AVATR - AUDIO VISUAL ASSISTANT TURNS REAL

Robert Granich, Herwig Weidle, Toni Homuth, Niels Baumbach

AvatR GbR

AvatR GbR Robert.Granich@AvatR.net

Kurzfassung:

Das Ziel des Projektes "AvatR – Audio visual assistant turns Real" ist die Entwicklung und Vermarktung einer multifunktionalen Software für die Bedienung neuer aber auch gegenwärtiger Multimedia- und Internet-Anwendungen.

Das Produkt erscheint dem Kunden als individueller persönlicher Helfer in Form eines sprachgesteuerten Avatars, welcher vor allem eine engere Identifikation zwischen Anwender und Software ermöglicht. Die kompakte Lösung bietet dem Anwender die Möglichkeit, verschiedene Dienste aus zahlreichen Anwendungsgebieten effizienter als bisher nutzen zu können. Die Verwendung zahlreicher Standard-Technologien, zusammengefasst zu einem einheitlichen Mensch-Maschine-Interface, wird dabei eine Vereinfachung der Erreichbarkeit und Bedienung neuer Medien ermöglichen. Zentraler Aspekt ist die umgängliche Handhabung der Benutzung von Internet-Diensten und Standardanwendungen am PC.

1 Das Produkt

1.1 Produktbeschreibung

Der Virtuelle persönliche Assistent AvatR ist eine Anwendungs-Software, welche sich problemlos auf einem handelsüblichen Windows-PC installieren lässt. Unsere Vision ist es, die Art und Weise den Computer zu benutzen zu revolutionieren. Um dies realisieren zu können, stellen wir dem Anwender einen Assistenten zur Seite, mit dessen Hilfe er sowohl aktuelle Internetdienste als auch herkömmliche PC-Anwendungen bedienen kann.

Sprache ist ein wesentlicher Bestandteil unserer Kommunikation und wird aus diesem Grund auch bei AvatR für die Interaktion mit dem Assistenten eingesetzt. Spracherkennung und Sprachausgabe ermöglichen es, dass zwischen Anwender und Assistent ein Dialog aufgebaut werden kann. Zur Verbesserung und zur Emotionalisierung der Kommunikation erhält die Software eine 3D-Animation in Form eines porträtierten Avatars – "Das Gesicht des virtuellen Assistenten". Der Nutzer erhält zusätzlich die Möglichkeit Aussehen und Stimme des Avatars seinen Vorlieben entsprechend zu gestalten. Beispielsweise stehen den Nutzern sowohl männliche als auch weibliche, individuelle Avatare zur Verfügung.



Abbildung 1: Die Bedienoberfläche der Anwendung

Die Konzeption der Software ist darüber hinaus so gestaltet, dass der Assistent intelligente Rückfragen an den Anwender stellt, um eventuellen Konfliktsituationen vorzugreifen. Somit erreichen wir die Loslösung des Nutzers von der üblichen Art den Computer zu nutzen, hin zu einer freien und konstruktiven Arbeit mit dem Selbigen.

Klare Befehle sowie kurze Sätze dienen der einfachen Steuerung der in AvatR integrierten Dienste, welche in seiner Grundausstattung wie folgt angeboten werden:

Mit der Verwendung von Diensten aus allen Bereichen der Informations- und Kommunikationstechnologie schaffen wir die interdisziplinäre Nutzung von AvatR.

Um dieses Ziel verwirklichen zu können, ist es die Hauptaufgabe von AvatR Anwendungen nicht mehr losgelöst von einander zu betrachten, sondern in ihrer Gesamtheit auf den Anwender abzustimmen. Wir bedienen uns dafür der Möglichkeiten der künstlichen Intelligenz und des Semantic Web¹. Auf dieser Basis ermöglichen wir die intelligente Verknüpfung der Dienste und somit den Schritt in eine automatisierte Benutzung des Computers. Ein Beispiel soll diesen Sachverhalt näher verdeutlichen:

Die Architektur der Software besitzt einen modularen Aufbau. Dies schafft den Vorteil, dass beliebige Dienste, auch durch Drittanbeiter, hinzugefügt werden können und verschiedene Komponenten, wie etwa Spracherkennung oder Sprachausgabe, ohne größeren Aufwand verändert oder ausgetauscht werden können und somit aktuellen Trends gefolgt werden kann. Dieses Konzept ermöglicht es dem Nutzer unter Anderem auch eigene Module und Dienste für den Assistenten hinzuzufügen und zu kreieren.

1.2 Kundennutzen

AvatR bietet die Möglichkeit sich von der bisherigen Nutzung des Computers loszulösen und in eine freiere Interaktion mit dem Assistenten zu treten. Das sich der Nutzer nicht mehr grundsätzlich am Computer aufhalten muss, um Informationen zu erhalten oder Unterhaltungsmedien bedienen zu können, ist ein wesentlicher Vorteil gegenüber herkömmlichen Nutzungskonzepten im Bereich der Informations- und Kommunikationstechnik. Die nahtlose Eingliederung des Assistenten in den Alltag schafft

-

¹ http://www.w3.org/2001/sw/

Freiräume und ermöglicht dessen konsequenten Einsatz beim Lösen verschiedenster Aufgaben. Zentraler Leistungsvorteil, der sich aus der Nutzung von AvatR für die Kunden ergibt, ist die Zeitersparnis, resultierend aus der einfacheren und gezielteren Steuerung. Der Anwender ist zudem in der Lage parallel zu gewöhnlichen Tätigkeiten, wie Frühstücken, Zeitung lesen, Koffer packen, die für ihn wichtigen Information, wie Wetterdaten, Fahrplan-Informationen oder Finanzmarktdaten unkompliziert und schnell abzurufen.

Die semantische Aufbereitung der Daten ermöglicht es in Verbindung mit dem Nutzungsprofil des Anwenders Schlussfolgerungen aus dessen Interessen zu ziehen. Somit können ihm Informationen angeboten werden, die im Bereich seines Interesses liegen. Dies bedeutet, dass auf der einen Seite der Wert der Informationen gesteigert wird. Es werden Verknüpfungen zwischen Informationen hergestellt, die es dem Nutzer ermöglichen einen weitaus größeren Interessenbereich einzusehen, ohne einen zusätzlichen Aufwand betreiben zu müssen. Auf der anderen Seite versuchen wir somit den konventionellen Weg der Informationsbeschaffung im Internet zu revolutionieren, jedoch ohne diesen gänzlich auszuschließen. Wichtigster Nutzen, der sich daraus ergibt ist die Bändigung der Informationsflut.

Nicht nur der individuell anpassbare AvatR und das Interessenprofil tragen zu Individualisierung der Software bei, sondern auch die Vielzahl an Diensten, die dem Kunden angeboten werden. Je nach Nutzerverhalten hat er die Möglichkeit die für ihn relevanten Anwendungen zu nutzen und seinen Assistenten um benötigte Anwendungen zu erweitern. Diese sind natürlich stark von der Art des Anwenders abhängig.

1.3 Technologisches Konzept

Anforderungen wie Anpassbarkeit des Produkts an Nutzerinteressen, Emotionalisierung und Erweiterbarkeit wurden in Analyse und Entwurf der Software aufgegriffen und deren Realisierung durch eine modulare Architektur erreicht.

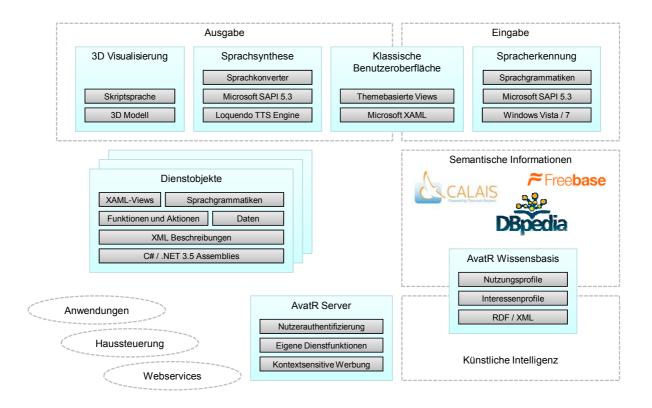


Abbildung 2: Kernmodule von AvatR

Durch die konsequente Softwareentwicklung unter Verwendung diverser von Microsoft bzw. dem $W3C^2$ definierter Standards sind die in der Abbildung dargestellten Module leicht modifizier – und austauschbar. Künftigen Entwicklungen kann durch die Software mit nur geringem programmiertechnischem Aufwand Rechnung getragen werden.

Deutlich wird die Realisierung einer multimodalen Benutzeroberfläche sowohl in Modulen zur Ausgabe als auch in den Formen der Dateneingabe durch den Anwender. Ausgaben an den Nutzer können durch klassische GUI Elemente (Bildschirmfenster, Texte), Sprachausgaben oder 3D Animationen (das "Gesicht" der Anwendung) erfolgen. Der Nutzer hat die Möglichkeit, sowohl Sprache als auch klassische Peripherie (Maus, Tastatur) zur Steuerung der Anwendungsfunktionen zu verwenden.

Eine Emotionalisierung der Anwendung wird durch die Integration eines 3D-Avatars erreicht, der die Sprachausgabe lippensynchron visualisiert. Eine definierte Skriptsprache zur Steuerung eines 3D Modells ermöglicht es darüber hinaus, sprachunabhängige Animationen wie Lächeln oder Verärgerung auszulösen. Die Bündelung dieser Funktionen in einem Modul und der Einsatz einer Skriptsprache zu dessen Steuerung ermöglicht wiederrum den unkomplizierten Austausch des Avatars und somit die Anpassung an Vorlieben des Nutzers.

Die Implementierung von Sprachsteuerung und Sprachsynthese basiert auf Schnittstellen von Microsoft SAPI³ in Version 5.3. Die bisher integrierten, zugekauften Komponenten zur deutschen Sprachausgabe (Loquendo TTS⁴) und die in Windows Vista integrierten Bestandteile zur Spracherkennung können in der Architektur der AvatR Plattform perspektivisch durch beliebige (und unter Umständen verbesserte) Module ersetzt werden, sofern diese ebenfalls auf SAPI basieren.

AvatR nutzt jedoch nicht nur vorhandene Sprachmodule von Drittanbietern, sondern verbessert diese nachhaltig. Der innerhalb der Sprachsynthese angesiedelte Sprachkonverter wandelt Texte und HTML Fragmente mittels eigener Definitionen in SSML⁵ konforme Ausgaben um. Hierdurch kann die Güte der Sprachsynthese erhöht werden.

Zentraler Bestandteil einer effizienten Spracherkennung sind Grammatiken. Diese beschreiben mögliche Befehle und Parameter, die der Nutzer der Anwendung mitteilen kann, um Aktionen auszulösen. Auch hier setzen wir auf definierte Standards: mit SRGS⁶ wird eine Empfehlung des W3C verwendet, um diese Sprachgrammatiken zu beschreiben. Deren Erstellung basiert bei AvatR nicht (wie sonst in sprachsteuerbaren Anwendungen üblich) nur auf vordefinierten Worten. AvatR bezieht dynamische und semantische Informationen aus verschiedenen Datenquellen mit ein und priorisieren Grammatiken (und somit Aktionen) durch die Auswertung von bereits gewonnenen Informationen des jeweiligen Nutzers.

Um die akustischen Komponenten der Spracherkennung möglichst effizient für deren Einsatz in der Applikation vorzubereiten, wird in einer Studie, mit dem Institut für Akustik und Sprachkommunikation der TU-Dresden, die Klassifizierung und Optimierung handelsüblicher Mikrofone durchgeführt. Die Auswertung der Ergebnisse soll dabei der Grundbaustein für eine Phase-Zwei-Studie legen. In dieser werden vor allem Störquellen in Alltagsszenarien erkannt, bewertet und aus dem Nutzsignal gefiltert werden.

Eine nutzerspezifische Anpassung wird aus einer lokalen Wissensbasis generiert. Die Anwendung sammelt während der Benutzung Daten und erstellt hieraus Nutzungs- und

⁵ Speech Synthesis Markup Language

² World Wide Web Consortium, http://www.w3.org

³ Speech Application Markup Language

⁴ Text-to-Speech, Sprachsynthese

⁶ Speech Recognition Grammar Specification

Interessenprofile. Diese werden in standardisierten Beschreibungssprachen wie XML, RDF⁷ und OWL⁸ lokal gespeichert und in verschiedenen Bestandteilen der Anwendung (neben der Spracherkennung auch in Dienstobjekten) wiederverwendet.

Dienstobjekte dienen im Kontext der AvatR Anwendung als Schnittstellen zu Funktionen, die sowohl das World Wide Web als auch lokal auf dem Rechner installierte Anwendungen oder im vernetzen Haushalt befindliche Geräte bieten. Dienste definieren nicht nur Daten, die Sie speichern und Funktionen, die auf Ihnen ausgeführt werden können. Sie legen auch Wege fest, Daten zu visualisieren und Funktionen auszuführen (Sprachgrammatiken).

Ein einfaches Beispiel für ein Dienstobjekt ist ein Wetterbericht. Dieser kann beispielsweise durch das Sprachkommando "Zeige mir das Wetter in Berlin heute" erstellt werden und bezieht nach Auslösen des Kommandos und auswerten der Eingabeparameter die entsprechenden Daten aus dem Internet. Die Visualisierung der Wetterdaten wird über XAML⁹ Beschreibungen realisiert, die grundsätzlich beliebig gestaltet werden können. Die Wege der Datenbeschaffung selbst werden in Programmcode definiert.

Zusammengenommen können diese Teilkomponenten eines Dienstes als Modul betrachtet werden, das sich nahtlos in die Bestandteile der Applikation integriert. Durch die von AvatR vorgegebene Standards und Dokumentationen ist es ohne Weiteres möglich, neue Dienste auch ohne Beeinflussung der Kernmodule in die Anwendung zu integrieren. Diese neuen Dienste gliedern sich automatisch ein und bieten dem Anwender Ihre Funktionen an.

Die leichte Integration ist wichtig, da nur so eine automatische Verkettung von Dienstobjekten ermöglicht werden kann: in Kombination mit den Informationen der Wissensbasis sowie externer semantischer Quellen werden dem Nutzer in jedem Zustand der Verwendung passende ergänzenden Dienste angeboten. Dadurch wird letztendlich eine Beschleunigung von Anwendungssteuerungen im Allgemeinen und Informationsbeschaffung im Speziellen erreicht.

Ein zentraler Server dient zur Realisierung verschiedener Aufgabenbereiche: eigenen Webdienste können ebenso angeboten werden wie eigens aufbereitet semantische Informationen und Datenquellen. Ebenso kann ein Server als Vertriebsweg (ausliefern von Softwareupdates oder neuen Modulen) und in diesem Zusammenhang Authentifizierungsplattform eingesetzt werden. Da durch die Verwendung der Plattform semantisch eindeutige Kontexte zugänglich sind, bietet sich der Server auch zur Auslieferung kontextsensitiver Werbebotschaften an.

2 Das Produkt

2.1 **Produktangebot**

Die durch AvatR angebotenen Leistungen und Dienste sind grundlegend in zwei Bereiche einzuteilen. Auf der einen Seite stellt die Software selbst die Basis für die Anwendung der mitgelieferten als auch der erweiterbaren Dienste bereit. Diese verfügt mit Sprachein- und Sprachausgabe, sowie dem Avatar und der Gesamtheit der Entwicklungen zur Funktionalität, über alle wichtigen Komponenten.

Die von der Basis-Software verwendeten Dienste stellen die Lieferanten von Informationen und Anwendungen dar. Die Dienste für AvatR werden in zwei unterschiedlichen Klassen

⁷ Resource Description Framework

⁸ Ontology Web Language

⁹ Extensible Application Markup Language

angeboten. Es wird Dienste geben, welche direkt mit der Software ausgeliefert bzw. während des Abonnements als monatliches Update dem Kunden angeboten werden. Als Beispiel sind hier Recherche- und Produktsuche-Dienste zu sehen. Weiterhin werden Dienste und Erweiterungspakete angeboten, die eine höhere Qualität an Informationen oder in deren Anwendungen versprechen. Dies können zum Einen erweiterbare Module wie eine verbesserte Spracherkennung oder eine Komponente zur Gesichtserkennung sein. Zum Anderen sind Premiumdienste, wie etwa ein für das Handelssystem XETRA entworfener Dienst, für die Implementierung in AvatR verfügbar.

Das Angebot des Produktes ist ebenfalls variabel. Der Kunde kann die Entscheidung treffen Komponenten für die Benutzung der Software, wie etwa Headsets, Mikrofone oder Webcams, einzeln oder als vollständiges Set hinzuzufügen.

2.2 Vertrieb

Der Vertrieb des Produktes wird in drei Kanälen realisiert. Es wird eine Downloadmöglichkeit auf der Homepage von AvatR¹⁰ zur Verfügung stehen. Über diese kann sich der kaufinteressierte Kunde das Programm auf seinem Computer speichern und benutzen. Dieser Distributionskanal besitzt den Vorteil, dass das Verteilen der im Produkt enthaltenen Lizenzen sehr einfach geregelt werden kann. Zudem ist die Form des Warenaustausches sehr kostengünstig und verbirgt dabei keinen Mehraufwand für den Kunden. Einzig und allein eine Vertrauensbasis für die Onlinebezahlung muss im Vorfeld durch eigene Kampagnen oder die Zusammenarbeit mit Vertrauenspartnern¹¹ geschaffen werden. Da AvatR kein statisches Produkt ist, sondern sich stetig weiter entwickelt und der Kunde weitere Neuheiten für das Produkt abonnieren soll, ist das Angebot über eine Onlineplattform optimal zu offerieren, da alle Möglichkeiten, inklusive Erklärung und Support, für den Kunden in einem Zug greifbar sind. Zudem stehen zusätzliche Möglichkeiten zur Visualisierung des Produktes zur Verfügung. Ein unter Umständen schwer definierbares Produkt wird so anschaulicher präsentiert, wodurch der Kunde wiederum zum Kauf bewegt wird.

Die Wahl der Zielgruppe der 30 bis 49 Jährigen erfordert es, das Produkt ebenfalls auf einem konventionellen Weg anzubieten, um Käufern, welche eher den Weg in den Fachmarkt als in den Online-Store wählen, gerecht zu werden. Für ein gelungenes Angebot in einem Fachmarkt ist die Möglichkeit der kompletten Präsentation ein wesentlicher Vorteil gegenüber einem Online-Angebot. Stellt man dem Kunden im Fachmarkt die Möglichkeit zu einem ausführlichen Test des Produktes zur Verfügung, kann sich dieser zum Einen von dem Produkt überzeugen und zum Anderen einen Einblick in die große Spanne hinzu kaufbarer Erweiterungen gewinnen. Ein geschulter Mitarbeiter, der auf Anfrage das Produkt vorführen und erklären kann, würde dieses Angebot zusätzlich aufwerten. Das dritte Standbein für einen erfolgreichen Vertrieb von AvatR stellen Testversionen der Software auf Zusatz-DVDs in Zeitschriften wie Computerbild, PC Magazin oder aber auch Medienmagazinen wie der Fernsehzeitung TV Movie dar. Die Leser dieser Zeitschriften werden sich eine Testversion von der beigefügten DVD herunterladen und in einer Trial-Version testen. Sofern die Leser Gefallen an dem Produkt finden, werden sie dieses abonnieren und alle Vorteile der Vollversion von AvatR nutzen. Diese Art des Verkaufes besitzt den Vorteil einer sehr unproblematischen Verteilung. Zudem wird dem Kunden ein einfaches Herangehen an die Software ermöglicht, da dieser nicht mehr als die Installation zu tätigen hat, um die Software testweise zu nutzen. Für das Angebot der Software ist zu entscheiden in welchem Umfang diese angeboten werden kann. Eine Möglichkeit besteht in einer dezimierten Version, die für

_

http://www.AvatR.net

¹¹ Bspw.: http://www.PayPal.com

AvatR lizenzkostenfrei ist und daher an eine Vielzahl an Nutzern ausgegeben werden kann. Die zweite Variante beinhaltet eine kurze Laufzeitversion in vollem Umfang.

2.3 Logistik

Der Vertrieb im Online-Store basiert auf der Zusammenarbeit mit der Firma Mediabit GmbH¹² aus Aschheim. Dieser Partner realisiert sowohl die Herstellung der DVDs und deren Verpackungen als auch die Lagerung und den Versand. Nicht nur das Produkt in Form einer DVD wird so in den Vertrieb integriert, sondern auch zusätzliche Komponenten wie Headsets, Mikrofone oder Webcams. Das bedeutet, dass die Logistik sowohl als Handelsverkauf zwischen Hersteller (Mediabit) und Markt (Saturn, Media Markt), als auch im Direktvertrieb zwischen dem Online-Store (AvatR) und dem Kunden realisiert wird.

Da es sich als wesentlich schwieriger darstellt das Produkt in einem Fachmarkt anzubieten als in einem Online-Store, haben wir uns für diese Aufgabe mit der Phoenix Organization, Inc.¹³ zu einer Kooperation entschlossen. Über diese Kooperation ist sowohl das Angebot im Markt selbst als auch die Verbesserung des Angebotes über einen Promoter im Markt realisierbar. Auf diesem Weg wird zudem überhaupt die Möglichkeit geschaffen sich in einem großen Markt zu positionieren.

Die Vorteile dieser Vertriebsstruktur sind:

- Kalkulierbare und geringere Kostenstruktur
- Nutzung des Know-How des Partners
- Beseitigung von Markteintrittsbarrieren

1

¹² http://www.mediabit.de/

¹³ http://www.phoenix-organization.de/