

# Conceptualització d'una solució gràfica per entorns audiovisuals basada en reconeixement facial

Memòria de Projecte Final de Grau

**Grau de Multimèdia**

Àrea Usabilitat i Interfícies

**Autor: Carlos Garcia Figuerola**

Consultor: Joan Coll Planas

Professor: Judit Casacuberta Bagó

Juny de 2020

# Crèdits/Copyright



Aquesta obra està subjecta a una llicència de:

[Reconeixement-NoComercial-SenseObraDerivada 3.0 Espanya de Creative Commons](https://creativecommons.org/licenses/by-nc-nd/3.0/es/)

# Abstract

En la actualitat, cada cop es més necessari l'ús de les noves tecnologies emergents com la intel·ligència artificial o la computació en el núvol sent el reconeixement i classificació d'objectes un dels camps amb més projecció i demanda de la societat actual. Aquest Treball de Fi de Grau busca aplicar el reconeixement facial mitjançant tècniques d'intel·ligència artificial i computació que permetin, a través de l'anàlisi de patrons, la detecció e identificació de rostres mitjançant la comparació amb d'altres patrons registrats en una base de dades.

Un dels propòsits del projecte és l'aplicació d'aquesta tecnologia en una empresa del sector dels mitjans audiovisuals amb la finalitat d'automatitzar els grafismes que intervenen en la confecció d'un programa informatiu diari d'un canal de televisió en el que es transmeten notícies d'interès públic. Per altra banda, es pretén abordar la necessitat de reduir la càrrega de treball dels redactors de l'àrea d'informatius mitjançant la automatització del procés que implica la confecció i execució dels gràfics en emissió.

El disseny d'un sistema de reconeixement facial mitjançant l'ús de tècniques d'intel·ligència artificial ha de permetre afrontar amb garanties la implementació d'un servidor de captació i reconeixement facial que s'integri dins del flux actual del sistema de retolació amb l'objectiu final d'integrar un sistema autònom de retolació que alliberi als redactors de la càrrega de treball i temps que implica la confecció d'un rètol de forma manual.

Com a resultat s'obté una proposta de disseny e implementació d'una solució gràfica basada en el reconeixement facial que s'integra en el flux de treball d'un programa informatiu diari permetent automatitzar la majoria de grafismes que intervenen en un programa informatiu i reduint significativament la càrrega de treball dels redactors de la redacció.

**Paraules clau:** intel·ligència Artificial, reconeixement facial, gràfics, radiodifusió, mitjans audiovisuals, Memòria, Treball de Fi de Grau, Word 2016

# Abstract (English version)

At present, it is increasingly necessary to use new emerging technologies such as artificial intelligence or cloud computing, being the recognition and classification of objects one of the areas with the greatest projection and demand in today's society. This final degree project seeks to apply facial recognition using artificial intelligence and computing techniques that allow, through pattern analysis, the detection and identification of faces by comparison with other patterns recorded in a database.

One of the purposes of this project is the application of this technology in a company in the audio-visual media sector in order to automate the graphics involved in the preparation of a daily news program on a television channel in which current news is broadcast. On the other hand, the aim is to address the need to reduce the workload of newsroom editors by automating the process that involves the preparation and execution of broadcast graphics.

The design of a facial recognition system through the use of artificial intelligence techniques must allow to face with guarantees the implementation of a server of capture and facial recognition that is integrated within the current flow of the graphic system with the ultimate goal of integrating a standalone graphics system that frees editors from the workload and time involved in making a graphics manually.

The result is a proposal for the design and implementation of a graphical solution based on facial recognition that is integrated into the workflow of a daily news program allowing the automation of most graphics involved in a news program and significantly reduce the workload of newsroom editors.

**Keywords:** Artificial intelligence, facial recognition, graphics, broadcasting, audio-visual media, Memory, Final Degree Project, Word 2016

# Índex

1. Introducció.....	7
2. Descripció/Definició/Hipòtesi .....	8
3. Objectius .....	9
3.1 Principals.....	9
3.2 Secundaris.....	9
5. Continguts .....	10
5.1 Anàlisi de les necessitats de l'usuari.....	10
5.2 Anàlisi del programari i maquinari.....	10
5.3 Anàlisi i disseny del prototipat, diagrames UML .....	10
5.4 Anàlisi Heurístic.....	10
6. Arquitectura de sistema .....	11
6.1 Sistema de Reconeixement .....	12
6.2 Sistema de Composició.....	12
6.3 Sistema d'Execució .....	13
7. Plataforma de desenvolupament .....	14
7.1 Programari.....	14
7.2 Maquinari.....	15
8. Planificació .....	17
9. Procés de treball/desenvolupament.....	19
9.1 Investigació.....	20
9.2 Definició.....	23
10. Diagrames UML .....	27
11. Prototips .....	30
11.1 Lo-Fi .....	30
12. Perfils d'usuari.....	33
13. Usabilitat/UX .....	34
13.1 Perfil Tècnic Administració .....	35
13.2 Perfil Usuari Operació .....	36
13.3 Anàlisi Heurístic.....	37
14. Seguretat.....	39
14.1 Seguretat contra intrusions (Intern).....	39
14.2 Seguretat contra intrusions (Extern).....	40
14.3 Seguretat en les electròniques de comunicacions.....	40
14.4 Seguretat en servidors i equips d'operació .....	41
14.5 Seguretat en cas de sinistre .....	41
14.6 Infografia.....	42

15. Requisits d'instal·lació/implantació .....	43
15.1 Sistema de reconeixement.....	43
15.2 Sistema de confecció de retolació.....	44
15.3 Sistema de confecció de retolació.....	45
16. Instruccions d'ús .....	46
16.1 Resum del Procés d'ús .....	46
16.2 Perfil d'usuari.....	47
16.3 Perfil Administrador TI.....	52
17. Projecció a futur .....	57
17.1 Reconeixement d'àudio.....	57
17.2 Reconeixement facial de personalitats.....	57
17.3 Optimització del Reconeixement facial .....	57
17.4 Catalogació del fons documental .....	58
18. Conclusions.....	59
Annex 1. Lliurables del projecte (PENDENT ) .....	60
Annex 2. Bibliografia .....	61
Annex 3. Vita.....	63

# Figures i taules

## Índex de figures

Figura: 1   Infografia representativa de la arquitectura del sistema.....	11
Figura: 2   Diagrama de Gantt de la planificació del projecte .....	18
Figura: 3   Diagrama Disseny Centrat en les Persones.....	19
Figura: 4   Diagrama del procés de reconeixement facial .....	20
Figura: 5   Procediment de configuració d'un enllaç amb una font de dades Avid Maestro   News.....	21
Figura: 6   Procediment de configuració d'un font de dades de tipus SQL .....	21
Figura: 7   Procés de càrrega d'una base de dades en Avid Maestro   News .....	22
Figura: 8   Resultat d'una consulta a la base de dades definida .....	22
Figura: 9   Diferents models de captadora SDi Blackmagic .....	23
Figura: 10   Imatgeria dels productes OpenCV i python.....	23
Figura: 11   Diagrama del procés de captura de senyal de vídeo i posterior emmagatzematge .....	24
Figura: 12   Diagrama del procés d'entrenament a partir de l'anàlisi de les captures realitzades .....	24
Figura: 13   Procés de reconeixement facial a través de la comparació amb el fitxer de mostres .....	25
Figura: 14   Procés de confecció de la retolació generada a través del sistema de reconeixement....	26
Figura: 15   Diagrama UML del procés d'aprenentatge.....	27
Figura: 16   Diagrama UML del procés de reconeixement facial a partir de l'arxiu de mostres .....	28
Figura: 17   Diagrama UML del procés de confecció i execució de la retolació .....	29
Figura: 18   Prototip en baixa definició de la interfície del servidor del sistema de reconeixement. ....	30
Figura: 19   Prototip en baixa definició del menú de captura de senyal SDi .....	31
Figura: 20   Prototip en baixa definició de la interfície de gestió del repositori de rostres capturats....	32
Figura: 21   Diagrama físic i lògic d'una configuració Virtual Switching System (VSS).....	40
Figura: 22   Infografia de les diferents mesures de seguretat a implementar en el projecte.....	42
Figura: 23   Infografia de les fases que intervenen en el procés de retolació automatitzada .....	46
Figura: 24   Menú que permet executar la monitorització d'una graella a l'aplicació Avid iNews .....	47
Figura: 25   Procediment de càrrega del monitor en els dispositius sincronitzats.....	48
Figura: 26   Unitat de xarxa requerida per accedir a la base de dades de Avid Maestro.....	49
Figura: 27   Icona del programa Avid Maestro Controller .....	49
Figura: 28   Interfície d'autenticació amb el servidor Avid Maestro Controller .....	49
Figura: 29   Interfície principal del programa Avid Maestro Controller.....	50
Figura: 30   Acció per connectar amb el servidor Maestro i carregar les graelles monitoritzades .....	50
Figura: 31   Graella de rètols carregada a Avid Maestro Controller .....	51
Figura: 32   Sistema de visualització i operació de la retolació a Avid Maestro Controller .....	51
Figura: 33   Activació de l'aprenentatge a la pàgina principal del sistema de reconeixement. ....	52
Figura: 34   Interfície de selecció i previsualització del canal de captura.....	53
Figura: 35   Simulació captura del sistema de reconeixement [Imatge: Freepik.com].....	53
Figura: 36   Accés al repositori d'usuaris del sistema de reconeixement de rostres.....	54
Figura: 37   Gestió de la base de dades del sistema de reconeixement. ....	54
Figura: 38   Icona de l'aplicació Avid Page Aditor.....	55
Figura: 39   Interfície de navegació a través del repositori de plantilles de retolació .....	55
Figura: 40   Interfície de configuració d'un rètol a Avid Maestro Page Editor.....	56
Figura: 41   Menú per a guardar la configuració del rètol .....	56

# 1. Introducció

Aquests darrers anys de la meua vida han estat marcats per grans sacrificis a nivell personal i professional. En el meu cas, cursar un Grau Universitari a la Universitat Oberta de Catalunya m'ha permès compaginar els estudis en el Grau de Multimèdia juntament amb la feina de tècnic informàtic en una empresa del sector de comunicacions i audiovisuals a nivell estatal.

Gràcies als coneixements adquirits en el Grau Multimèdia juntament amb la experiència obtinguda a nivell laboral, m'han permès afrontar amb garanties els diferents projectes professionals sorgits aquests últims anys tot aplicant nous conceptes i procediments adquirits.

Arrel d'aquest conceptes, procediments i eines adquirits que sorgeix la oportunitat de desenvolupar un projecte en la seva fase de conceptualització que permeti donar solució a un problema recurrent de la redacció d'informatius a la empresa de comunicacions i audiovisuals en la que actualment exerceixo la meua feina diària.

La redacció d'informatius d'una empresa del sector audiovisuals és un lloc amb una gran càrrega de feina i estrès. Tots treballen de forma coordinada per tal de confeccionar un conjunt de notícies amb la pressió de finalitzar la notícia abans de l'inici de la emissió en directe. És per això que, a mesura que s'apropa l'hora límit d'emissió, augmenta el clima de tensió dels redactors, realitzadors i ajudants de realització de la redacció. Aquest clima de tensió provoca que, a vegades, els treballadors cometin alguns errors o, fins i tot, alguna notícia aparegui incompleta.

Un dels problemes analitzats és la falta de retolació o grafismes en les diferents notícies de l'informatiu. Per tal de solucionar aquest problema desenvoluparem un projecte que implica la creació d'una aplicació de reconeixement facial de vídeo que detectaria la cara dels reporter que està donant la notícia i generaria un rètol (grafisme) amb el nom del periodista que ha aparegut a la peça de vídeo.

La idea, per tant, és la d'automatitzar una part de la feina dels reporters / grafistes a l'àrea d'informatius d'una televisió. D'aquesta manera descongestionaríem una part de la feina dels periodistes i evitaríem, alhora, possibles problemes derivats de la pròpia creació dels gràfics d'un informatiu.



## 2. Descripció/Definició/Hipòtesi

El Projecte consisteix en la conceptualització i disseny d'una aplicació que s'integraria dins del procés de creació de notícies de la redacció d'informatius de la empresa de comunicacions i audiovisuals on treballa i amb la finalitat de descongestionar la càrrega de treball dels integrants de la redacció mitjançant la implementació d'un procés d'automatització que implica la generació de rètols o grafismes mitjançant sistemes de reconeixement facial dels diferents reporters que integren la plantilla de la empresa.

Gràcies a la idea d'implementació d'una automatització de grafismes, els redactors podran alliberar la feina repetitiva de firmar cadascuna de les peces o notícies que integren els diferents informatius diaris que es realitzen.

Actualment, la redacció d'informatius disposa d'un ventall d'eines integrades en el sistema d'edició per tal de realitzar la retolació però aquests procediments no deixen de ser repetitius i costosos per a l'empresa.

La implementació de l'automatització mitjançant d'un aplicatiu permetria l'anàlisi del vídeo de la notícia mitjançant un procediment de reconeixement facial que detectaria les dels reporters que integren la notícia, les compararia amb una base de dades i posteriorment donaria l'ordre de generar un rètol (grafisme) amb el nom del reporter que ha aparegut a la peça de vídeo. Aquest grafisme apareixeria automàticament al final de la peça estalviant recursos i temps a la empresa.

En funció de la efectivitat definida en el projecte inicial podria ampliar-se amb la implementació d'anàlisi i reconeixement facial de personalitats de la vida quotidiana (Polítics, Cantants o personalitats) o fins i tot realitzar el reconeixement de l'àudio i la posterior transcripció en text per tal de crear retolació amb traducció simultània d'idiomes.

# 3. Objectius

## 3.1 Principals

Objectius clau del TF.

- Conceptualitzar i dissenyar una aplicació que permeti automatitzar la tasca de creació de rètols alhora de firmar les notícies d'un informatiu
- Alleugerir la càrrega de treball dels redactors i grafistes gràcies a la introducció d'un procés d'automatització de gràfics.
- Integrar la nova automatització al flux actual de producció de notícies de l'informatiu.

## 3.2 Secundaris

Objectius addicionals que enriqueixen el TF i que poden patir variacions.

- Integrar nous fluxos de treball alhora de desenvolupar la tasca de creació de notícies per a l'informatiu.
- Ampliar les capacitats d'automatització mitjançant la introducció de diferents variables i escenaris que permetin aplicar altres patrons de reconeixement (Grans personalitats, polítics, famosos).
- Ampliar el projecte mitjançant la implementació de reconeixement d'àudio (Traducció en temps real, minutat de notícies per al departament de documentació)

## 5. Continguts

El Propòsit del projecte és la conceptualització i disseny d'una solució tecnològica que permeti la creació automatitzada de retolació o grafics mitjançant un sistema de reconeixement facial i que, alhora, alliberi d'una part de la feina que realitza la plantilla de redactors d'una empresa d'audiovisuals.

Així doncs, el projecte abasta des de la investigació i anàlisi de les necessitats de l'usuari, l'anàlisi de les tecnologies necessàries, passant per la definició de les funcionalitat si escenaris d'ús, i finalitzant amb el desenvolupament d'un prototipat en baixa fidelitat, anàlisi d'heurística i l'anàlisi d'una hipotètica implementació en una escenari real.

En quant a l'estructuració del projecte podem dir que està dividit en diferents fases amb els seus objectius o resultats ben definits. Pel que fa al desenvolupament del projecte podem diferenciar-ne la fase d'anàlisi de les necessitats de l'usuari, la part de l'anàlisi i disseny del software i tecnologies per a la funcionalitat del projecte juntament amb la part d'anàlisi del hardware necessari. Per altra banda procedirem amb la fase d'anàlisi i disseny d'un prototipat en baixa fidelitat juntament amb la definició de diagrames UML i anàlisi d'heurística (Usabilitat) que permeti afrontar amb garanties una possible implementació en un escenari real de producció diari.

### 5.1 Anàlisi de les necessitats de l'usuari

En aquesta fase analitzarem qui és l'escenari actual amb el que conviuen els usuaris, profunditzarem en les seves necessitats i adaptarem el projecte per tal de donar resposta als seus requeriments.

### 5.2 Anàlisi del programari i maquinari

En aquesta fase analitzarem quines són les necessitats en quan a la implementació de la nova tecnologia de reconeixement facial i definirem quines necessitats cal introduir per tal de donar viabilitat al projecte en curs.

### 5.3 Anàlisi i disseny del prototipat, diagrames UML

Desenvoluparem un prototipat en baixa definició de les diferents interfícies implicades en el procés de retolació de notícies posant èmfasi en la fase de detecció, normalització o extracció de les característiques que garanteixin l'èxit de l'algorisme de reconeixement facial.

Per tal de reforçar la informació de tots el processos que intervenen en l el projecte tecnològics, procedirem a crear un diagrama UML des de la fase de captura d'imatge, que ens permetrà iniciar el reconeixement facial, fins a la fase final amb el llançament del rètol quan el vídeo de la notícia estigui en emissió.

### 5.4 Anàlisi Heurístic

Per a finalitzat realitzarem la avaluació de la usabilitat del sistema d'interfícies i processos que intervenen en les diferents fases del projecte mitjançant l'anàlisi heurístic.

## 6. Arquitectura de sistema

Per tal d'implementar un sistema de reconeixement facial dins d'un entorn d'edició de vídeo de la redacció d'informatius caldrà dissenyar una arquitectura client-servidor juntament amb la implementació d'una sèrie de bases de dades per tal d'emmagatzemar i comparar tota la informació que es va generant dins del sistema.

En el projecte que ens ocupa es crea a partir d'una arquitectura servidor quedant restringia l'arquitectura client a la fase inicial i final sent l'usuari el que inicia el procés de reconeixement i el que el finalitza donant validesa a les dades o grafismes resultants.

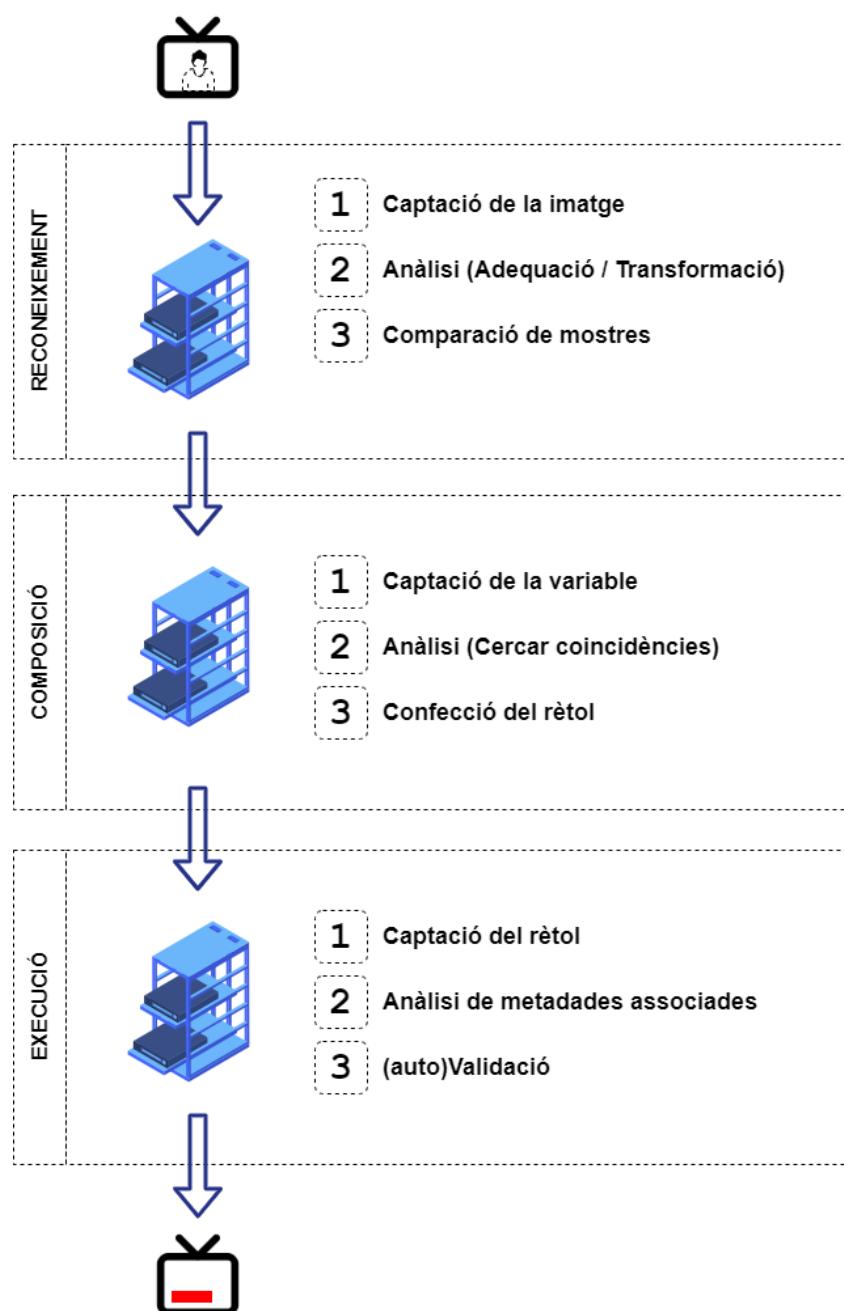


Figura: 1 | Infografia representativa de la arquitectura del sistema

## **6.1 Sistema de Reconeixement**

### **6.1.1 Client**

És el client el que es presta a iniciar el procediment de reconeixement al ser enregistrat amb una càmera de vídeo.

### **6.1.2 Servidor**

És en aquesta fase en la que el servidor realitza el procés de captació de les imatges des de la font de vídeo, realitza un anàlisi de detecció facial, continua amb una adequació o transformació i, finalment, aplica l'algorisme de reconeixement que generarà un resultat en forma de variable.

### **6.1.3 Base de dades**

És el client el que es presta a iniciar el procediment de reconeixement al ser enregistrat amb una càmera de vídeo per tal d'esbrinar si existeix una coincidència entre la imatge captada i les mostres emmagatzemades a la base de dades.

## **6.2 Sistema de Composició**

### **6.2.1 Servidor**

El sistema de composició monitoritza la captació de dades que provenen del sistema de reconeixement. Aquestes dades en forma de variable són analitzades i comparades amb els registres d'una base de dades que conté la informació necessària per a la confecció de la retolació o grafisme. És en aquest punt que el sistema confeccionarà el rètol i l'emmagatzemarà en un repositori temporal.

### **6.2.2 Base de dades**

Tal i com hem dit anteriorment, en aquesta fase de composició, la base de dades ens servirà per emmagatzemar tots els camps d'informació relacionats amb tots els identificadors d'usuaris juntament amb la informació o variables de composició que s'hagin definit.

## **6.3 Sistema d'Execució**

### **6.3.1 Servidor**

El sistema d'execució monitoritza el repositori de rètols generats en el sistema de composició i analitza les metadades que té associat cada rètol. Aquestes metadades permeten representar la informació de tots els rètols en una graella per a ser llançades *on air* de forma automàtica o prèvia validació del usuari.

### **6.3.2 Base de dades**

El sistema d'execució fa ús d'una base de dades que emmagatzema diferents camps d'informació relacionades amb els rètols preparats per a ser llançats en emissió.

### **6.3.3 Client**

En cas que els rètols no estiguin configurats per a ser llançats automàticament, l'usuari serà l'encarregat de validar el rètol realitzant una verificació prèvia a que sigui llançat en emissió uns segons abans de finalitzar la notícia.

# 7. Plataforma de desenvolupament

## 7.1 Programari

### 7.1.1 OPEN CV

És una biblioteca de codi obert amb un repositori d'algorismes de reconeixement de vídeo. Entre les serves funcions podem destacar les següents:

- Processat d'imatges (transformació d'imatge geomètrica.
- Anàlisi de vídeo (estimació de moviment, sostracció del fons i seguiment d'objectes)
- Captura de vídeo.
- Detecció d'objectes (cares, ulls i objectes variats)

### 7.1.2 Microsoft SQL Server

Sistema de gestió de bases de dades relacional amb funcionalitats d'emmagatzemar, gestionar i definir relacions entre dades.

Serà necessària la seva implementació en la fase de reconeixement per tal de capturar les imatges que seran tractades i adequades posteriorment per tal de garantir l'eficàcia del sistema de reconeixement.

Per altra banda, també serà necessària a la fase de composició ja que el sistema disposarà d'un repositori de plantilles de grafismes, rètols i variables d'informació requerides per a la confecció dels diferents rètols i el seu posterior emmagatzematge.

Per últim, a la fase d'execució també serà necessària una base de dades per tal de mostrar la graella de rètols que es disposen en el servidor per tal de ser llançat en emissió de forma automatitzada o prèvia validació per part d'un ajudant de realització en el control d'emissió.

### 7.1.3 Avid Maestro | News

Es tracta d'una *suite* de disseny gràfic que permet la creació de retolació avançada en entorns *broadcast* propietària de la empresa Avid Technology, Inc

Els diferents mòduls emprats per aquest projecte són:

- **Maestro | Designer:**  
Crearem i composarem les escenes, animacions i gràfic que conformaran el disseny del gràfic alhora que ens servirà per a definir àrees on crearem els diferents caps que contindran enllaços a la informació continguda en registres de les bases de dades.

- **Maestro | Page Editor:**

Procedirem a definir quines variables o registres de la base de dades introduïrem a cadascun dels camps del *templates* que hem confeccionat a Maestro | Designer.

És en aquest programari on definirem la lògica que realitzarà la acció de consultar a la base de dades que conté la informació necessària per a confeccionar el rètol. Per altra banda, en aquest programari també definirem si el rètol tindrà definit un automatisme d'entrada i sortida de rètol per tal d'implementar l'automatització.

- **Maestro | Controller:**

És el programari que contindrà la graella de rètols confeccionats i enllestits juntament amb l'ordre d'execució automatitzada i sincronitzada amb els clips de vídeo.

## 7.2 Maquinari

### 7.2.1 Servidor Sistema de reconeixement

Aquest servidor correrà bàsicament el programari OPEN CV. Segons els requeriment de la plataforma per a la implementació de reconeixement de text que fan referència A. Sharma, V. Ravid i M. Beyeler (2019), aquests són relativament assequibles. Els autors recalquen en el seu llibre que les necessitats mínimes són 4GB de memòria RAM, una CPU *multicore* i que no és necessari la utilització d'una GPU per a realitzar càlculs.

Altres referències consultades, MJRoBot (2018), fan ús d'una Raspberry Pi 3 B el qual és un SBC<sup>1</sup> de baix cost amb una potència limitada, juntament amb una càmera que captura a 15 fps<sup>2</sup>.

En el nostre cas en concret, implementaríem un maquinari adaptat a un entorn de producció *broadcast* amb unes especificacions que garanteixin la captura de vídeo en temps real de 25 fps per a fonts d'imatge 4k<sup>3</sup> a través d'una capturadora SDI<sup>4</sup> i no tant per la capacitat de procés necessari per al reconeixement facial en si.

### 7.2.3 Servidor Confecció Retolació

Necessitarem d'un equip amb capacitat de processat i emmagatzematge adien per a gestionar una base de dades SQL juntament amb el requeriments indicats pel programari Avid Maestro Designer i Page editor. Degut a la criticitat del sistema en el entorn *broadcast*, garantirem la tolerància a errors del servei implementat un sistema redundat que garanteixi la operativa del servei en tot moment.

<sup>1</sup> Simple Computer Board

<sup>2</sup> Frames per second

<sup>3</sup> DCI 4k (full frame) [4096 × 2160 pixels]

<sup>4</sup> *Serial Digital Interface*. Interfície digital de vídeo estandaritzat per broadcast.



### **7.2.3 Servidor Maestro Engine**

Es tracta d'un renderitzador de gràfic avançats que permet produir gràfic, platós virtuals, realitat augmentada o *video wall*<sup>5</sup>.

En el nostre cas, serà l'encarregat de donar l'ordre de generar la retolació per a que aparegui superposada en el vídeo de la notícia que està en emissió.

<sup>5</sup> Multi-monitor que permet crear un array de pantalles formen una única pantalla.

## 8. Planificació

	ACTIVITAT	DURADA
1	PAC 1	
1.1	Definició formal del Projecte	
1.1.1	Definició de la idea i creació de títol	4 dies
1.1.2	Redacció de la Introducció del projecte	2 dies
1.1.3	Descripció, definició del projecte	2 dies
1.1.4	Definició d'objectius	1 dia
1.1.5	Definició de la metodologia a seguir	1 dia
1.1.6	Planificació del projecte	3 dies
2	PAC 2	
2.1	Correccions i <i>feedback</i> amb professor	
2.1.1	Correcció PAC 1	2 dies
2.2	Investigació i Desenvolupament del projecte	
2.2.1	Anàlisi de les necessitats de l'usuari	2 dies
2.2.2	Anàlisi de les tecnologies necessàries	3 dies
2.2.3	Definició de funcionalitats i escenaris d'ús	2 dies
2.2.4	Definició del continguts del projecte i procés de desenvolupament	4 dies
2.2.5	Inici del desenvolupament de prototipat en baixa fidelitat	3 dies
2.2.6	Redacció i entrega PAC2	4 dies
3	PAC 3	
3.1	Correccions i <i>feedback</i> amb professor	
3.1.1	Correcció PAC 2	2 dies
3.2	Desenvolupament del projecte	
3.2.1	Desenvolupament de prototipat en baixa fidelitat	4 dies
3.2.2	Definició de diagrames uml	3 dies
3.2.3	Definició i anàlisi d'Heurística (Usabilitat)	4 dies
3.2.4	Definició de Seguretat	1 dia
3.2.5	Anàlisi possible implementació real	2 dies
4	LLIURAMENT FINAL	
4.1	Correccions i <i>feedback</i> amb professor	4 dies
4.2	Finalització Projecte	6 dies
4.3	Finalització Memòria	6 dies
4.4	Elaboració Presentació	6 dies

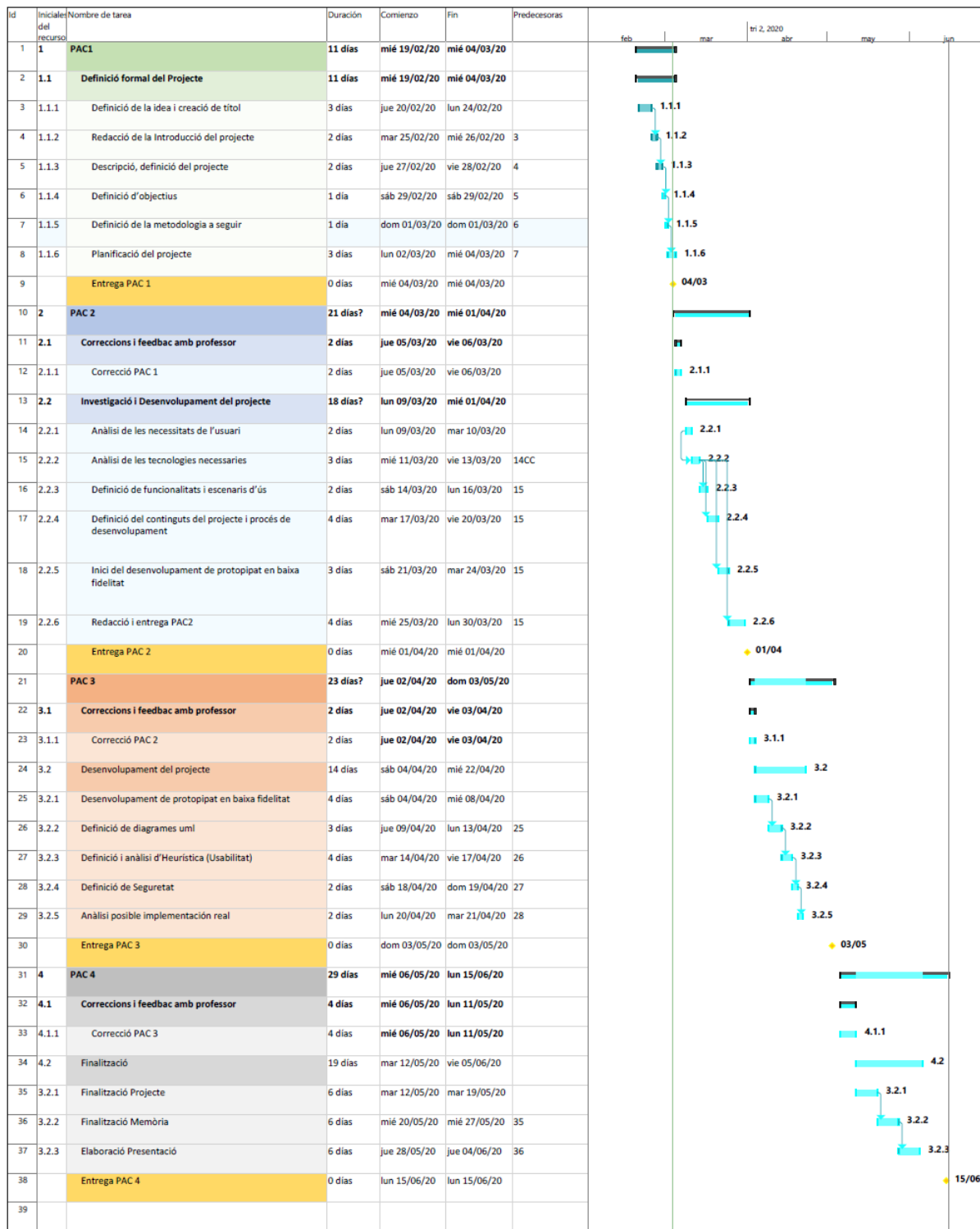


Figura: 2 | Diagrama de Gantt de la planificació del projecte

## 9. Procés de treball/desenvolupament

Aquest projecte neix com a impuls per a cobrir una necessitat per part del usuari de la redacció d'informatius d'una empresa d'audiovisuals. Per a la creació d'aquest projecte s'ha decidit seguir un model de disseny centrat en l'usuari (DCU<sup>6</sup>) en el que el propi usuari es situat en el centre del procés de creació de cadascuna de les fases en les que dividim el projecte.

Una altra de les particularitats del DCU és que aquestes etapes es desenvolupen de forma iterativa poden tornar a fases anteriors amb l'objectiu de garantir l'èxit en la consecució del projecte.



Figura: 3 | Diagrama Disseny Centrat en les Persones

En el nostre cas concret, el TFG es centra en les fases d'investigació i definició del projecte intentant profunditzar en la conceptualització i disseny d'una aplicació de reconeixement d'imatge per tal d'automatitzar la generació de la retolació d'una notícia audiovisual.

Per tal de desenvolupar la idea concebuda pel treball de recerca procedirem amb la definició d'un pla de treball. Aquest passa per una primera fase d'investigació en el que s'ha buscat informació tècnica d'anàlisi de imatge posant èmfasi als aspecte més rellevants pel que fa a la tècnica de reconeixement facial. En aquesta fase també hem procedit a cercar inspiració en altres projectes relacionats amb el reconeixement facial i procediments d'anàlisi d'imatge. Per altra banda hem procedit a realitzar un estudi de mercat per tal d'analitzar quines són les necessitats i objectius reals que cal assolir en el projecte.

Una vegada recopilada la informació hem procedit amb la fase de disseny del procediment mitjançant la creació d'un guió i diagrames conceptuals que ens serveixin com a guia en el desenvolupament de prototip en baixa definició.

Tot el procés d'anàlisi, disseny i avaluació estarà basat en un mètode de disseny centrat en l'usuari (DCU).

<sup>6</sup> ABRAS, CH. et ali 2004

## 9.1 Investigació

### 9.1.1 Usuaris

Tenint en compte el model DCU, en aquesta fase inicial del projecte es fa necessària una feina de recerca i anàlisi dels usuaris i les seves necessitats. En aquest cas partim d'una petició del servei informatiu de la empresa audiovisual on treballa. El client requereix d'un sistema de creació i confecció de retolació automatitzat per tal que l'usuari es beneficiï d'una eina que permeti reduir la càrrega de treball dels periodistes de l'àrea d'informatius i grafistes de l'empresa.

Després de realitzar les diferents reunions calendaritzades amb els departament de redacció i àrea de grafisme d'informatius s'arriba a les següents conclusions en que a les necessitats i objectius.

- Reduir la càrrega de treball que implica la creació de la retolació no essencial.
- Facilitar la integració de retolació automatitzada que permeti introduir noves funcionalitats gràfiques tals com: Transcripció de text, reconeixement d'àudio o anàlisi i detecció d'objectes.
- Reduir el nombre d'errors de retolació produïts per factor humans.
- Simplificar el procés de llançament de la retolació en emissió gràcies a la automatització de la retolació.
- Millorar l'eficiència del sistema de confecció de retolació del sistema informatiu.

### 9.1.2 Sistema de reconeixement facial

Per tal de garantir les necessitats que demanen el usuari procedirem a investigar en el camp del reconeixement facial per tal que puguem automatitzar un sistema de detecció dels rostres dels periodistes de la redacció d'informatius i d'aquesta manera puguem enviar aquesta identificació de la persona al servidor que genera la retolació que apareix a l'emissió dels informatius.

Una de les opcions més factibles és utilitzar una llibreria de codi obert **OPENCV** per a realitzar la detecció facial juntament amb l'adequació de la captura i el processat per tal de reconèixer la identitat mitjançant la comparació de resultats amb una base de dades.



Figura: 4 | Diagrama del procés de reconeixement facial

Per altra banda s'investiga sobre els possibles mètodes o tècniques de detecció de rostres. Finalment, després de la investigació, s'arriba a la conclusió que l'escenari més comú d'enregistrament de notícies

es realitzarà en òptimes condicions de il·luminació i posició amb el que descartem altres tècniques i triem la opció més simple i efectiva que ens ofereix OPENCV<sup>7</sup>: *HaaR Cascade* .

### 9.1.3 Sistema de confecció del rètol

Ens basarem en eines que actualment ja estan en producció. En el nostre cas tenim un sistema *Avid Maestro | News*. Procedim a investigar en la automatització de la captura de les dades de fonts externes al sistema *Maestro | News*.

Segons la investigació realitzada, els sistema permet la vinculació dels camps de dades contra fonts externes de tipus ODBC, EXCEL, XML, JSON, TEXT, RSS o SQL.

En el nostre cas, explorarem la opció SQL per la senzilla raó que ja existeixen els recursos i els servidors dedicats a aquest funció.

►► **To define a data source:**

1. Open the **Data Sources** tab.
2. Right-click in the Data Sources column and select **Add Data Sources**.  
The **Data Source** dialog box is displayed.
3. Define your data source as required.

<b>Factory name</b>	Set the format of the data source; each type requires installation of its own dedicated driver.
<b>Type</b>	Set the data source file type.

**(All other settings derive from the data source type)**

4. Define your data source according to its type. For more information, see:
  - [ODBC Data Source](#)
  - [Database \(SQL\) Data Source](#)
  - [Excel Data Source](#)
  - [Text Data Source](#)
  - [News \(RSS\) Data Source](#)
  - [XML Data Source](#)
  - [JSON Data Source](#)
  - [SMH Data Source](#)

Figura: 5 | Procediment de configuració d'un enllaç amb una font de dades Avid Maestro | News

►► **To define a database data source:**

1. In the **Data Source** dialog box, select DATABASE from the **Type** list.
2. Under **Enter database server location**, type the path.
3. Type the database file name under **Enter database name** (<Server name>\<Instance name>).
4. Type a **User name** and **Password** if necessary (this should be available from the database administrator).
5. Type the **Data source name** as it will appear in the **Data Sources and Queries** dialog box.
6. Click **Test Connection** to verify the connection has been made.
7. Click **OK** to save the data source information.

Figura: 6 | Procediment de configuració d'un font de dades de tipus SQL

<sup>7</sup> L'èxit en el procés de reconeixement queda estretament lligat a les condicions d'il·luminació, posicionament i resolució de la captura. El nostre escenari de captura sempre serà favorable ja que treballarem amb condicions controlades juntament amb el material tècnic que el garanteixi.

► To define a query:

1. Under **Data Sources**, select the required data source.

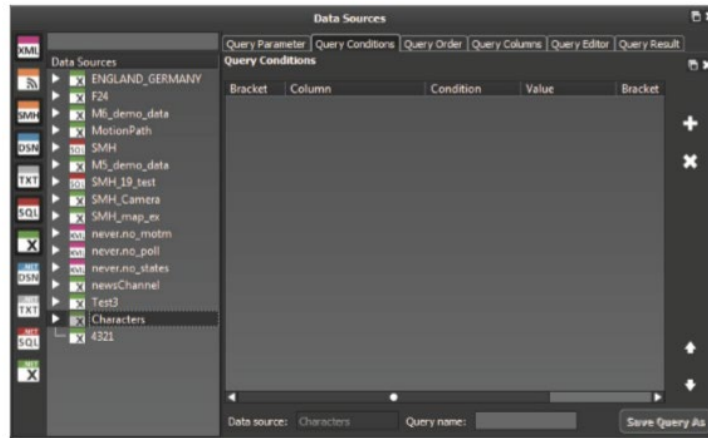


Figura: 7 | Procés de càrrega d'una base se de dades en Avid Maestro | News

Open the **Query Result** subtab. All defined cells from the selected table(s) are displayed. You can select the entire column by selecting the check box in the column header or you can select rows in the Range selection mode (default).

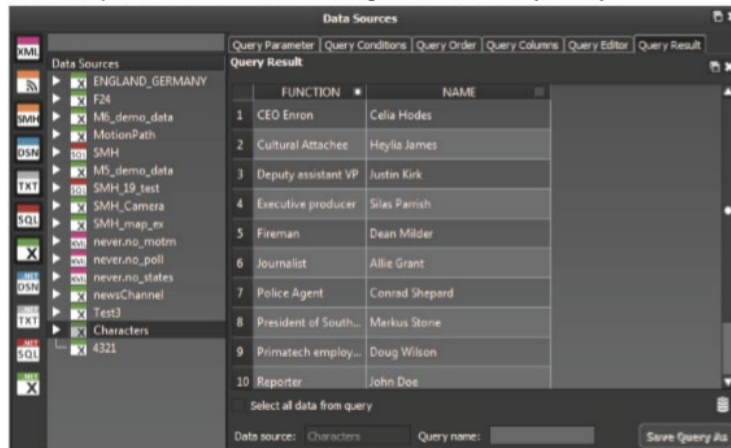


Figura: 8 | Resultat d'una consulta a la base de dades definida

## 9.2 Definició

### 9.2.1 Sistema de captura de vídeo SDI

El punt de partida del projecte a desenvolupar és la captura de la senyal del programa que està en emissió. Partim del supòsit que s'ha iniciar la retransmissió d'un informatiu en directe. El departament de control central, que és qui redirigeix els senyals d'àudio i vídeo als diferents equipaments, en fa arribar dos enllaços de senyal SDI a l'entrada del servidor de reconeixement facial.

Per tal de poder capturar i extreure els *frames* de vídeo es fa necessari la utilització d'una capturadora *broadcast* SDI HD. Farem ús d'una capturadora de la marca Black Magic Decklink per la seva facilitat d'implementació i compatibilitat amb la majoria de plataformes Mac / Windows / Linux.

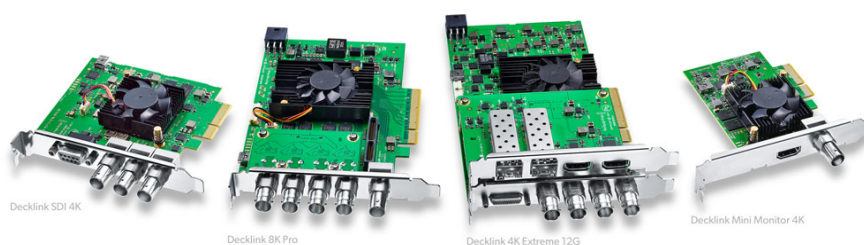


Figura: 9 | Diferents models de capturadora SDI Blackmagic

### 9.2.2 Sistema de reconeixement facial

Definirem un sistema de reconeixement basat en un servidor amb uns requisits mínim d'una CPU multi nucli de 8 cores juntament amb 16 GB de memòria RAM <sup>8</sup> muntat sobre un sistema d'emmagatzemament amb tolerància a errades i optimitzat per a la lectura i escriptura (Raid 10)<sup>9</sup>

En quant al sistema operatiu, OPENCV és una llibreria que funciona tant en Mac, Windows com Linux. En el nostre cas hem decidit optar per Linux gràcies al seu baix cost econòmic i la relativa senzillesa en quant a la configuració i posada en marxa del procés de reconeixement.

Una vegada preparat el maquinari amb el sistema operatiu procedirem a la instal·lació de la llibreria *OPENCV* juntament amb el llenguatge de programació *PYTHON*.



Figura: 10 | Imatgeria dels productes OpenCV i python

<sup>8</sup> *Randon Access Memory*

<sup>9</sup> Sistema de múltiples discs que serveixen per a incrementar el rendiment de lectura/escriptura alhora que permet la tolerància a errades gràcies a la redundància de les dades emmagatzemades.



Pel que fa a la llibreria *OPENCV*, aquesta conté diferents classificadors en cascada per tal de realitzar la feina de detecció de rostres. En el nostre escenari utilitzarem la detecció facial mitjançant el mètode *HaaR*<sup>10</sup> que permet la detecció de rostre, enquadrament, retall i posterior emmagatzematge en format imatge dins dels sistema de fitxers que haguem definit juntament amb un identificador (ID).



Figura: 11 | Diagrama del procés de captura de senyal de vídeo i posterior emmagatzematge

Posteriorment a la captura, configurarem el mòdul de reconeixement pel que fa al entrenament o inserció del models predeterminats que utilitzarà el sistema per tal de comparar amb la resta de rostres que analitzi. Per a dur a terme aquest apartat serà necessari la programació d'un scripts en *Python* que executi una funció ja definida a *OPENCV* amb el que capturarem les mostres que quedaran emmagatzemades en un arxiu. Aquest procés només caldrà fer-ho quan haguem d'introduir un rostre nou d'un periodista en el sistema de reconeixement.



Figura: 12 | Diagrama del procés d'entrenament a partir de l'anàlisi de les captures realitzades

<sup>10</sup> Mètode de reconeixement d'objectes proposat per Paul Viola i Michael Jones a *Rapid object Detection using Boosted Cascade of Simple features* (2001)

Una vegada emmagatzemades les mostres i realitzats el càlculs dels descriptors procedirem a configurar un *script*<sup>11</sup> **recognizer.predict** que compari entre les mostres analitzades (trainer.xml) i les captures en temps real que obtenim a través de la capturadora BlackMagic. El resultat d'aplicar la funció serà una variable en forma de probabilitat de coincidència amb alguna de les mostres emmagatzemades. Aquest resultat s'emmagatzemarà en forma de registre en una base de dades.

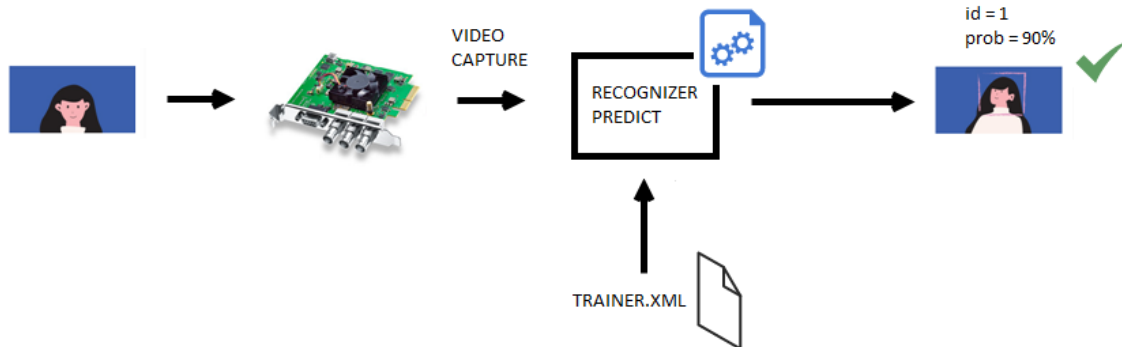


Figura: 13 | Procés de reconeixement facial a través de la comparació amb el fitxer de mostres

### 9.2.3 Confecció del Rètol

Configurarem un servidor amb uns requisits semblants al sistema de reconeixement definit en el anterior apartat. En aquest cas, i per requeriments del proveïdor del programari *Avid Technology*, utilitzarem un sistema operatiu *Windows Server*.

<sup>11</sup> En aquest cas utilitzem el terme script per a referir-se al codi font escrit en llenguatge de programació Python que contindrà les ordres o processos que es realitzaran al executar l'arxiu.

En quant a l'aplicació que confecciona el rètol a partir d'una plantilla existent, configurarem el camp *nom d'usuari* per tal que enllaci amb la base de dades externa. L'aplicació estarà monitoritzant els nous registres que apareguin, fruit del reconeixement facial, y l'aplicació *Maestro Page Editor* captarà els valors del nou registre i confeccionarà el rètol seguint la lògica que s'hagi definit i, posteriorment, s'enviarà al servidor d'emissió.

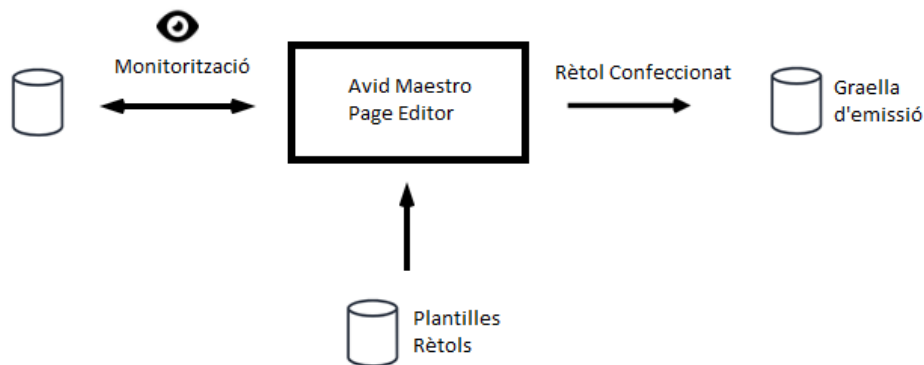


Figura: 14 | Procés de confecció de la retolació generada a través del sistema de reconeixement

#### 9.2.4 Sistema de llançament del Rètol en emissió

Configurarem un equip que monitoritzarà la graella o pauta d'emissió en el que es veu reflectit els diferents rètols confeccionat. És des de aquest equip que definirem quins dels rètols que apareixen a la graella són auto executats al final de cada notícia o si serà l'operador que, de forma manual, executarà l'ordre de llançament.

Gràcies a l'execució d'un servei, el software *Avid Maestro Controller* permet vincular la graella de rètols amb els diferents clips de vídeo de la graella d'emissió. D'aquesta manera, el sistema pot vincular les dues graelles i obtenir dades creuades que permetran interpretar les metadades<sup>12</sup> del clip de vídeo en quant a la durada del clip executant el rètol en el moment adient abans de finalitzar la notícia.

Ja sigui de forma manual o automàtica, el programari *Maestro Controller* enviarà l'ordre al sistema *Maestro Engine* que renderitzarà, en temps real, el rètol i l'incrustarà en el vídeo de la notícia.

<sup>12</sup> Ens referim al conjunt de dades en forma d'informació tècnica o personal que van incorporades en el propi arxiu i que poden ser llegits per aplicacions de tercers.

# 10. Diagrames UML

## 10.1 Flux: Procés d'Aprenentatge Reconeixement Facial

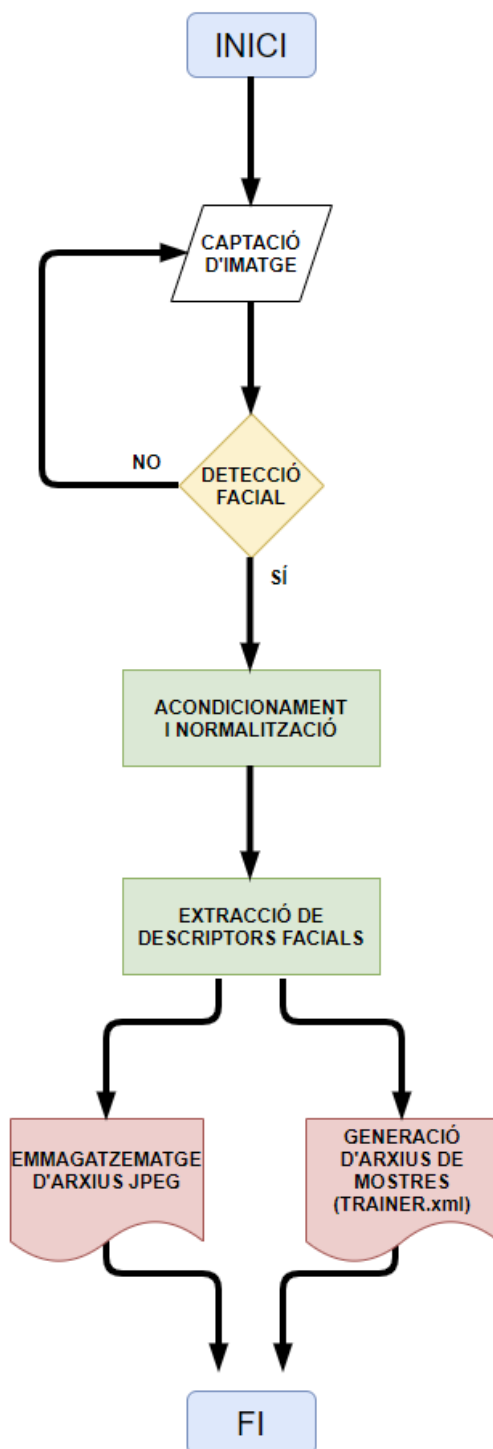


Figura: 15 | Diagrama UML del procés d'aprenentatge

## 10.2 Flux: Procés de Reconeixement Facial

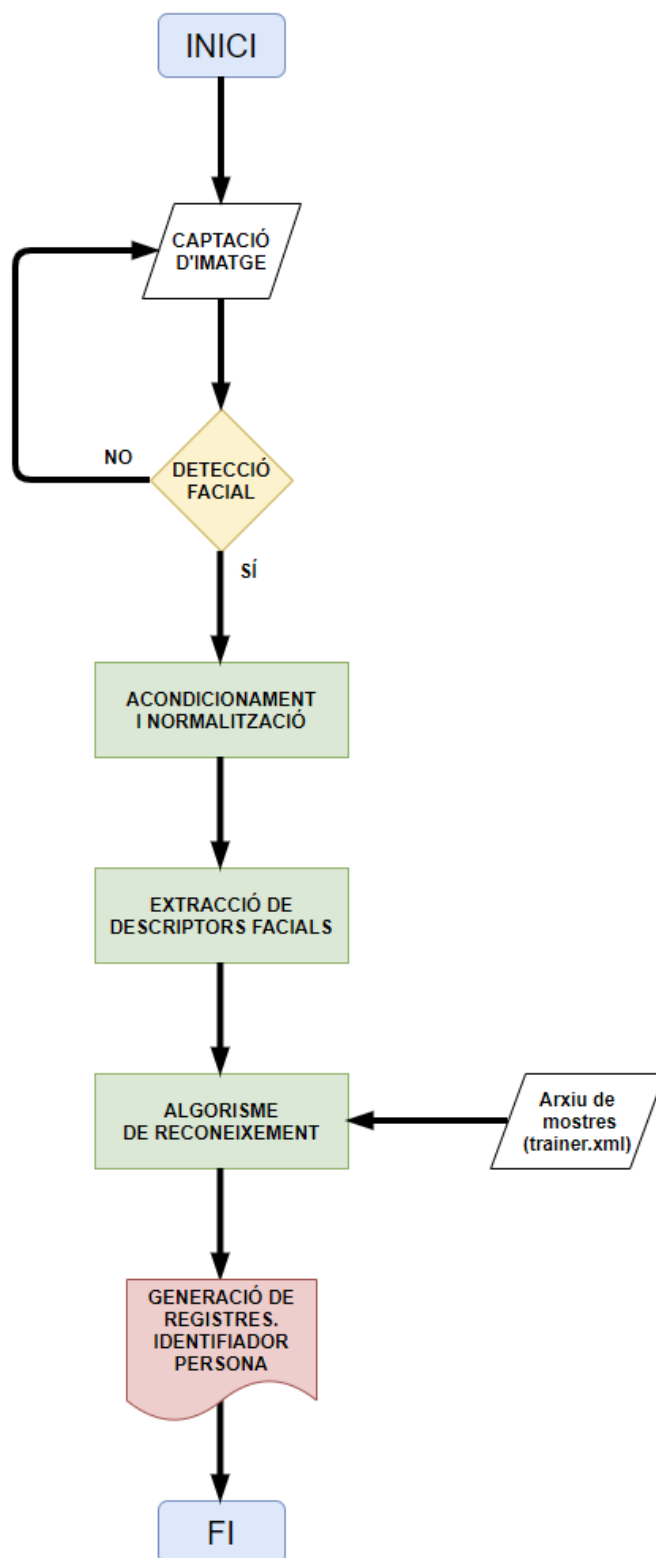


Figura: 16 | Diagrama UML del procés de reconeixement facial a partir de l'arxiu de mostres

### 10.3 Flux: Confecció i Execució de la retolació

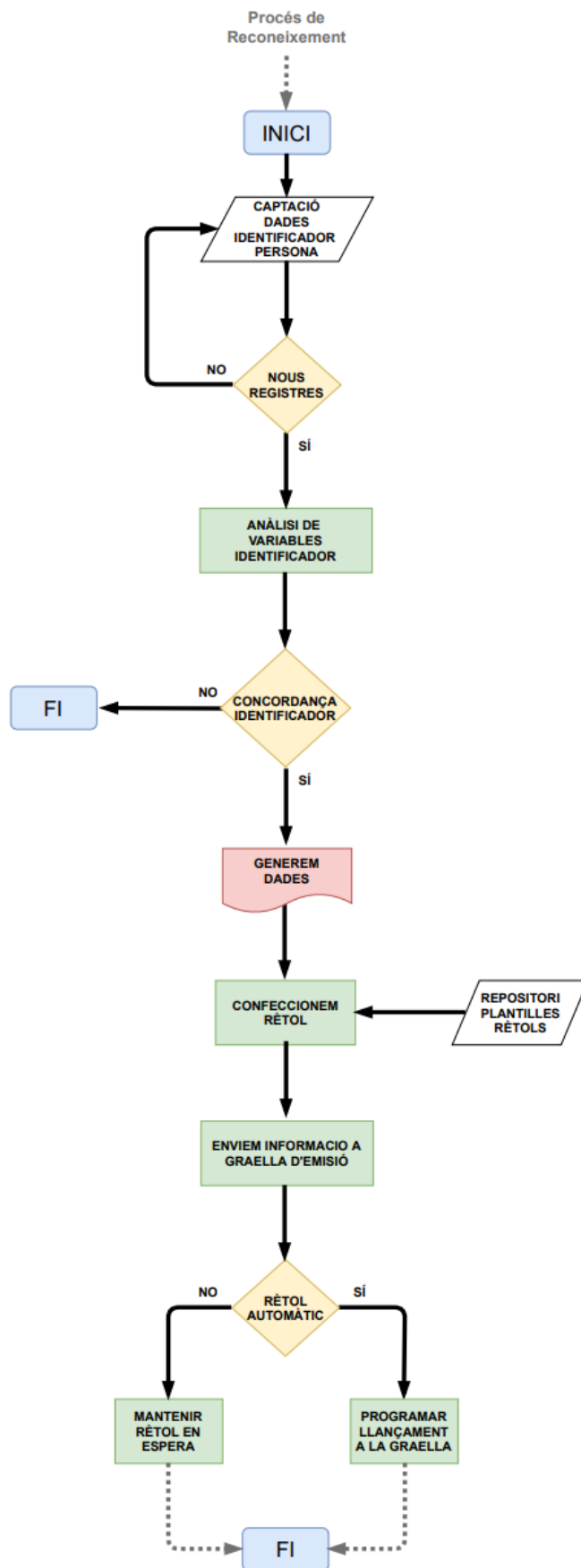


Figura: 17 | Diagrama UML del procés de confecció i execució de la retolació

# 11. Prototips

## 11.1 Lo-Fi

### 11.1.1 Pantalla Principal. Sistema de Reconeixement Facial

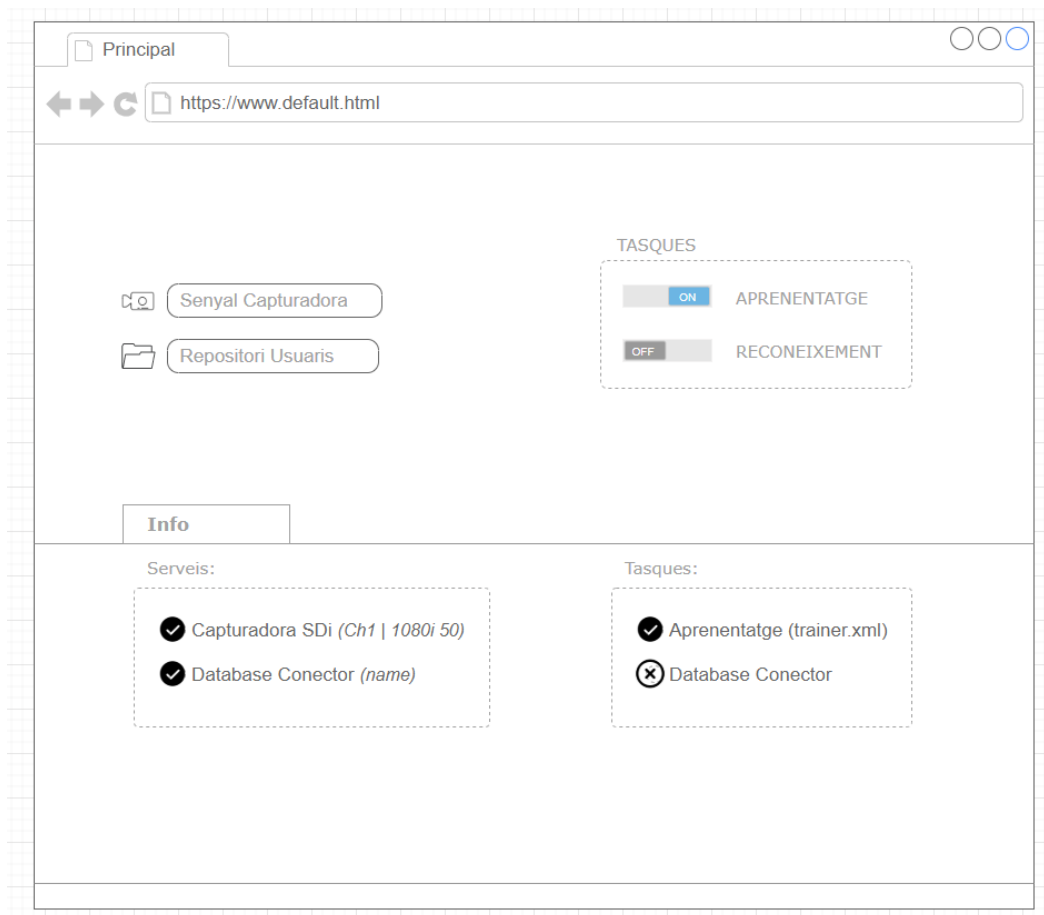


Figura: 18 | Prototip en baixa definició de la interfície del servidor del sistema de reconeixement.

En el menú principal podrem accedir:

#### A. **Senyal Capturadora:**

Podrem seleccionar quina entrada volem capturar en cas que existeixi més d'una. Alhora, la pàgina ens servirà per a verificar si la senyal d'entrada és correcte.

#### B. **Repositori Usuaris:**

En aquest menú podrem accedir a la base de dades que relaciona els identificadors (ID) de cada usuari amb els fitxers de rostres capturats. Alhora, des d'aquesta pàgina podrem eliminar els diferents registres creats.

#### C. **Tasques (Aprentatge):**

Aquesta acció permetrà activar el sistema d'aprenentatge (trainer.xml) de la llibreria *OPENCV* que permetrà iniciar l'aprenentatge d'un nou rostre que vulguem introduir en el repositori.

#### D. Tasques (Reconeixement):

Aquesta acció permetrà activar el sistema de reconeixement (*Recognizer.Predict*) de la llibreria *OPENCV* que analitzarà les imatges del dispositiu de captura per tal de detectar rostres i comparar-los amb la base de dades en busca de coincidències.

#### E. Info:

Llista un resum de l'estat dels diferents serveis per tal de facilitar una monitorització més àgil.

### 11.1.2 Pantalla Senyal Capturadora

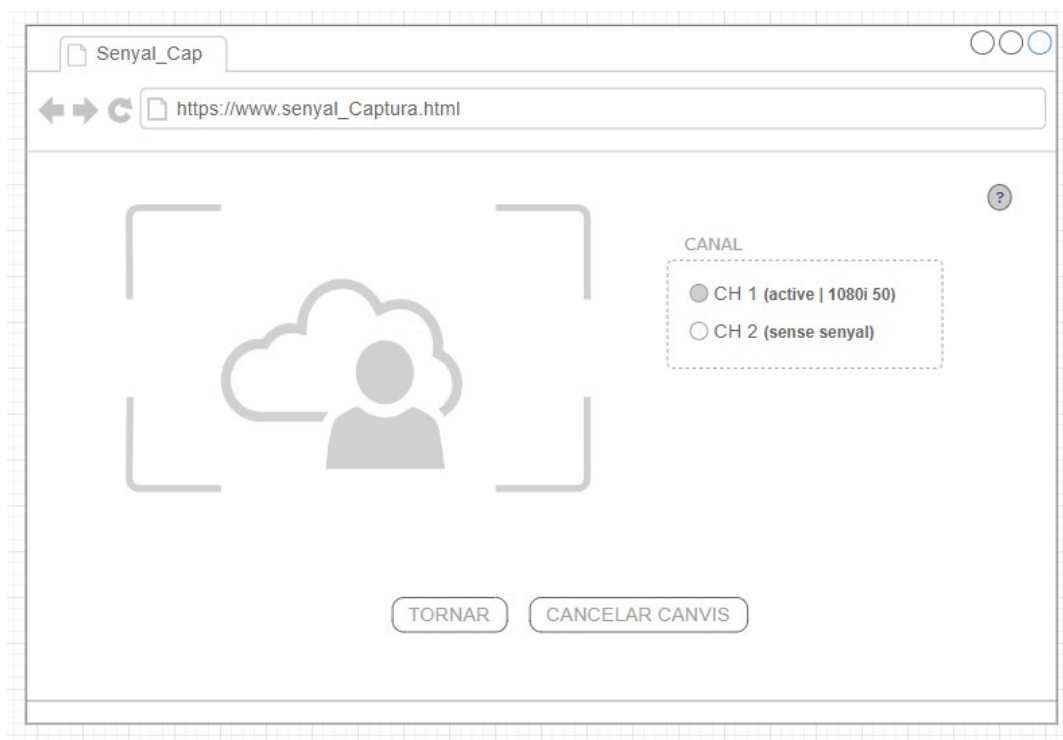


Figura: 19 | Prototip en baixa definició del menú de captura de senyal SDI

Aquest menú ens permetrà visualitzar la senyal d'entrada de la capturadora per tal de garantir que el senyal d'entrada està funcionant correctament. En cas de tenir configurats diferents canals d'entrada, l'operador podrà modificar-ho.



### 11.1.3 Repositori d'usuaris

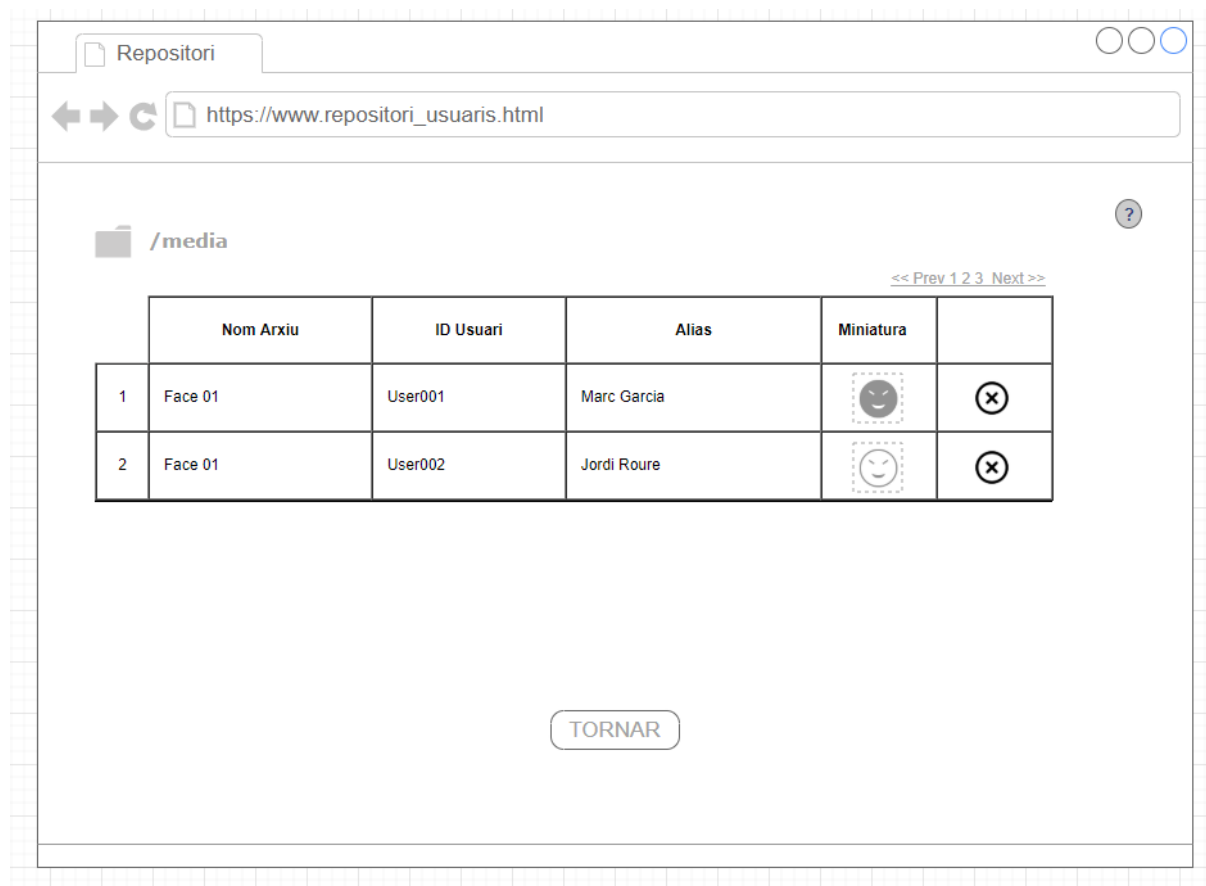


Figura: 20 | Prototip en baixa definició de la interfície de gestió del repositori de rostres capturats

Aquest menú ens permet visualitzar la base de dades de tots els usuaris que s'han capturat en mode aprenentatge. D'aquesta manera podem consultar els registres emmagatzemats i visualitzar si les captures dels rostres estan associats correctament al nom de l'usuari que li correspon. En cas de no ser així, la interfície permet eliminar el registre.

## 12. Perfils d'usuari

Podem distingir dos perfils diferenciats pel tipus d'operació que realitzen. Per una banda tenim els perfils d'usuari d'operació i, per l'altra, el perfil tècnic d'administració:

### A. Redactor / Periodista:

Aquest perfil d'usuari es prestarà amb el seu rostre a iniciar el procés de auto reconeixement i execució automàtica de la retolació del programa que estarà en emissió. La interacció amb el sistema és pràcticament transparent ja que, simplement, ha de prestar-se davant de la càmera i començar a retransmetre la notícia.

### B. Realitzador:

És un tipus d'usuari que intervé en el sistema de forma puntual en la fase de preparació de la emissió. Uns minuts abans de l'hora prevista d'inici del l'Informatiu haurà d'executar l'ordre de monitorització al sistema *Avid iNews* per tal que els diferents servidors que controlen el vídeos i la retolació quedin enllaçats entre si. Amb aquesta acció es permetrà disposar d'una eina d'interacció que englobarà l'estat de tots els recursos que intervenen a la emissió ordenats en el temps.

### C. Ajudant de Realitzador:

Perfil que s'encarrega d'operar el sistema de retolació en el control d'emissió mitjançant la posada en marxa del equip *Avid Maestro Controller* que permet la execució de l'aplicació de control de rètols en emissió. El seu rol és el de validar els rètols generats pel sistema de reconeixement facial aprovant o descartant les entrades que apareguin a la graella d'emissió.

### D. Tècnic Administrador TI:

És un tipus de perfil que no intervé en el procés d'operació al control d'emissió. Les seves tasques estan relacionades amb la operació durant la fase d'aprenentatge del sistema de reconeixement facial, la gestió del repositori d'imatges emmagatzemades juntament amb la seva vinculació amb la base de dades del sistema de reconeixement i la creació de les plantilles de rètols que intervenen en el procés de retolació automatitzat.

Per altra banda també realitzaran una tasca proactiva vetllant perquè tot el conjunt de sistemes sigui operatiu al cent per cent.

# 13. Usabilitat/UX

Amb el disseny d'aquest projecte s'ha intentat marcar una sèrie d'objectius i principis d'usabilitat per tal d'oferir al usuari un sistema que s'adapti a les seves necessitats i que, alhora, s'adapti a la feina diària del conjunt de treballadors de la redacció d'informatius.

Donat que el disseny del projecte s'ha definit en vers a un model DCU<sup>13</sup>, la usabilitat del projecte també hauria de tenir en compte el perfil d'usuari en el procés d'analitzar el problemes de visualització, funcionalitat, ús i interacció que poden sorgir durant les diferents fases de creació del projecte.

Dit això, per tal d'assolir un disseny usable ha sigut necessari l'aplicació dels principis clau de la usabilitat com són la coherència, la interacció, la retroalimentació, el control o les opcions.

Com hem dit abans, per tal d'analitzar la usabilitat aplicada en el projecte cal distingir dos casos totalment diferenciats. Aquest dos escenaris són, per una banda, el disseny de la interfície de control i gestió del sistema de captura, aprenentatge i reconeixement del sistema relegat a un entorn d'usuari més tècnic i per altra banda tenim la interfície d'usuari relegada a un perfil d'operació en el que l'usuari interactuarà amb el sistema per tal de realitzar les accions d'execució i gestió de la retolació del programa informatiu en emissió.

Basant-nos en la definició dels criteris o heurístics proposats per Pearl (1984) i, centrant-nos en el mètode desenvolupat per Nielsen i Molich (1994), procedirem a avaluar un conjunt d'heurístics proposats en el seu llibre *Enhancing the explanatory power of usability heuristics*. Aquest anàlisi consisteix en l'avaluació dels elements de la interfície creades en el prototipat de baixa definició per tal de mesurar si existeixen problemes d'usabilitat en la fase inicial de disseny.

Dividirem l'anàlisi en dos escenaris diferenciats. Per una banda realitzarem l'estudi de la interfície associada a un perfil d'usuari tècnic administrador del sistema de reconeixement i, per l'altra, analitzarem la interfície associada un perfil d'usuari operador que interactuarà amb la interfície d'automatització de rètols.

Per últim, realitzarem un test heurístic amb alguns dels principis claus de Nielsen per tal d'analitzar el compliment dels diferents criteris d'usabilitat de les interfícies prototipades.

<sup>13</sup> Diseny Centrat en l'Usuari

## 13.1 Perfil Tècnic Administració

En aquest cas parlarem d'un perfil d'administrador del sistema d'aprenentatge i reconeixement del sistema. Dins dels diferents models d'avaluació de la usabilitat ens hem centrat en el mètode d'avaluació heurística el qual és molt conegut i utilitzat alhora d'analitzar la usabilitat per inspecció, és a dir, dut a terme mitjançant experts en usabilitat seguint els principis heurístics<sup>14</sup>. Creiem que l'aplicació d'aquest tipus de test sense usuaris és el que més s'adapta a la fase de disseny de prototipat en baixa definició sobre el qual parteix aquest anàlisi.

La interfície principal de gestió del sistema de reconeixement facial implementa el principi de visibilitat de l'estat del sistema gràcies a la utilització d'indicador d'estat *on/off* en funció dels processos inicialitzats alhora que dona informació de l'estat dels servei de captura SDi pel que fa al canal de funcionament (*Ch1, Ch2,...*) i les característiques del senyal d'entrada (*1080i, 720p,...*).

Per altra banda, la interfície permet a l'usuari tenir la llibertat de control de les accions portades a terme amb la implementació d'enllaços per desfer accions, tornar a les pàgines anteriors o eliminar/modificar registres de la base de dades.

Un altre dels principis implementats és el de prevenció d'errors donant informació a l'usuari en quant a senyal de vídeo que està activa en cada canal de la targeta capturadora per tal de garantir la correcta elecció del senyal d'entrada al tècnic que està interactuant amb la interfície.

També s'ha aplicat el disseny estètic i minimalista reduint el contingut d'informació irrellevant o innecessària i centrant-se només en mostrar allò que és útil per a la interactivitat usuari-màquina.

Per últim s'ha implementat un enllaç d'ajuda per tal de garantir a l'usuari una documentació que pugui resoldre qualsevol dubte que interfereixi en la consecució dels objectius pels quals s'hagi dissenyat la interfície.

<sup>14</sup> O dit d'una altra manera: Mètode d'avaluació sense usuaris.

## 13.2 Perfil Usuari Operació

En aquest cas en centrarem en el perfil d'usuari que realitza la operació de gestió del sistema de retolació durant la emissió d'un programa informatiu.

La interfície principal de gestió no és més que una graella de les diferents peces o entrades de producció associades a la pauta que està en emissió. La particularitat de la interfície és la d'agrupar el conjunt de entrades de producció amb els vídeos associats juntament amb la retolació definida per a cada entrada de producció seguint un orde i estructura en funció de la línia de temps d'emissió.

En aquest cas, la interfície mostra la informació de tots els registres associats a cada entrada de producció mantenint un estat de monitorització en temps real amb tots els sistemes associats. D'aquesta manera, la interfície manté informat a l'usuari d'operació del control de realització de tota la informació i processos que estan en execució i d'aquells que es van afegint en temps real.

De la mateixa manera que la interfície permet la visualització de tota la informació en temps real, també permet la llibertat i control de les diferents entrades de vídeo i retolació sent el propi usuari operador de control de realització qui té la capacitat de cancel·lar o executar qualsevol acció.

Per altra banda, la interfície planteja certs paral·lelismes amb utilització de simbologia (*play*, *stop*, *pause*) que utilitzem en el món real i que fan més intuïtiva la assimilació de conceptes i el seu aprenentatge.

Al ser una aplicació crítica en quant a operativa, aquesta implementa mesures per prevenir la execució accidentada d'accions o la d'incloure missatges de confirmació abans d'executar certes accions de risc.

Tal i com hem esmenat anteriorment, la interfície està dissenyada per ser executada en entorns crítics en el que l'usuari estarà sotmès a certs graus d'estrès. És per això que la interfície permet la flexibilitat i eficiència en el seu ús mitjançant la parametrització de dreceres de teclat que facilitin l'execució de les ordres de control.

Per últim podem observar que la interfície ens dona la possibilitat d'accedir a una pàgina d'ajuda en el que observem diferents consells i procediments més habituals en la operació diària juntament amb un fitxer amb tota la documentació referida a l'aplicació i la seva interfície.

## 13.3 Anàlisi Heurístic

Un anàlisi heurístic no deixa de ser una tècnica més per a la valuació de la usabilitat d'una interfície. Dins de l'anàlisi d'usabilitat poder trobar els test amb i sense usuaris. Cal dir que cadascun dels mètodes té les seves virtuts i els seus defectes i és per això que creiem que la millor manera de garantir l'èxit de l'anàlisi és la realització dels dos models.

Donat la fase inicial en quant al prototipat en baixa definició de la interfície del sistema de reconeixement facial que hem definit, hem considerat que el test amb usuaris no és el model d'avaluació més adient i és per això que creiem que un model sense usuaris pot encabir-se perfectament en la fase de desenvolupament en el que està immers aquest projecte.

En el nostre cas, l'anàlisi heurístic permet avaluar la usabilitat a través d'un expert per tal d'obtenir conclusions i possibles millores a implementar.

Procedirem a avaluar la usabilitat del sistema d'interfícies prototipat en baixa fidelitat en busca de les possibles millores. En aquest cas, l'anàlisi consistirà en l'aplicació del deu principis heurístics de usabilitat desenvolupats per Jakob Nielsen

### 13.3.1 Visibilitat de l'estat del sistema

- [A millorar] No existeix un menú a la part superior de cada pàgina que indiqui a quina secció es troba l'usuari. És necessari incloure un menú per tal de mantenir informat l'usuari de la seva ubicació.
- [A millorar] El lloc Web no ens mostra amb quines credencials hem accedit al sistema de reconeixement. Caldrà, doncs, donar visibilitat a l'usuari del perfil amb el que s'ha autenticat a la interfície.
- [Bé] Es valora positivament que la interfície en informi de l'estat del sistema mitjançant imatgeria o botons *on/off*

### 13.3.2 Adequació entre el sistema i el món real

- [Bé] S'utilitza simbologia amb cert paral·lelisme a les convencions que observem al món real.
- [Bé] La informació apareix de manera estructurada i seguint un ordre lògic. La disposició del registres d'informació del repositori d'imatges està organitzat seguint una ordenació per camps que facilita l'enteniment per part de l'usuari.

### 13.3.3 Llibertat i control per part de l'usuari

- [Bé] La interfície permet cancel·lar els canvis realitzats o tornar a la pàgina anterior gràcies a la utilització d'uns botons.

#### 13.3.4 Consistència i estàndards

- [Bé] Existeix iconografia en el Web que segueix les regles que la indústria ha establert per a la interfície d'usuari.
- [Bé] Existeix consistència en quant a la tipografia emprada i no es fa ús de diverses tipografies. Entenem que hauria d'existir una variació d'estil tipogràfic entre les etiquetes i el cos del text.

#### 13.3.5 Prevenició d'errors

- [Bé] La interfície permet el control d'errors gràcies a la implementació del botó "cancel·lar canvis" que prevenen l'usuari de qualsevol error.
- [Bé] A la interfície de selecció de senyal de la captador, el sistema informa si existeix senyal en cada canal. D'aquesta manera l'usuari rep una informació vital per tal de garantir la correcta selecció del canal i permet prevenir-lo de cometre l'error de selecció de canal incorrecte.

#### 13.3.6 Reconeixement abans que record

- [Bé] L'ús d'imatgeria estàndard permet a l'usuari la execució d'accions gràcies a la associació visual amb la tasca a executar.

#### 13.3.7 Flexibilitat i eficiència en l'ús

- [A millorar] No existeixen o no es contemplen dreceres de teclat que facilitin una interacció més ràpida per a un conjunt d'usuaris experts. És necessari que la interfície implementi la possibilitat de parametritzar l'ús d'aquestes dreceres tant per usuaris novell com avançats.

#### 13.3.8 Disseny estètic i minimalista

- [Bé] La interfície no conté informació irrellevant o duplicada cosa que permet centrar l'usuari en el contingut rellevant.

#### 13.3.9 Ajuda els usuaris a reconèixer i diagnosticar els errors i recuperar-se'n

- [A millorar] La interfície hauria de mostrar informació o generar informacions emergents indicant l'estat del sistema o alertant d'algun error o problema i suggerir possibles solucions.

#### 13.3.10 Ajuda y documentació

- [A millorar] La interfície mostra una icona d'ajuda a la part superior dreta de la interfície Web que permet l'usuari d'accedir a una documentació o glossari d'ajuda amb la resolució dels dubtes i procediments claus per a la consecució de les diferents accions que es realitzen amb la interfície del sistema de reconeixement facial.

# 14. Seguretat

Un dels factors claus en el desenvolupament i implementació d'un projecte TIC és la seguretat informàtica. Per tal de garantir l'èxit del projecte i, sobretot, la seva implementació en un entorn de producció 24/7 com és la redacció d'informatius d'una empresa d'audiovisuals caldrà analitzar i dissenyar un entorn de funcionament robust amb la premissa d'intentar minimitzar qualsevol risc de caiguda del sistema ja sigui per factor interns o externs a l'empresa.

En aquest cas es procedeix a realitzar una sèrie d'actuacions en diferents àmbits de la infraestructura dissenyada per tal de garantir la seguretat del sistema i el seu funcionament 24/7.

## 14.1 Seguretat contra intrusions (Intern)

En aquest cas parlarem de la seguretat definida per a intentar minimitzar el riscs de seguretat produïts pels usuaris que operen dins de la empresa realitzant tasques de manteniment (administradors TIC) o la operativa de producció diària (Usuaris d'operació) en els servidors de reconeixement facial, el servidors de composició de retolació o el equips d'operativa de llançament de retolació en el controls d'estudi durant la emissió del programa.

Les mesures de seguretat implementades vindran condicionades pels requisits del diferents programaris que intervenen durant tot el procés. En el nostre cas, les mesures implementades per garantir la seguretat són la instal·lació de programari Antivirus que protegeixi el sistema operatiu de qualsevol atac local produït per l'execució de codi maliciós via dispositius d'emmagatzematge massiu o a través de la xarxa de connexió local.

Per altra banda, implementarem mesures de seguretat en quant a l'accés a la operació del sistema i interfícies mitjançant la obligatorietat de realitzar una autenticació al sistema utilitzant unes credencials d'usuari prèviament consensuades amb l'administrador TI de l'empresa. Més concretament, la idea és implementar la autenticació dels usuaris mitjançant un directori actiu de Microsoft (AD) per tal de centralitzar la gestió dels usuaris i els seus rols. Gràcies a la convergència de tecnologies, optarem per vincular l'autenticació *Active Directory – LDAP* a la interfície de configuració i gestió del programari de reconeixement facial a través de Internet Information Services (IIS) juntament amb el programari *Avid Maestro* amb el que interactuen el usuaris al que va dirigit aquest projecte.



## 14.2 Seguretat contra intrusions (Extern)

En aquest cas parlarem de les mesures encaminades a intentar minimitzar el riscs de seguretat procedents del exterior de les infraestructures físiques de la empresa. En el nostre projecte no està definit un flux de treball que necessiti establir comunicacions amb l'exterior del centre. Encara i així, creiem necessari l'ús d'un sistema de monitorització en temps real que protegeixi de les connexions de l'exterior. Implementarem un tallafocs que realitzi un filtratge del trànsit en les comunicacions d'entrada i sortida tenint especial cura de bloquejar l'accés des de l'exterior al direccionament de xarxa dels servidors i estacions de treball per a diversos serveis i protocols (TCP/IP Capa 3 OSI).

## 14.3 Seguretat en les electròniques de comunicacions

Per tal d'interconnectar tots els equips del sistema, procedirem a utilitzar electròniques de comunicacions *Ethernet* en capa 2 i capa 3. La seguretat de les comunicacions ve donada per la garantia de funcionament ininterromput 24/7 amb la implementació de dissenys de topologia que minimitzin les possibles falles de les electròniques.

Les mesures de redundància en la topologia de la xarxa interna passen per la implementació de electròniques redundades a nivell lògic i físic. Per una banda dissenyarem una topologia de xarxa amb dos electròniques i les enllaçarem entre elles. Per altra banda, implementarem redundància a nivell lògic en aquestes dos electròniques. Configurarem *Cisco Virtual Switching System (VSS)* que ens permet treballar dos equips independents com si fossin un. D'aquesta manera afegim alta disponibilitat, escalabilitat i facilitat en el manteniment al treballar amb una sola configuració virtual que engloba els dos equips. Gràcies a aquesta virtualització d'electròniques podrem implementar la utilització d'un doble enllaç contra els diferents servidors que integren el projecte de tal manera que si una de les electròniques cau, sempre segueix quedant un segon enllaç actiu que permeti mantenir el sistema en funcionament.

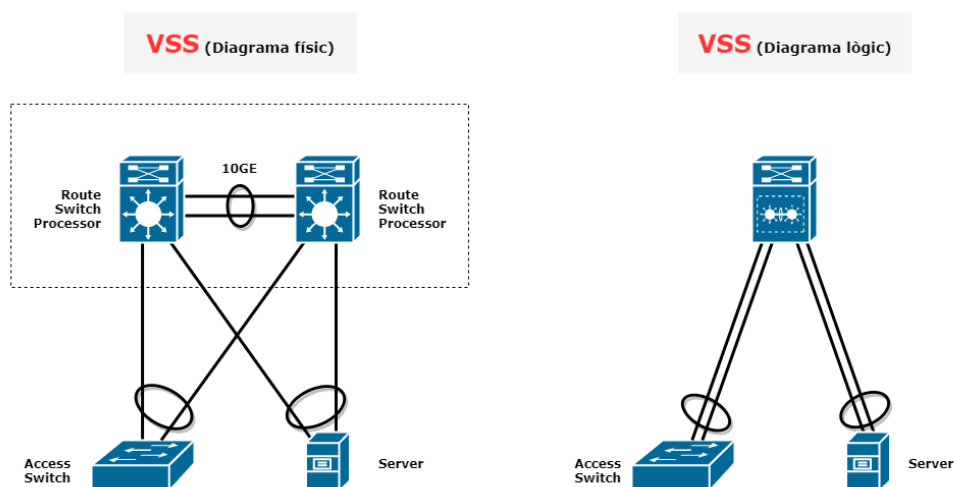


Figura: 21 | Diagrama físic i lògic d'una configuració Virtual Switching System (VSS)

## 14.4 Seguretat en servidors i equips d'operació

Gràcies a la implementació de les electròniques de comunicacions Ethernet VSS podem configurar un enllaç redundat contra les electròniques mitjançant una agregació d'enllaços de comunicacions. Aquestes configuracions reben el nom de *Team* o *Bonding* i permeten la virtualització de dos o més enllaços *ethernet* físics en un de lògic i habiliten la redundància a talls de connexió. També obtenim un increment en l'ample de banda i velocitat de transmissió.

Per altra banda realitzarem la instal·lació per duplicat dels servidors que intervenen en el sistema de reconeixement. D'aquesta manera, afegirem redundància i permetem que el sistema segueixi funcionant en cas que un dels servidors tingui una aturada. La configuració de la redundància passa per configurar un sistema de monitoratge entre els dos equips de tal manera que si un d'ells detecta que el servidor principal està fora de servei, aquest pren el rol de principal i activa tots els serveis per tal de passar a ser l'equip Principal.

Per últim, implementarem la redundància amb el servidor que conté la base de dades i el repositori de fitxers del sistema. En aquest cas la redundància del servidor passa per muntar un sistema RAID5<sup>15</sup> que implica la utilització d'un conjunt de discos d'emmagatzematge formant un sol grup lògic aportant característiques com l'increment de la velocitat d'escriptura i lectura juntament amb l'afegit de la redundància que permet la possibilitat de recuperació del sistema en cas de errors o falles d'algun dels discos que formen el conjunt.

## 14.5 Seguretat en cas de sinistre

Encara que tenim tot el sistema redundat i garantim el funcionament 24/7, no hem parlat d'un factor no menys important que és la recuperació de la informació en cas de pèrdua total o sinistre. En aquest cas, procedirem a implementar un sistema de còpies de seguretat de la informació del repositori de fitxers del sistema de reconeixement juntament amb la informació configurada a les bases de dades utilitzades en el procés de reconeixement. Mitjançant un programari, establirem un sistema recursiu de còpies de seguretat en cintes magnètiques LTO que procediríem a emmagatzemar en un lloc amb certes garanties de conservació per tal de garantir la recuperació en cas de sinistre.

<sup>15</sup> Tipus de d'emmagatzemament de dades que dota de paritat al sistema. Aquesta arquitectura permetrà recuperar-se d'un error de CRC sense perdre la integritat del sistema de dades.

## 14.6 Infografia

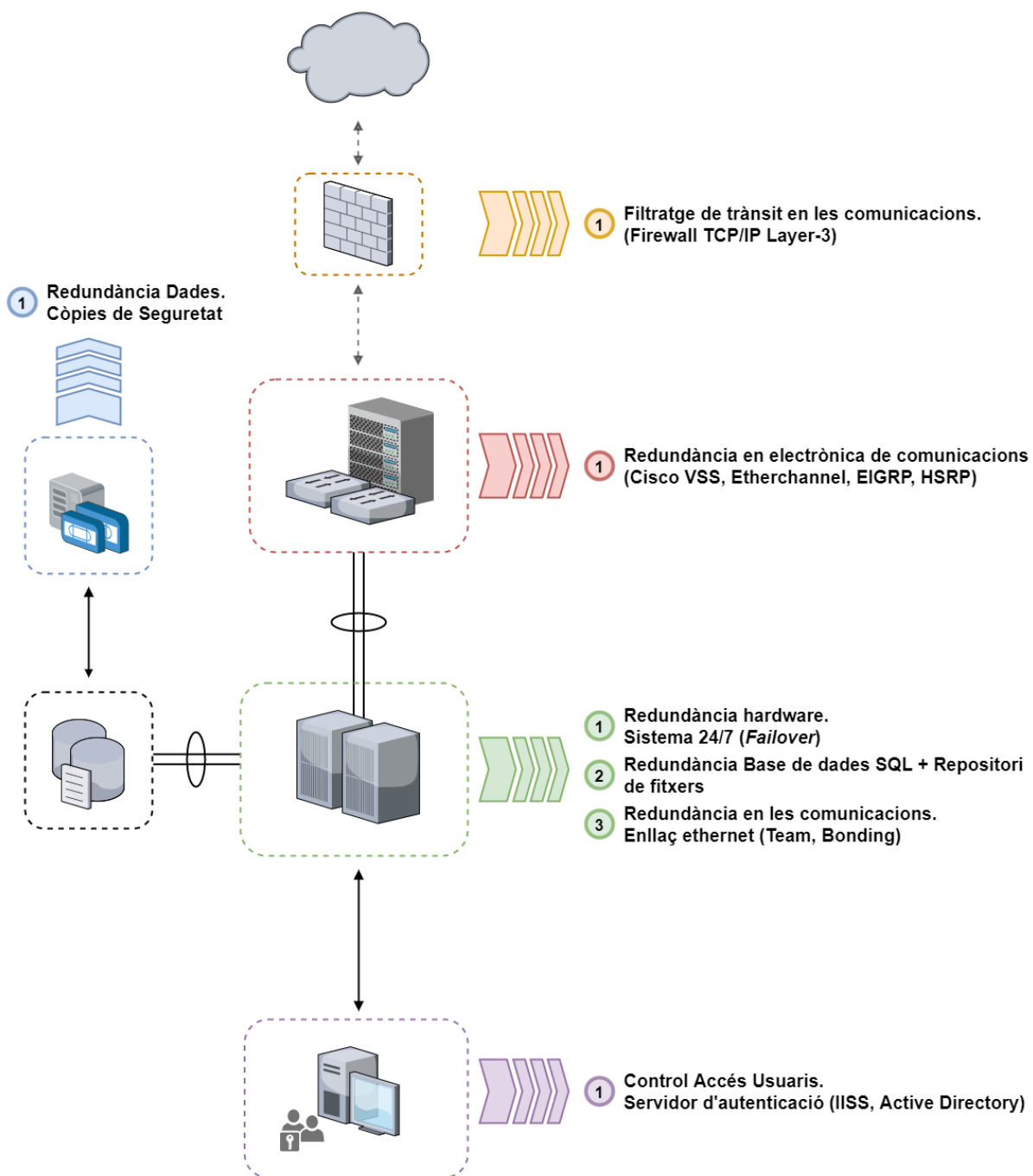


Figura: 22 | Infografia de les diferents mesures de seguretat a implementar en el projecte.

# 15. Requisits d'instal·lació/implantació

Els requisits d'implantació venen determinats pel conjunt de requeriments de programari i maquinari per a cadascun dels servidors que implementen en el projecte juntament amb la formació associada. D'aquesta manera podrem garantir la correcta utilització del sistema de reconeixement projectat.

## 15.1 Sistema de reconeixement

- Maquinari
  - Dos servidors enracables amb uns requisits mínims de CPU *multicore* Intel (4 cores), 4 GB RAM, dos connexions de xarxa *ethernet* 1Gb i una ranura d'expansió PCI-EX 1x en placa base.
  - Dos capturadores SDi PCI-EX en 1080. Com a mínim, un canal d'ingesta i un de reproducció. Ha de ser compatible amb llibreria OPENCV
  - Dos cables BCN 3G-SDi SMPTE 424 per a senyal 1080.  
Quatre cables BCN-6G-SDi SMTPE ST2081 per a senyal 4k
  - Cinc cables *ethernet* Categoria 6. Quatre d'ells per a connexions en *Bonding*<sup>16</sup> contra les electròniques de xarxa i un cinquè que farà d'enllaç *failover* entre equips per establir la redundància del servei.
  
- Programari
  - Dos llicències per a Sistema operatiu Windows Server 2016 Standard
  - Biblioteca OPENCV
  - Microsoft SQL Server 2017 Express
  - Microsoft Web Server IIS
  - Phyton 3.8.2 (Windows)
  
- Formació/Coneixements
  - Documentació oficial d'instal·lació OPENCV i Phyton en entorns Windows Server 2016
  - Manual d'instal·lació del Rol Web Server IIS en entorns Windows Server 2016
  - Documentació i manual d'instal·lació de la Capturadora SDi per a Windows Server
  - Manual d'instal·lació Microsoft Server 2017 Express.

<sup>16</sup> Gràcies a la suma de dos o més interfícies de xarxa podrem augmentar l'ample de banda. Per altra banda també dotarem el sistema de redundància en cas que una de les interfícies caigui per tal de garantir que no existeix pèrdua de connexió en cap moment.

## 15.2 Sistema de confecció de retolació

- Maquinari
  - Dos servidors enracables amb uns requisits mínims de CPU *multicore* Intel (4 cores), 8 GB RAM, dos connexions de xarxa *ethernet* 1Gb
  - Una targeta gràfica NVIDIA Geforce GTX 650 2GB o similar.
  - Cinc cables *ethernet* Categoria 6. Quatre d'ells per a connexions en *Bonding* contra les electròniques de xarxa i un cinquè que farà d'enllaç *failover* entre equips per establir la redundància del servei
  
- Programari
  - Dos llicències per a Sistema operatiu Windows Server 2016 Standard
  - Prerequisits per part de l'integrador del sistema de retolació Avid:
    - HASP Driver
    - Render Engine
    - Microsoft Visual c++ 2005 / 2010
  
- Formació/Coneixements

Encara que el projecte contempla la automatització en la fase de creació de la retolació, caldrà dissenyar i confeccionar les diverses plantilles que continguin els diversos camps que apuntin als registres de la base da dades d'on recuperar la informació resultant de l'anàlisi del sistema de reconeixement facial.

Per tal de realitzar aquesta fase de disseny i creació de plantilles caldrà realitzar una formació específica de l'eina *Avid Maestro 4Designer* alhora que serà imprescindible la entrega de la documentació relacionada amb la formació per tal de poder confeccionar plantilles en un futur.

## 15.3 Sistema de confecció de retolació

- Maquinari
  - Dos servidors enracables amb uns requisits mínims de CPU *multicore* Intel I5 (3470), 8 GB RAM, dos connexions de xarxa *ethernet* 1Gb
  - Tarjeta Gráfica NVIDIA amb arquitectura Maxwell 2.0 amb un mínim de 2 GB RAM
  - Dos monitors FullHD<sup>17</sup> per l'aplicació de control i Previ de Retolació
  - Dos cables *ethernet* Categoria 6 per a connexions amb electròniques de xarxa
  
- Programari
  - Dos llicències per a Sistema operatiu Windows Server 2016 Standard
  - Avid Maestro Controller 7.0
  
- Formació/Coneixements

Es requereix d'una formació específica en Avid Maestro Controller 7.0 per a un perfil d'usuari que farà les funcions d'un ajudant de realització que gestionarà la graella d'execució d'entrades de producció i retolació (automàtica o manual).

<sup>17</sup> Resolució de 1080p i format 16:9

# 16. Instruccions d'ús

Existeixen dos perfils d'ús per al projecte desenvolupat. Per una banda definirem una guia explicativa del funcionament del sistema de reconeixement orientada a un perfil d'usuari que treballarà diàriament amb les diferents eines del sistema de producció d'un informatiu. Per altra banda, definirem una guia amb les instruccions, pas a pas, orientat per a un perfil d'usuari Administrador TI que s'encarregarà del procés de captura i gestió dels registres del sistema de reconeixement.

## 16.1 Resum del Procés d'ús



Figura: 23 | Infografia de les fases que intervenen en el procés de retolació automatitzada

## 16.2 Perfil d'usuari

Partim de la premissa que l'usuari no intervé en el procés del sistema de reconeixement fins a la etapa final d'execució del rètol en emissió. És en aquesta darrera fase en el que els usuaris han estat treballant en la confecció d'una graella de notícies o pauta d'emissió que conté les diferents entrades de producció o notícies organitzades seguint un ordre d'emissió.

L'eina per a realitzar aquesta tasca de confecció de la graella d'emissió és *Avid iNews* i permet la generació i modificació d'entrades juntament amb la creació i modificació de retolació manual.

### 16.2.1 Preparació Emissió

Uns minuts abans de la emissió del programa caldrà que el realitzador carregui la graella confeccionada en *Avid iNews* per tal que els diferents servidors que controlen el vídeos i la retolació queden enllaçats entre si. D'aquesta manera, qualsevol modificació de la graella afectarà immediatament als diferents servidors associats degut a aquest vincle o enllaç. El nom que rep aquesta acció en *iNews* és *Monitorizar* i poden executar l'acció aquells usuaris que tinguin permisos especials sobre la graella del programa.

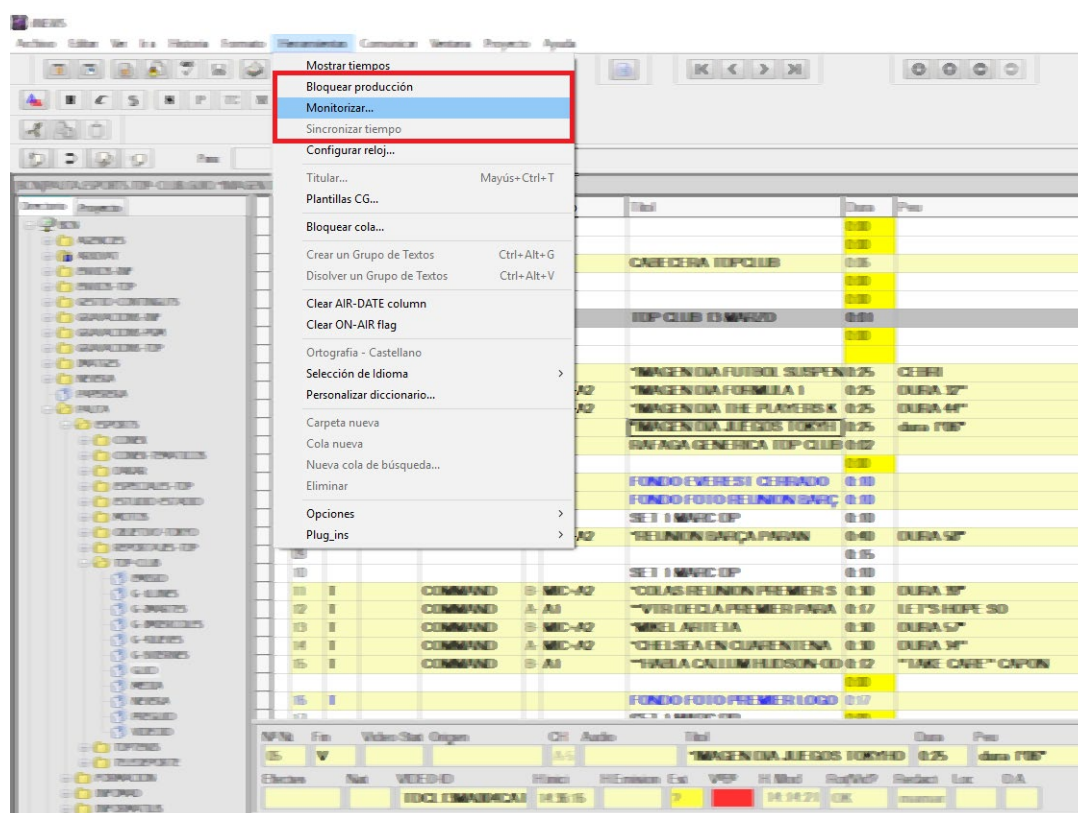


Figura: 24 | Menú que permet executar la monitorització d'una graella a l'aplicació Avid iNews



Seguidament ens apareixerà un menú indicant l'estat de la monitorització de del conjunt de servidors o dispositius que quedaran sincronitzats amb la graella. Procedirem a carregar els dispositius en cas que el monitor estigui *OFF*

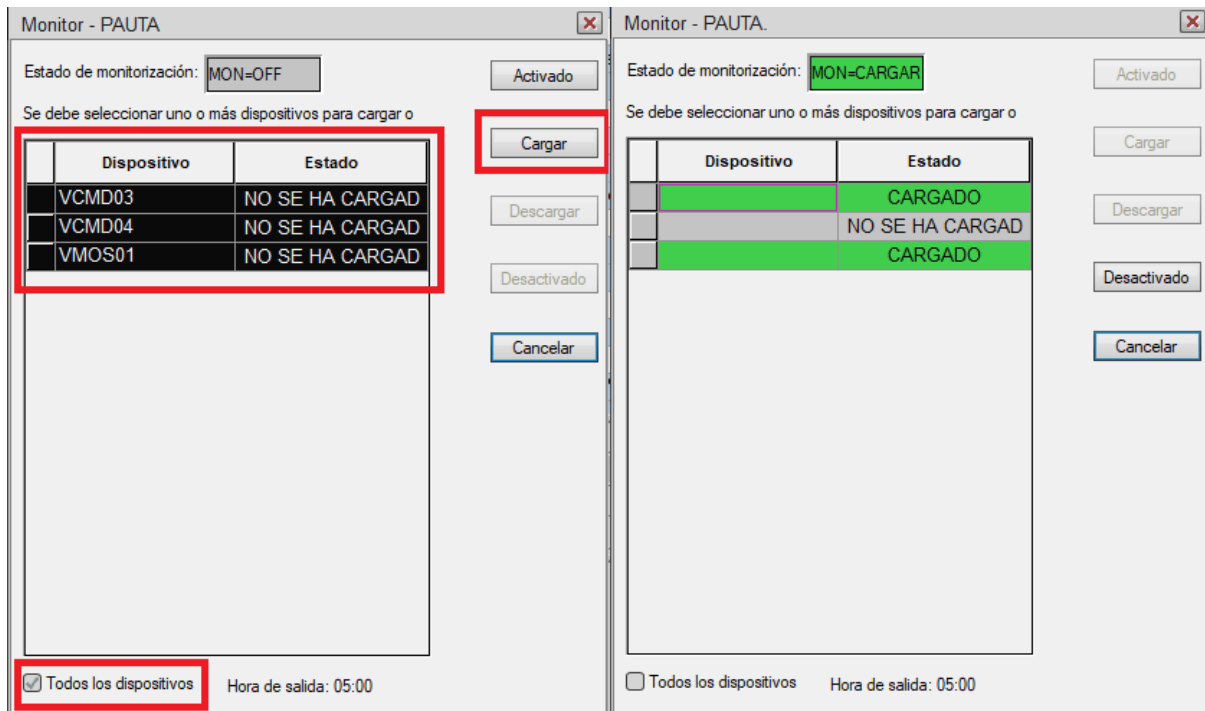


Figura: 25 | Procediment de càrrega del monitor en els dispositius sincronitzats

Una vegada carregat el monitor han quedat vinculats els servidors que emmagatzemen el vídeos *VCMDXX* juntament amb el servidor que controla la retolació *VMOSXX*. Qualsevol modificació de la graella quedarà sincronitzada amb el servidors de vídeo i rètols associats.

## 16.2.2 Sistema de Retolació en Emissió.

Accedim a l'equip de retolació en el control del estudi i ens cerciorem que el servidor té connectivitat amb la unitat de xarxa que conté la base de dades i els rètols.

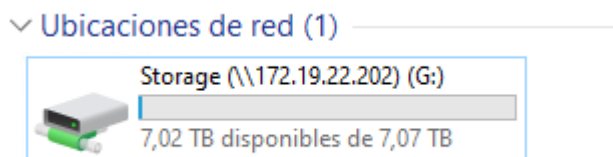


Figura: 26 | Unitat de xarxa requerida per accedir a la base de dades de Avid Maestro

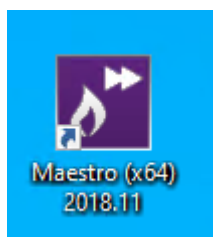


Figura: 27 | Icona del programa Avid Maestro Controller

Una vegada verificada la connexió amb la base de dades procedim a executar l'aplicació de control de rètols en emissió *Avid Maestro Controller* que ens permetrà la gestió i monitorització de la retolació programada en la graella.

Ens demanarà usuari, contrasenya i adreça del servidor que conté la base de dades de rètols.



Figura: 28 | Interficie d'autenticació amb el servidor Avid Maestro Controller

Una vegada introduïdes les credencials accediren a la interfície de treball amb la base de dades dels rètols emmagatzemats.

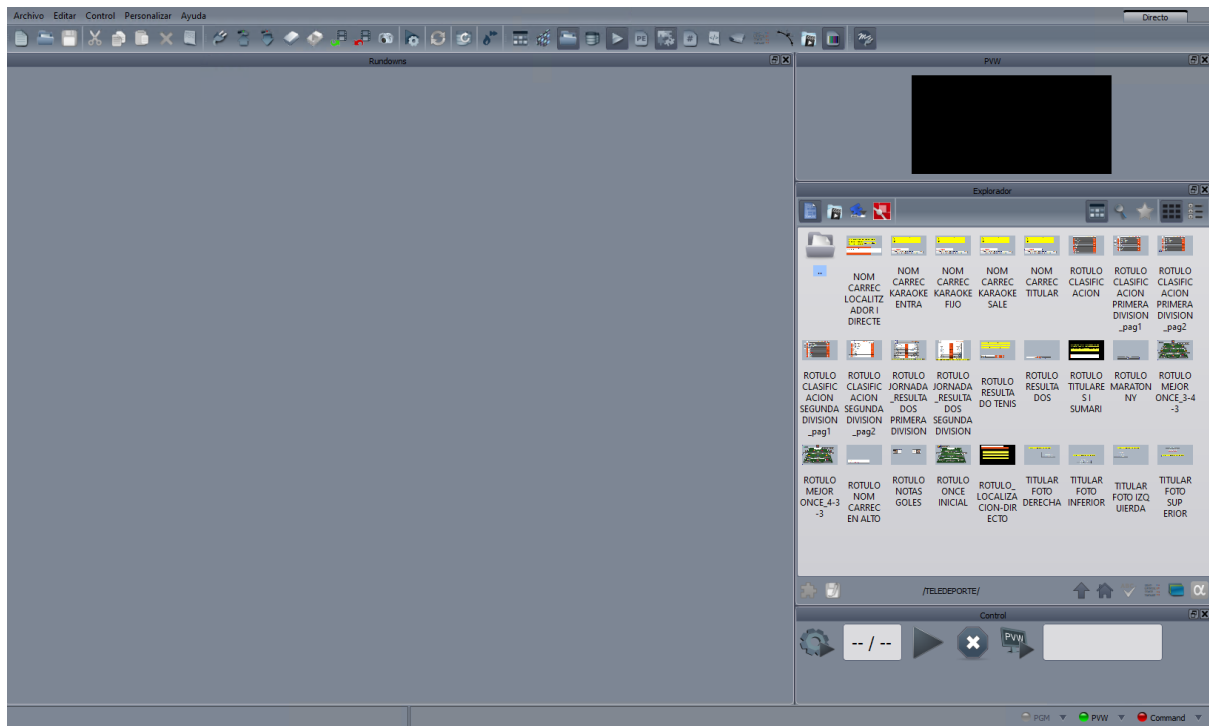


Figura: 29 | Interfície principal del programa Avid Maestro Controller

Tot seguit procedirem a connectar l'equip per tal de carregar les graelles o pautes que estiguin monitoritzades. Procedirem a executar la icona que tenim en el menú d'opcions de la barra superior.

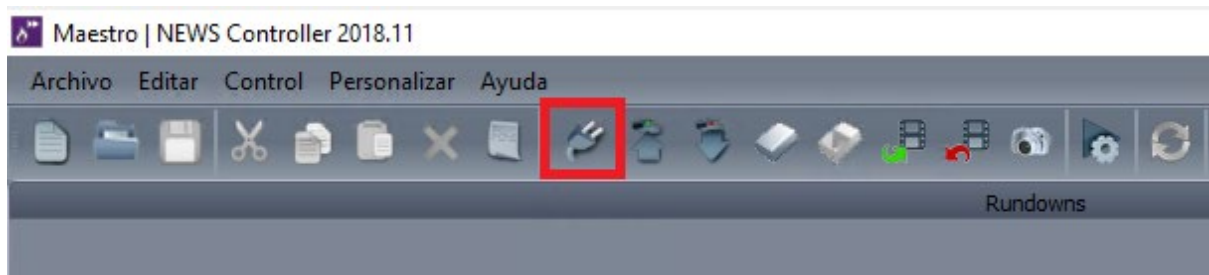


Figura: 30 | Acció per connectar amb el servidor Maestro i carregar les graelles monitoritzades

Al establir la connexió amb el servidor ens apareixerà la graella confeccionada a *Avid iNews* amb la particularitat que la informació que apareix queda restringida només al títols de cada peça que continguin rètols confeccionats manualment per part dels redactors o els rètols confeccionats de forma automàtica gràcies al sistema de reconeixement desenvolupat.

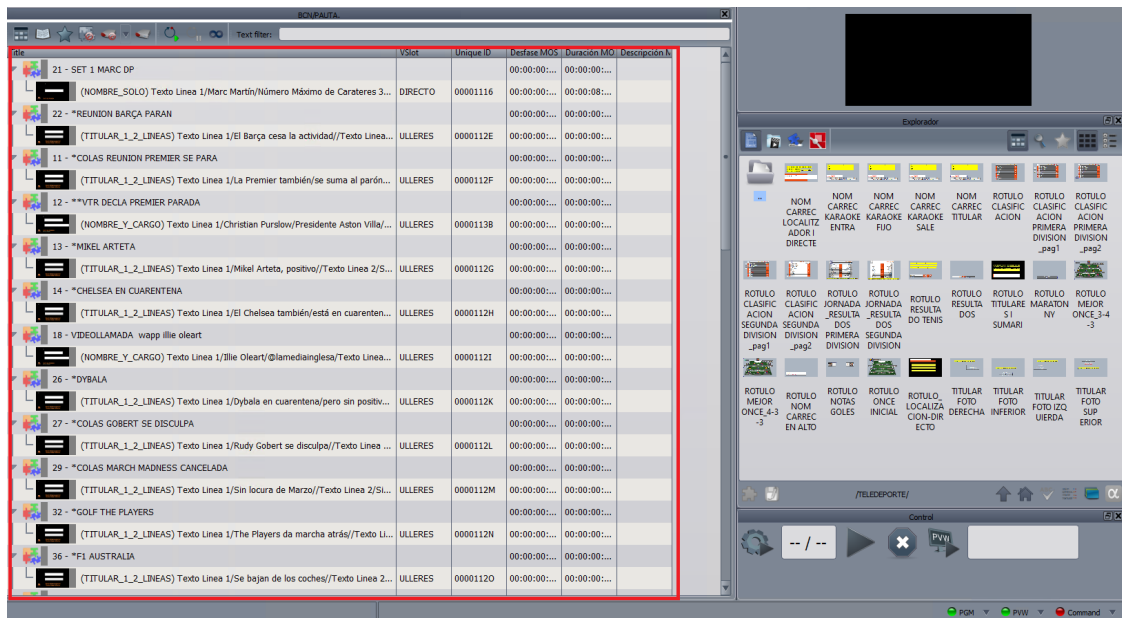


Figura: 31 | Graella de rètols carregada a Avid Maestro Controller

Tant si els rètols són llançats manualment o automàticament, el sistema informa en tot moment de la línia de temps actual juntament amb la previsualització del rètol que està en espera per ser llançat en emissió. Alhora, l'ajudant de realització té el control sobre l'escena amb la possibilitat de reproduir o parar qualsevol dels rètols del sistema d'emissió.

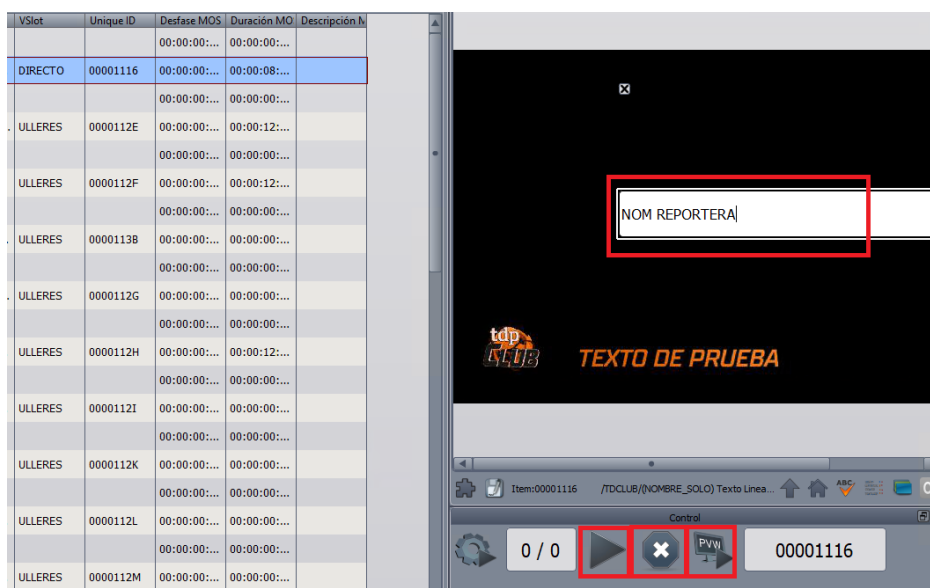


Figura: 32 | Sistema de visualització i operació de la retolació a Avid Maestro Controller

## 16.3 Perfil Administrador TI

Un administrador TI no intervindrà en el procés final de retolació sinó que s'encarregarà de diferents tasques relacionades amb la fase de captura d'imatges, gestió de les bases de dades del sistema de reconeixement i la creació de les plantilles de rètols que intervindran en el procés de retolació automatitzat.

### 16.3.1 Captura d'imatge (Nou registre)

Procedirem a la captura d'un nou rostre per a emmagatzemar a la base de dades del sistema de reconeixement. En aquest cas, accedirem al servidor de la interfície del sistema de reconeixement i habilitarem el sistema per a l'aprenentatge d'un nou rostre captat a través de la senyal SDi de la capturadora.

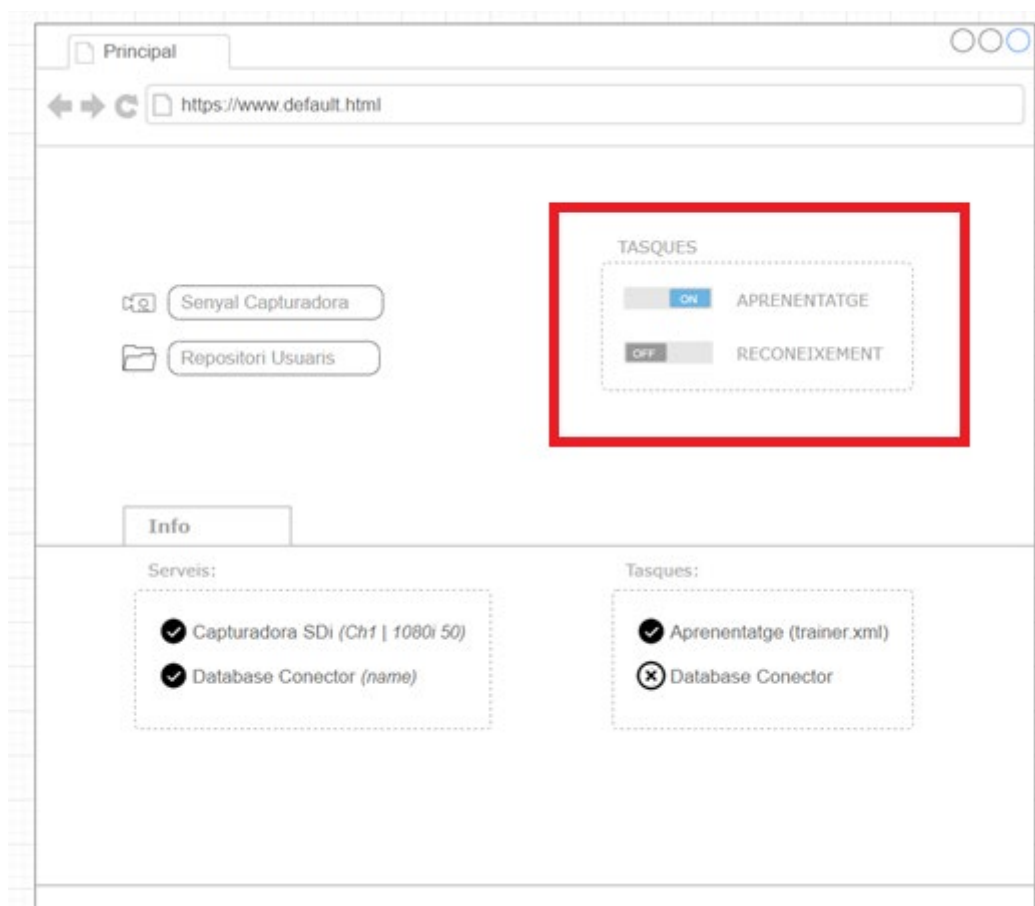


Figura: 33 | Activació de l'aprenentatge a la pàgina principal del sistema de reconeixement.

És important haver verificat que la senyal d'entrada de la captadora sigui la correcta. Accedirem al menú de *Senyal Captadora* i seleccionarem el canal d'entrada que contingui la senyal de vídeo amb el rostre del reporter o reportera a introduir en el sistema. Una vegada seleccionat el canal adient podem tornar al menú principal.

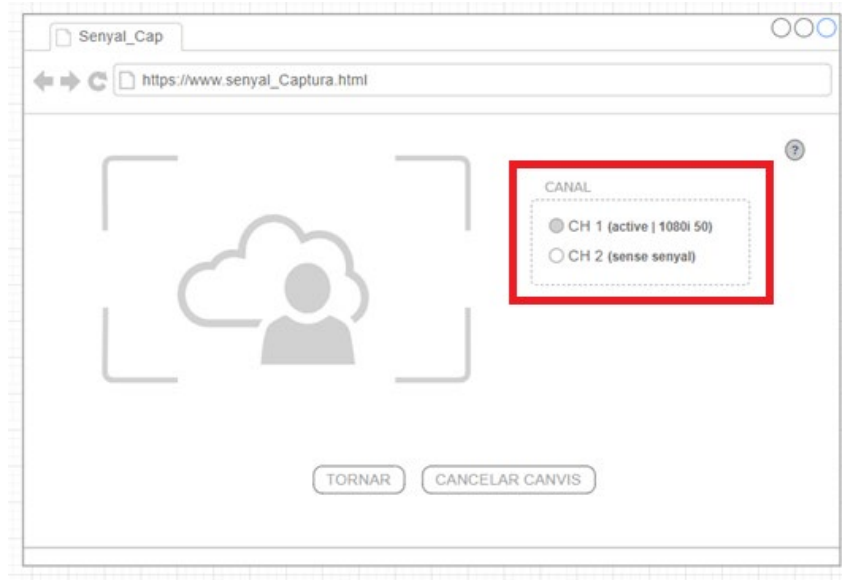


Figura: 34 | Interfície de selecció i previsualització del canal de captura

El sistema està en disposició de realitzar la detecció, captura, adequació i posterior emmagatzematge de les imatges del rostre juntament amb un identificador.

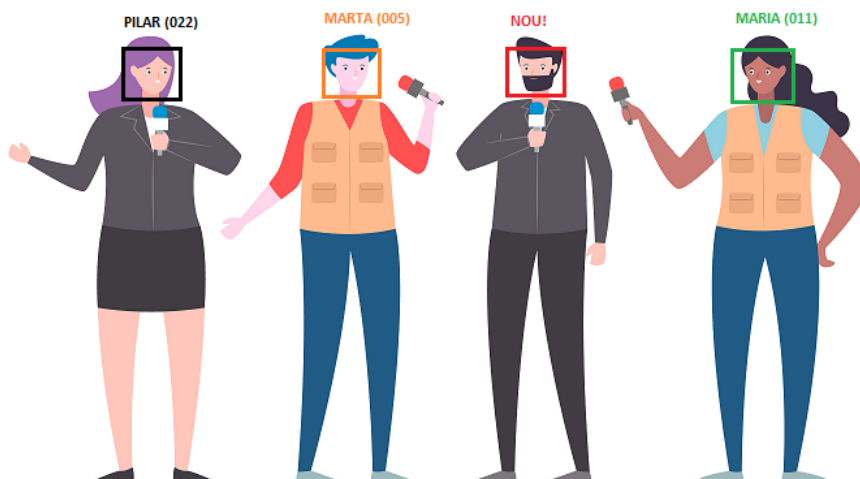


Figura: 35 | Simulació captura del sistema de reconeixement [imatge: Freepik.com]

Una vegada finalitzat el procés de captura accedim al repositori d'usuaris per tal de revisar la base de dades de captures en busca d'errades o falsos positius.

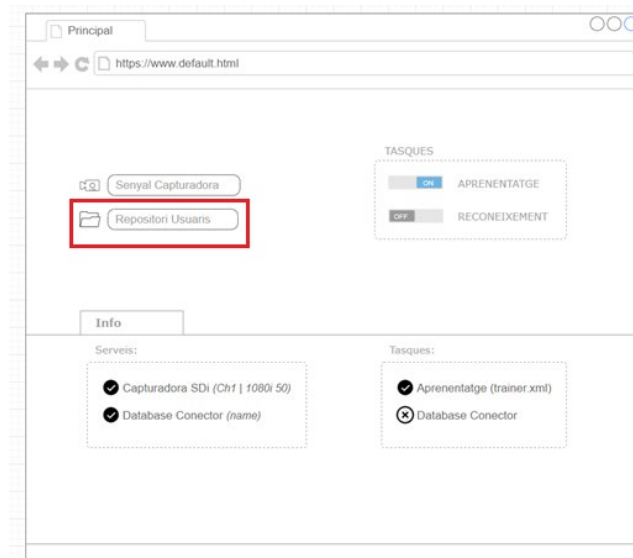


Figura: 36 | Accés al repositori d'usuaris del sistema de reconeixement de rostres.

Una vegada verificats el registres, definirem un alias o identificador que ens servirà per a la confecció automatitzada del rètol.

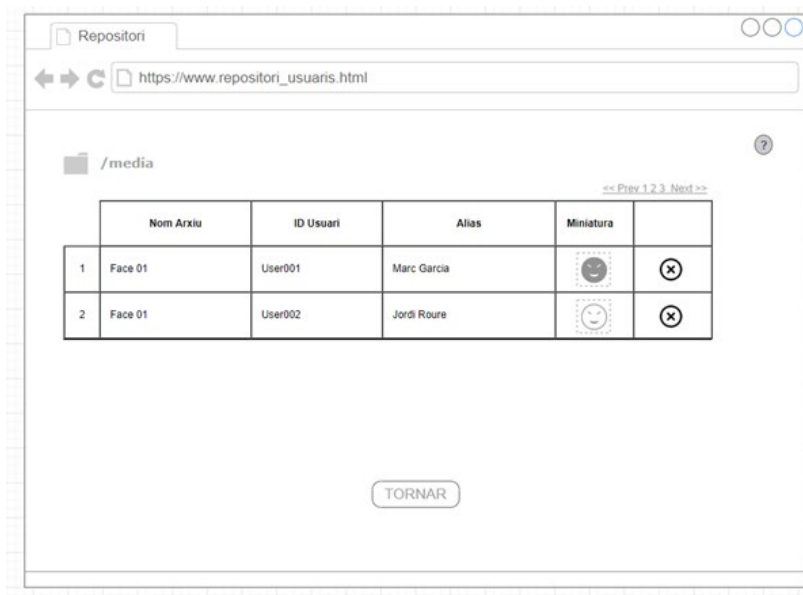


Figura: 37 | Gestió de la base de dades del sistema de reconeixement.

Una vegada haguem finalitzat l'enregistrament de nous registres procedim a tornar al menú principal i desactivar el sistema d'aprenentatge.

### 16.3.2 Definició de plantilles de rètols

Una vegada emmagatzemats els registres procedirem a la vinculació de la base de dades de registres amb les diferents plantilles de rètols que té el sistema. Per això accedirem al servidor *Avid Maestro* i accedirem a la aplicació *Maestro Page Editor*.

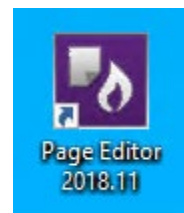


Figura: 38 | Icona de l'aplicació Avid Page Editor

Procedim a obrir una de les plantilles de rètols del programa corresponent per tal de crear una vinculació d'un cap a la base de dades SQL.

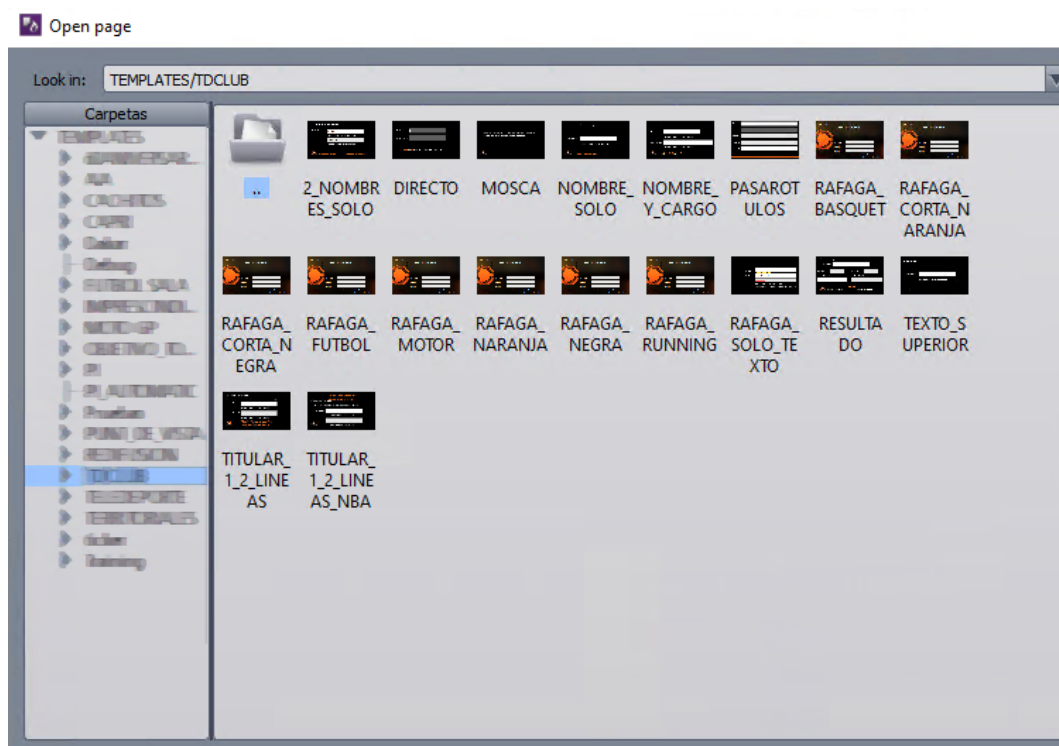


Figura: 39 | Interfície de navegació a través del repositori de plantilles de retolació



Una vegada oberta la plantilla, definim l'origen de la font de dades per el camp de nom del presentador. En aquest cas vincularem el camp a una consulta de SQL que accedirà a la base de dades que el sistema de reconeixement va actualitzant amb els nous camps i els noms dels usuaris a mesura que els va detectant amb el servidor del sistema de reconeixement. Per altra banda, definirem la programació de la plantilla per tal que el rètol es posi en previ uns segons abans del final del vídeo que està en emissió.

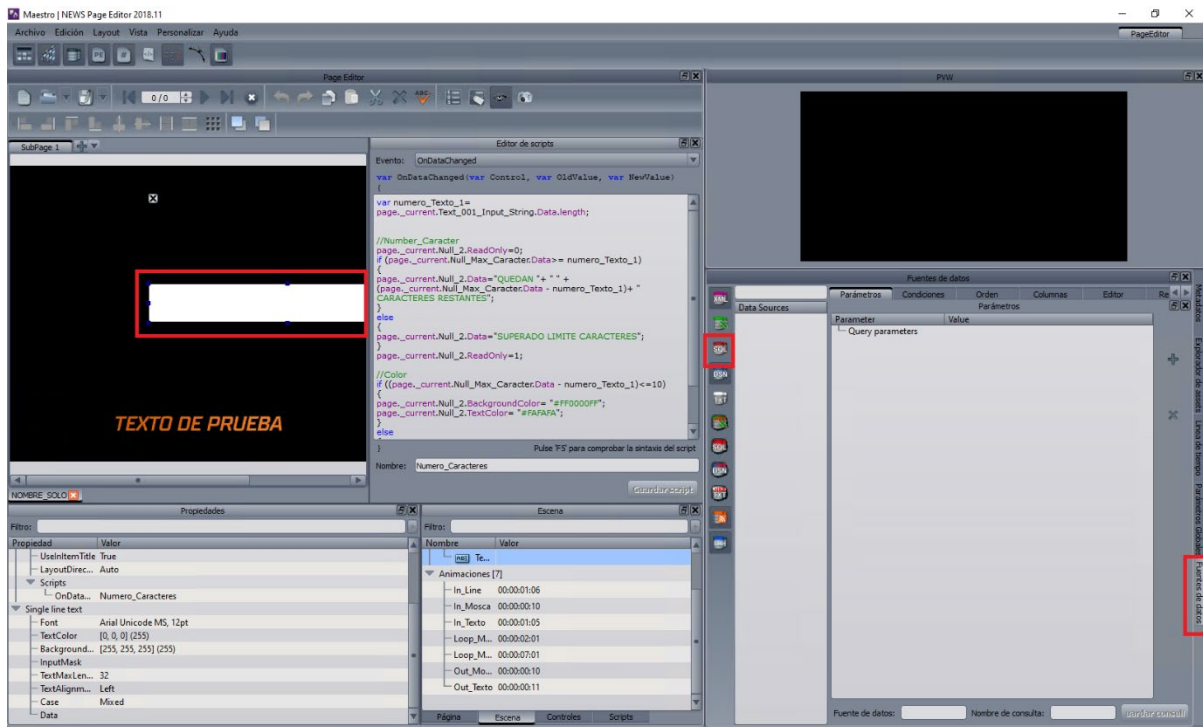


Figura: 40 | Interfície de configuració d'un rètol a Avid Maestro Page Editor

Una vegada configurat, guardem la plantilla a la base de dades del programa.

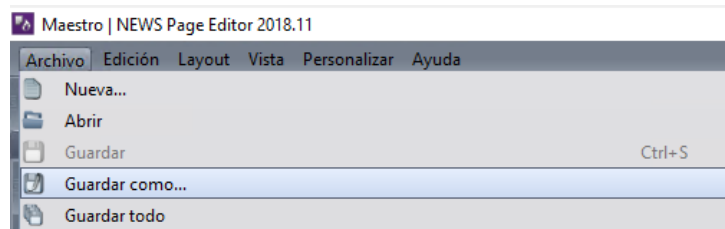


Figura: 41 | Menú per a guardar la configuració del rètol

# 17. Projecció a futur

En tot procés de desenvolupament sempre existeixen noves propostes i funcionalitats que no tenen cabuda dins del projecte actual. Aquest apartat intenta resumir aquelles idees o funcionalitats que han quedat fora de l'abast inicial però que creiem que tenen viabilitat com a projecte complementari de futur.

## 17.1 Reconeixement d'àudio

A més de l'ús del vídeo per al reconeixement facial podríem implementar un sistema de reconeixement d'àudio. Una de les funcionalitats seria la traducció simultània d'idiomes i la seva posterior implementació en la retolació de les notícies. Un dels punts forts d'aquesta funcionalitat és que no requeria d'equipament addicional. Només caldria investigar i desenvolupar el codi font en OPENCV que permeti l'anàlisi de l'àudio capturat.

## 17.2 Reconeixement facial de personalitats

La implementació del reconeixement facial de polítics, cantants o personalitats podria donar lloc a la possibilitat d'automatitzar la retolació i descarregar de feina la redacció d'informatius de la empresa d'audiovisuals. Com en el anterior punt, no seria necessari cap equipament addicional. En el nostre cas caldria cercar vídeos en el fons documental de l'empresa o imatges a Internet i processar-les a través del sistema d'aprenentatge desenvolupat.

## 17.3 Optimització del Reconeixement facial

Durant el procés d'investigació s'ha pogut determinar diverses implementacions i escenaris alhora d'executar un sistema de reconeixement facial amb garanties. El cas és que existeixen diferents mètodes de reconeixement facial que juguen amb les variables de precisió i la velocitat. En el nostre cas s'ha optat per un mètode HaaR<sup>18</sup> que es caracteritza per la velocitat però que no és gaire precís en situacions complexes. En el nostre escenari, les captures estan controlades gràcies a la posada en escena d'il·luminació i l'ús de plans mitjos sense gaire moviment amb el que no la considerem una situació complexa.

Encara i així creiem que podríem optimitzar el sistema mitjançant el mètode DNN<sup>19</sup>. Aquest és més recomanat per a captura de vídeo en temps real amb certa velocitat i precisió gràcies a l'ús de l'acceleració mitjançant targetes gràfiques amb suport CUDA<sup>20</sup>. Creiem que és una opció de futur i més si s'actualitza el sistema amb la convivència de formats 4k o 8k.

<sup>18</sup> *HaaR Cascade* és un algoritme de detecció d'objectes mitjançant l'aprenentatge automàtic amb el que permet identificar objectes mitjançant l'anàlisi de les característiques de cada imatge.

<sup>19</sup> *Deep neural networks* permet una major eficiència en el reconeixement alhora que permet augmentar la velocitat de processat gràcies a l'ús de la unitat de procés gràfic.

<sup>20</sup> Plataforma de computació en paral·lel desenvolupat per la empresa Nvidia que permet utilitzar la GPU per a realitzar càlculs aritmètics.

## 17.4 Catalogació del fons documental

L'actual panorama de la documentació audiovisual està en constant transformació que fa que els esdeveniments es desencadenin a una velocitat vertiginosa. L'actual societat de la informació demanda cada cop més continguts i, per tant, sorgeix una necessitat de preservar la informació d'una manera fiable per tal de poder preservar-la amb garanties. A més, el fet d'emmagatzemar grans quantitats d'informació fa que la catalogació del material esdevingui fonamental per a garantir la seva preservació i agilitat en la consulta i recuperació. És per això que creiem que la automatització de tasques de catalogació mitjançant Intel·ligència Artificial permetria alleugerir la càrrega de treball del documentalistes.

Mitjançant les eines de reconeixement facial implementades en aquest projecte juntament amb la idea de l'anàlisi de l'àudio i el reconeixement d'objectes mitjançant *Deep Learning* podríem resoldre la necessitat del departament de catalogació de l'empresa d'audiovisuals en quant a la catalogació de la mèdia amb la creació de metadades associades pel que fa a:

- Detecció de Persones (Descripció, Nom, URL de la Imatge oficial...)
- Detecció del idioma (Tipus idioma, text OCR, Transcripció del àudio)
- Detecció visual de contingut d'escenes.
- Detecció de sentiments
- Detecció d'efectes d'àudio (silencis)
- Detecció de contingut ofensiu de tipus adult o racial.
- Detecció de *black frames*

Creiem que amb aquesta funcionalitat podríem alliberar de més d'un 50% la càrrega de treball dels documentalistes amb el que es podrien centrar més en tasques més especialitzades.

# 18. Conclusions

La realització d'aquest treball de final de grau m'ha permès desenvolupar una idea gestada al llarg dels últims anys del Grau Multimèdia.

Gràcies a la consecució de les diferents assignatures he pogut obtenir els coneixements suficients per tal d'afrontar amb certes garanties aquella idea que inicialment semblava impossible.

En el meu cas, he tingut la sort de poder compaginar una feina a temps complert en una empresa de comunicació audiovisual juntament amb els estudis de la Universitat Oberta de Catalunya. Fruit d'aquest model d'estudi he pogut assolir amb èxit els meu objectiu que no és més que formar-me per tal de donar valor afegit al meu currículum formatiu per tal de potenciar les meves possibilitats a nivell professional a la empresa per la qual treballo actualment.

El present projecte representa amb claredat aquest simbiosi entre el món laboral i l'acadèmic. La idea de la conceptualització neix d'una necessitat laboral que, gràcies al coneixement adquirit en el grau multimèdia, em permet dissenyar una possible solució.

El treball realitzat pretén fomentar les bases que permetin desenvolupar la idea d'un sistema de reconeixement facial aplicat en un entorn real com pot ser una empresa d'audiovisuals del sector de la comunicació.

A nivell personal, crec que he pogut desenvolupar la idea inicial gràcies al fet de treballar per a la consecució d'un projecte real amb l'objectiu de donar una solució factible a una necessitat diària. La meta real d'aquest objectiu és la que m'ha permès tirar endavant un projecte que ha resultat molt enriquidor en tots els sentits i, alhora, m'ha servit per a motivar-me a continuar amb el desenvolupament d'un projecte factible i viable a curt termini.

# Annex 1. Lliurables del projecte

El projecte inclou els següents lliurables:

1. Memòria del projecte [ *PAC\_FINAL\_mem\_GarciaFiguerola\_Carlos.pdf* ]
2. Presentació del projecte en PowerPoint [ *PAC\_FINAL\_prs\_GarciaFiguerola\_Carlos.pptx* ]
3. Autoinforme d'avaluació [ *Informe\_Autoavaluacio\_TFG\_GarciaFiguerola\_Carlos.pdf* ]
4. Presentació en vídeo [ *Present@* ]
5. Arxius del projecte:
  - a. Diagrama de Gantt de la planificació del projecte [ *Planificacio\_Gantt.pdf* ]
  - b. Diagrames UML:
    - i. Procés d'aprenentatge [ *UML\_Proces\_Aprenentatge\_RF.png* ]
    - ii. Procés de reconeixement [ *UML\_Proces\_Reconeixement.png* ]
    - iii. Procés de confecció i execució [ *UML\_Confeccionament\_Execucio.png* ]
  - c. Prototips en baixa fidelitat:
    - i. Pantalla Principal. Sistema de Reconeixement Facial [ *LoFi\_Principal.png* ]
    - ii. Pantalla Senyal Capturadora [ *LoFi\_Capturadora.png* ]
    - iii. Pantalla Repositori d'usuaris [ *LoFi\_Repositori\_Usuaris.png* ]
  - d. Infografia de la seguretat del projecte [ *Seguretat\_Infografia.png* ]
  - e. Infografia de les Instruccions d'ús [ *Instruccions\_Us\_Infografia.png* ]

## Annex 2. Bibliografia

ABRAS, CH.; MALONEY-KRICHMAR, D.; PREECE, J. (2004). «User-centered design». Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications.

Ortega Santamaría, Sergio. Introdució a la usabilitat i la avaluació. PID\_00176604. FUOC

Avid Maestro | News [en línia] (febrer de 2020) Avid Technology, Inc.

[Obtingut de: <https://www.avid.com/products/maestro-news>]

Berger, W [en línia] (març de 2020). WordPress

*Deep Learning Haar Cascade Explained*

[Obtingut de: <http://www.willberger.org/cascade-haar-explained/>]

BlackMagic DeckLink [en línia] (febrer de 2020) Blackmagic Design Pty. Ltd.

[Obtingut de: <https://www.blackmagicdesign.com/products/decklink>]

Daga, H [en línia] (març de 2020). A Medium Corporation

*Face detection and Image capture with OpenCV and deploying it to AWS Cloud*

[Obtingut de: <https://medium.com/@harshdaga97/face-detection-and-image-capture-with-opencv-and-deploying-it-to-aws-cloud-596fbabb534f>]

Doxygen [en línia] (març de 2020) OpenCV.

*Cascade Classifier tutorial*

[Obtingut de: [https://docs.opencv.org/3.4/db/d28/tutorial\\_cascade\\_classifier.html](https://docs.opencv.org/3.4/db/d28/tutorial_cascade_classifier.html)]

Dwivedi, D [en línia] (març de 2020).

*Face Recognition for Beginners*

[Obtingut de: <https://towardsdatascience.com/face-recognition-for-beginners-a7a9bd5eb5c2>]

Geitgey, A [en línia] (febrer de 2020).

*Machine Learning is Fun! Part 4: Modern Face Recognition with Deep Learning*

[Obtingut de: <https://medium.com/@ageitgey/machine-learning-is-fun-part-4-modern-face-recognition-with-deep-learning-c3cffc121d78>]

Gupta, V [en línia] (març de 2020). Learn OpenCV

*Face Detection – OpenCV, Dlib and Deep Learning ( C++ / Python )*

[Obtingut de: <https://www.learnopencv.com/face-detection-opencv-dlib-and-deep-learning-c-python/>]

MJRoBot [en línia] (febrer de 2020).

*Real-Time Face Recognition: An End-to-End Project*

[Obtingut de: <https://www.hackster.io/mjrobot/real-time-face-recognition-an-end-to-end-project>]

Nielsen i Molich (1994).Nielsen Norman Group.

*10 Usability Heuristics for User Interface Design.*

[Obtingut de : <https://www.nngroup.com/articles/ten-usability-heuristics/>]

Paul Viola i Michael Jones [en línea] (març de 2020).

*Rapid object Detection using Boosted Cascade of Simple features*

[Obtingut de: <https://www.cs.cmu.edu/>]

Python [en línia] (febrer de 2020) Python Software Foundation.

[Obtingut de: <https://www.python.org>]

Rosebrock, A [en línea] (març de 2020). PyImageSearch

*Face recognition with OpenCV, Python, and deep learning*

[Obtingut de: <https://www.pyimagesearch.com/2018/06/18/face-recognition-with-opencv-python-and-deep-learning/>]

Rosebrock, A [en línea] (març de 2020). PyImageSearch

*How to use OpenCV's "dnn" module with NVIDIA GPUs, CUDA, and cuDNN*

[Obtingut de: <https://www.pyimagesearch.com/2020/02/03/how-to-use-opencvs-dnn-module-with-nvidia-gpus-cuda-and-cudnn/>]

Rosebrock, A [en línea] (març de 2020). PyImageSearch

*Object detection with deep learning and OpenCV*

[Obtingut de: <https://www.pyimagesearch.com/object-detection-with-deep-learning-and-opencv/>]

Vakhnenko, H [en línea] (març de 2020).

*How to Make an Image And Face Recognition App*

[Obtingut de: <https://agilie.com/en/blog/how-to-make-an-image-and-face-recognition-app>]

## **Annex 3. Vita**

Carlos Garcia Figuerola és administrador de tecnologies de la informació a Radio Televisión Española amb seu a Sant Cugat del Vallès. La seva tasca principal és la gestió dels sistemes que intervenen en la producció audiovisual juntament amb el suport tècnic als usuaris.

La seva formació professional ha estat estretament lligada al sector de la tecnologia i les ciències informàtiques cursant un cicle formatiu de grau superior en administració de sistemes informàtics i l'actual Grau Multimèdia al qual pertany aquest treball.



