

Problemes i solucions

Heroku i la immutabilitat

DFs

Canvis de paradigmes

Aprenentatges

Valoració del client

Demostració

El project

Motivació

Introducció

Descripció del sistema

nvestigació

Google

Versionista

Visualping

Google Docs

Distill Web Monito

El Far

Valoració del client

- El projecte
- Motivació Introducció

Introducció Descripció del sistem

El Far

Kenan Rhoton

- Investigació
- Google
 Versionista
 Visualping
 Google Docs
 WebSite-Watcher
 Distill Web Monitor
- Com funciona?
- Estructura del sisten Eines emprades Metodologia
- Problemes solucions

Heroku i la immutabilitat PDFs

Aprenentatges

Valoració del client

Demostració

- Objectius bàsics complerts
- Gran estalvi de temps amb l'eina
- ▶ Ús habitual d'ella
- Necessitat de manteniment

El proje

Introducció

nvestigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monitor

Com funciona?

Estructura del sistem Eines emprades Metodologia

Problemes solucions

Heroku i la immutabilitat PDFs

prenentatges

Valoració del client

Demostració

Com funcional

Estructura del sistema

Eines emprades

Docalogoment

Droblemes i solucion

Problemes i solucions

Heroku i la immutabilitat

PDFs

Canvis de paradigmes

Aprenentatges Valoració del clier

Demostració

El project

Motivació

Introduccio

Descripció del sistema

nvestigació

Google

Versionista

Visualping

WehSite-Watche

Distill Web Monitor

El Far

Demostració

http://mighty-ridge-1447.herokuapp.com

El Far

Kenan Rhoton

El projec

Introducció

Investigació

Google
Versionista
Visualping
Google Docs
WebSite-Watcher
Distill Web Monito

Com funciona

Estructura del sisten Eines emprades Metodologia Desplegament

Problemes solucions

mmutabilitat

Aprenentatges

Valoració del clier

Demostració