Cas: Vocaloid, un projecte de col·laboració entre el Grup de Recerca en Tecnologia Musical (MTG-UPF) i Yamaha (Japó)

El 1997 Yamaha va plantejar un repte a l'Institut de l'Audiovisual, que va ser la llavor del que després esdevindria el Grup de Recerca en Tecnologia Musical (MTG), vinculat a la UPF: el karaoke vivia una època d'esplendor al Japó, i Yamaha, que tenia una posició de lideratge en el sector, es plantejava desenvolupar una solució que permetés que qualsevol *amateur* pogués cantar com un cantant professional.

El repte plantejat es va concretar en un contracte d'un any. Els primers resultats positius del projecte van facilitar l'extensió del contracte a un any més. Aquest primer projecte va acabar amb un prototip que no era viable econòmicament: el programari desenvolupat permetia canviar la veu del cantant *amateur* per la d'un professional, però aquest enfocament requeria que totes les cançons fossin enregistrades prèviament per un o diversos cantants professionals, la qual cosa encaria molt el procés de producció. A més a més, Yamaha treballava paral·lelament amb una altra empresa per trobar una solució al mateix repte des d'una altra perspectiva que va ser viable.

L'any 1999, en canvi, l'MTG fou el que va plantejar a Yamaha un nou repte: una solució que permetés que músics *amateurs* o professionals que no saben cantar poguessin posar veu a les seves cançons per mitjà d'un sintetitzador de veu. El 2003 va sortir al mercat la primera versió de Vocaloid, però fins a la segona versió, el 2007, coincidint amb l'esclat dels ídols musicals basats en avatars virtuals i la cultura de l'animació, no va esdevenir un èxit comercial rotund, de manera que es convertí en un dels productes estrella de Yamaha.

El producte:

L'aparició de Vocaloid al mercat va ser una revolució en les aplicacions de programari de síntesi de veu, ja que permetia generar cantants virtuals, i contribuí a una nova indústria de l'entreteniment. La tecnologia desenvolupada entra dins la categoria de síntesi concatenativa, la qual es basa en la unió i el processament de fragments vocals extrets de les veus humanes cantades.

El programari proporciona a l'usuari la capacitat de sintetitzar cançons simplement escrivint-ne la lletra i la melodia. Per a crear una cançó, s'utilitza una interfície de rotlle de piano per a incorporar la melodia i les lletres que es poden posar en cada nota. El programari pot canviar l'accent de les pronunciacions, afegir efectes com el *vibrato* o el canvi de ritme i el to de la veu.

El Vocaloid es ven com un «cantant a dins d'una capsa» dissenyat per a actuar com a substitut del cantant real. Aquest enfocament ha creat una indústria pròpia, dirigida tant a músics professionals com a *amateurs*, i permet que altres empreses puguin emprar el sintetitzador de veu per a enregistrar, crear i modelar veus que després posen a la disposició dels músics.

La col·laboració:

La col·laboració entre l'MTG i Yamaha es va iniciar fa més de quinze anys. L'origen de la col·laboració cal anar a cercar-lo en la relació professional entre Xavier Serra, màxim responsable actual de l'MTG, i el qui fa alguns anys fou president executiu de Yamaha, els quals havien treballat junts en un centre de recerca de Yamaha als Estats Units.

La fase inicial de col·laboració amb Yamaha a partir de 1997 va ser clau perquè que es consolidés el Departament de Tecnologia i el grup de recerca en una universitat molt jove. En aquests quinze anys, la col·laboració s'ha consolidat, i el projecte estrella resultant ha estat Vocaloid, que ja va per la tercera versió i se n'està treballant en la quarta. Ara bé, al marge de Vocaloid, en tot aquest període han conviscut d'una manera estable dos o tres projectes de col·laboració menors al voltant de les tecnologies de processament de veu.

La forma de col·laboració durant aquests quinze anys i escaig s'ha estructurat normalment en projectes de dos o tres anys de durada emmarcats en l'estratègia comercial de Yamaha. L'empresa té una clara orientació comercial que es basa a invertir un màxim dos anys en R+D+I, per a acabar disposant d'un producte que es pugui traslladar ràpidament al mercat. Aquests períodes d'R+D+I es divideixen en anys pressupostaris, al final dels quals es prenen decisions sobre la continuïtat i viabilitat dels projectes. Pel que fa al seguiment de cada projecte, mensualment hi ha reunions de control per a retre comptes i tractar els aspectes més rellevants.

Els termes de la col·laboració entre Yamaha i l'MTG s'han basat en contractes de provisió de servei. En el primer contracte s'acordà una modalitat de copropietat de les patents, però ja bastant abans de la primera versió de Vocaloid Yamaha va requerir ser el propietari exclusiu de les gairebé cinquanta patents desenvolupades i permeté que els investigadors de l'MTG hi constessin com a inventors.

Observació i avaluació:

Els més de quinze anys de col·laboració entre Yamaha i l'MTG han tingut la proactivitat del grup de recerca com un dels elements clau perquè, a partir del projecte principal, hagin anat sorgint altres oportunitats entorn de la veu i del so. Aquest marc de treball ha permès de mantenir equips més o menys estables de quatre persones.

Pel que fa a la producció acadèmica, aquesta col·laboració ha estat el context per a diverses publicacions i quatre tesis, malgrat la dificultat de compaginar la recerca i l'evolució del projecte amb la redacció de les tesis. Igualment ha estat una oportunitat per al desenvolupament professional mitjançant l'intercanvi d'investigadors, especialment de l'MTG cap a Yamaha.

Així mateix, la relació consolidada ha fet que el grup de recerca de l'MTG actués com un consultor d'innovació per a Yamaha a l'hora d'analitzar les estratègies

de l'empresa en l'àmbit d'aplicacions de música en l'àmbit de noves tecnologies quant a xarxes socials, mobilitat, etc.

Oportunitats i reptes:

Del projecte principal i del saber fer acumulat n'han anant sorgint noves branques de recerca, que l'MTG ha desenvolupat de manera autònoma i que han arribat crear una empresa derivada. . Aquesta empresa es dedica a crear «llibreries» de veus i sons a partir de tecnologies de processament de la veu desenvolupades pel grup de recerca i fent servir el sintetitzador de veu Vocaloid.

Els resultats i la continuïtat de la col·laboració són especialment rellevants, tenint en compte la gran competitivitat, tant comercial com de recerca, que hi ha en el món de les tecnologies de veu i musicals. Tanmateix, l'experiència recorda que, malgrat que la relació universitat-empresa està molt consolidada, cal mantenir sempre l'excel·lència, perquè en qualsevol moment pot aparèixer un competidor que posi en risc la col·laboració.

Igualment la col·laboració amb Yamaha mostra que la universitat ha de facilitar i reconèixer l'enfocament de l'orientació de la recerca envers l'enginyeria i el servei. Una empresa com Yamaha està essencialment orientada al desenvolupament d'aplicacions comercials, i això es reflecteix en els calendaris dels projectes, normalment limitats a dos anys de desenvolupament. El repte és que, en aquest cas l'MTG, respongui als requeriments, la qualitat i els calendaris que demana l'empresa.