



US006654722B1

(12) **United States Patent**
Aldous et al.

(10) **Patent No.:** **US 6,654,722 B1**
(45) **Date of Patent:** **Nov. 25, 2003**

(54) **VOICE OVER IP PROTOCOL BASED SPEECH SYSTEM**

(75) Inventors: **Anne M. Aldous**, Davie, FL (US);
Joseph Celi, Jr., Boca Raton, FL (US);
Brett Gavagni, Sunrise, FL (US);
Kyriakos Leontiades, Boca Raton, FL (US);
Bruce D. Lucas, Yorktown Heights, NY (US); **David E. Reich**, Jupiter, FL (US)

(73) Assignee: **International Business Machines Corporation**, Armonk, NY (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 681 days.

(21) Appl. No.: **09/596,769**

(22) Filed: **Jun. 19, 2000**

(51) **Int. Cl.**⁷ **G10L 11/00**

(52) **U.S. Cl.** **704/270.1**

(58) **Field of Search** 704/270, 270.1,
704/275, 276; 370/237; 379/88.02, 221.01,
219; 709/204

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,581,600 A * 12/1996 Watts et al. 379/88.02

5,881,135 A * 3/1999 Watts et al. 379/88.02
5,916,302 A * 6/1999 Dunn et al. 709/204
5,933,490 A * 8/1999 White et al. 379/221.01
5,983,190 A * 11/1999 Trower et al. 704/276
6,014,437 A * 1/2000 Acker et al. 379/219
6,154,445 A * 11/2000 Farris et al. 370/237
6,269,336 B1 * 7/2001 Ladd et al. 704/270

* cited by examiner

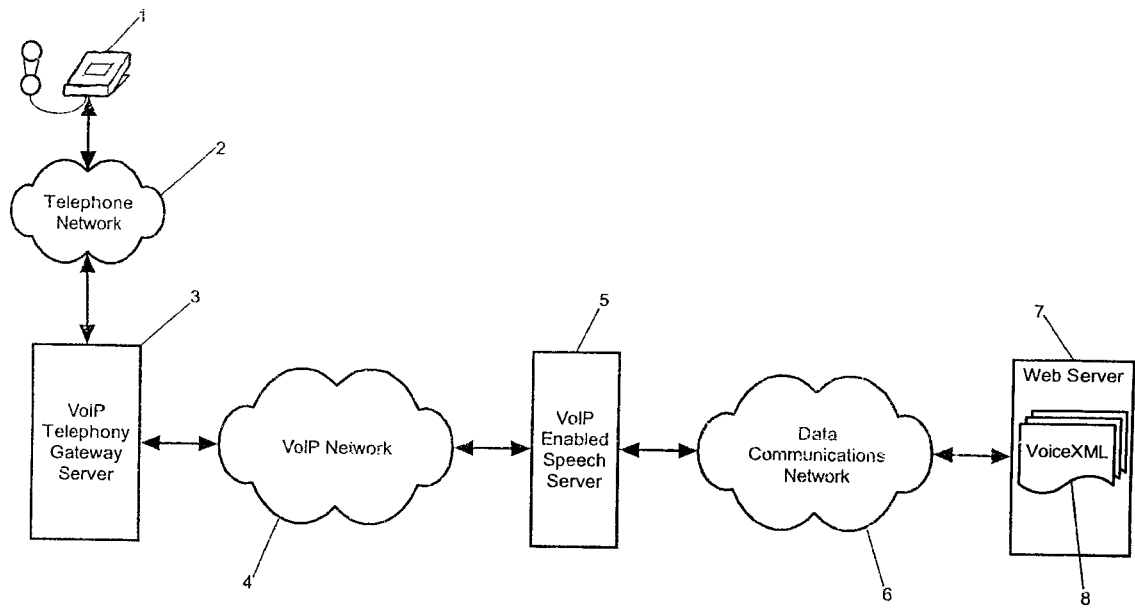
Primary Examiner—David D. Knepper

(74) *Attorney, Agent, or Firm*—Akerman Senterfitt

(57) **ABSTRACT**

A VoIP-enabled speech server can include a speech application which can be configured to communicate with a VoIP telephony gateway server over a VoIP communications path. The VoIP-enabled speech server can also include a VoIP-compliant call control interface to the VoIP telephony gate server, the VoIP-compliant call control interface establishing the VoIP communications path. In operation, the speech application can receive VoIP-compliant packets from the VoIP telephony gateway server over the VoIP communications path. Subsequently, digitized audio data can be reconstructed from the VoIP-compliant packets, and the digitized audio data can be speech-to-text converted. Additionally, text can be synthesized into digitized audio data and the digitized audio data can be encapsulated in VoIP-compliant packets which can be transmitted over the VoIP communications path to the telephony gateway server.

21 Claims, 3 Drawing Sheets



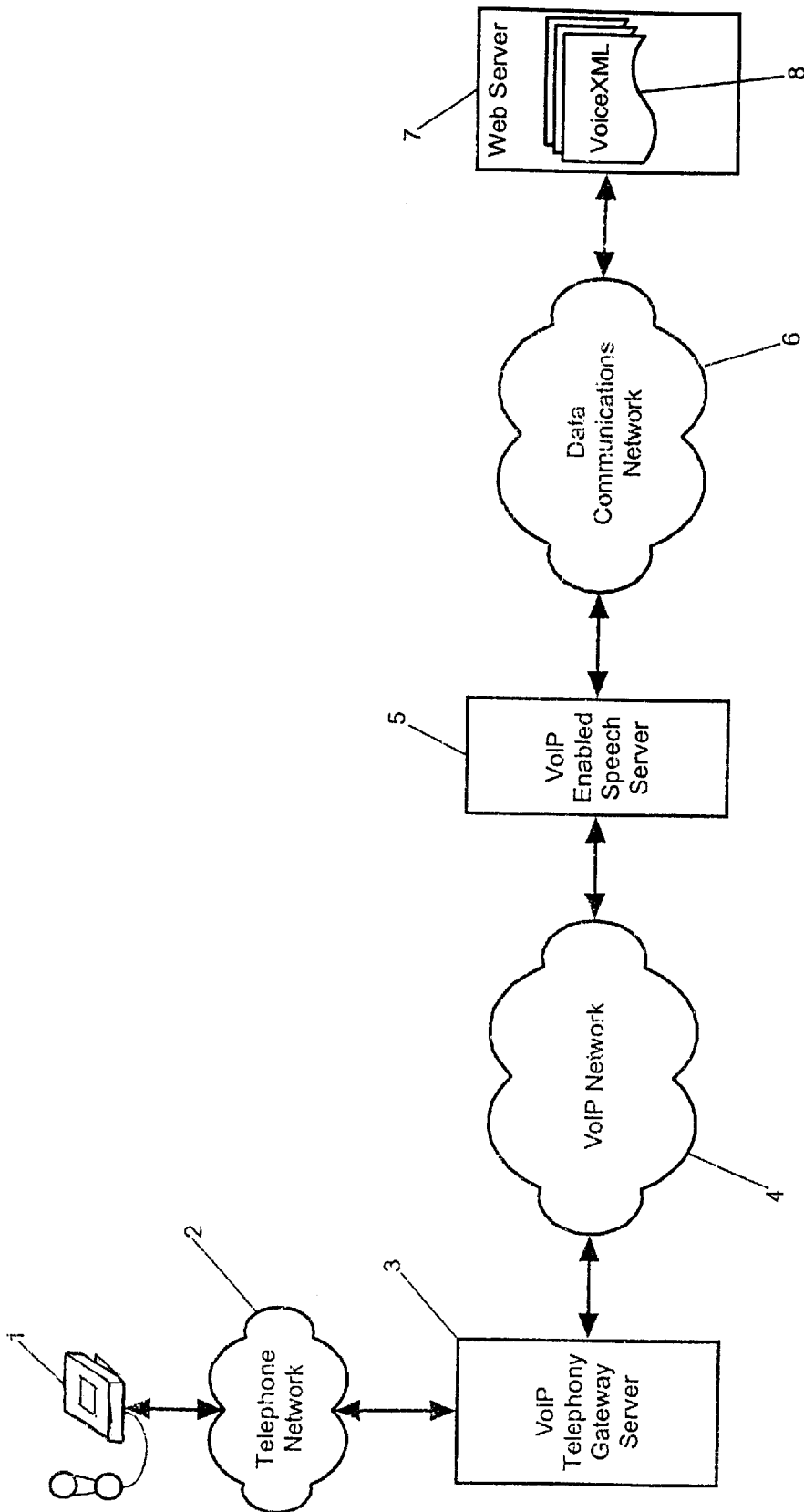


FIG. 1

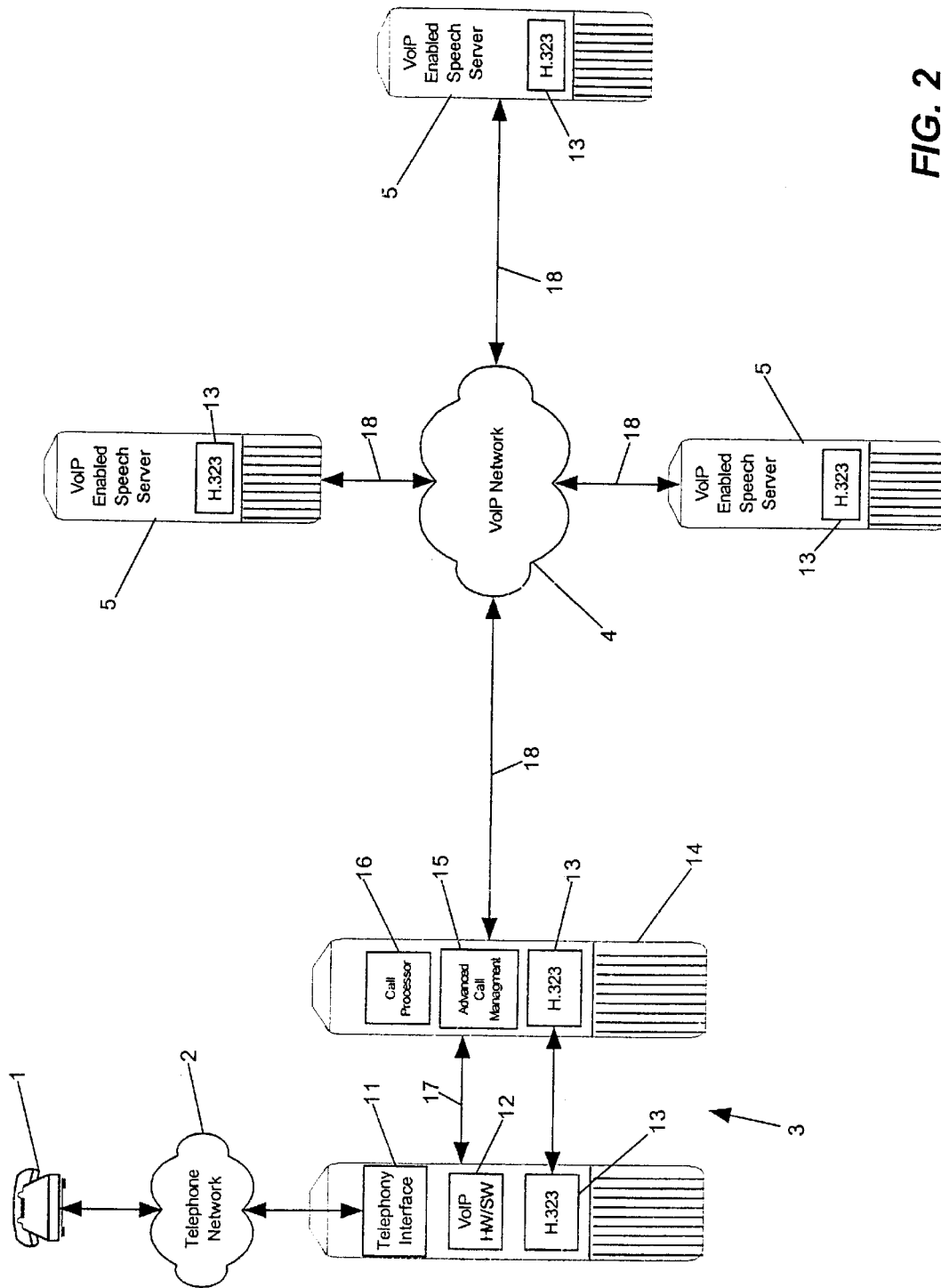


FIG. 2

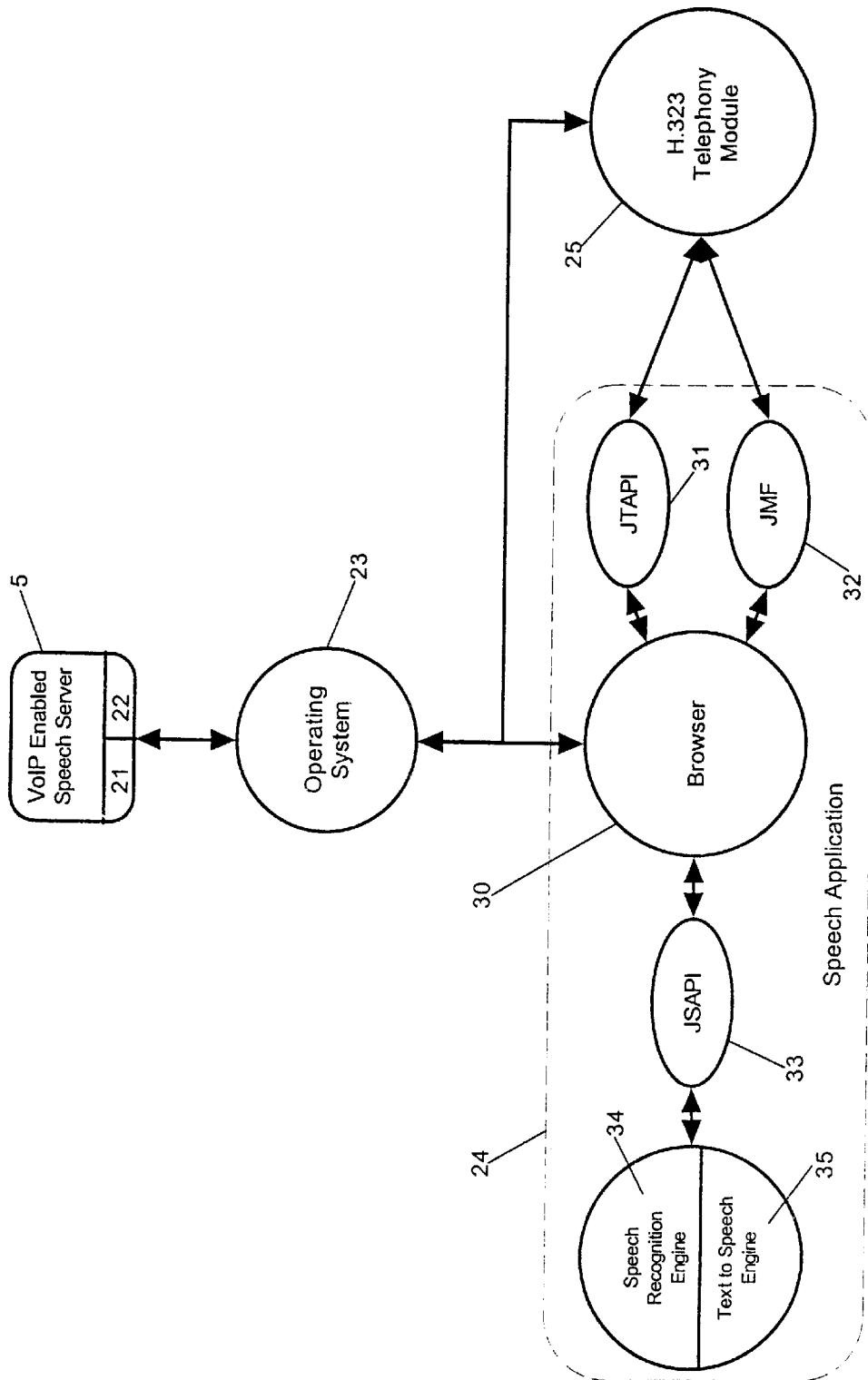


FIG. 3

VOICE OVER IP PROTOCOL BASED SPEECH SYSTEM

BACKGROUND OF THE INVENTION

1. Technical Field

This invention relates to the field of voice recognition and more particularly to a speech application for use in a Voice over IP protocol network.

2. Description of the Related Art

LAN telephony, which means “the integration of telephony and data services provided by packet-switched data networks,” is the technology that takes person-to-person communication to a high new level and associated costs to a lower level. LAN telephony enables a more flexible and cost-efficient use of many applications, for example automated call distribution, interactive voice response, voice logging, etc. This is in contrast to the relatively limited integration offered by the current voice/data integration paradigm, computer-telephony integration in which voice traffic is kept separate from data traffic and carried over circuit-switched links. Whereas the old paradigm for integrating data and voice has been to use the circuit-switched telephony fabric for data communications, the obvious drawbacks of the relatively low bandwidth available to data traffic, the inefficiency of circuit-switched data communications due to the “bursty” nature of data traffic, and the limited voice/data integration possibilities have led to present topologies in which IP data servers are bundled with proprietary PBXs or voice circuit switches in order to provide a loose integration between circuit and packet-switched networks and voice is carried by the circuit-switched network.

One of the most common uses of LAN telephony is in the enterprise Internet/Intranet environment, referred to as IP telephony. The Voice over IP (“VoIP”) protocol is the protocol upon which voice traffic can be transmitted across IP networks. In a VoIP network, analog speech signals received from an analog speech audio source, for example a PSTN or a microphone, are digitized, compressed and translated into IP packets for transmission over an IP network. Several well-known protocols implement the VoIP protocol specification including H.323, Session Initialization Protocol (“SIP”) and Master Gateway Control Protocol (“MGCP”).

A common application for IP telephony is the integration of voice mail (“v-mail”) and electronic mail (“e-mail”). Another application can include voice logging by financial or emergency-response organizations. Additionally, automated call distribution (“ACD”) can be facilitated whereby an ACD server performs value-based queuing of incoming telephone calls. Finally, interactive voice response systems can incorporate IP telephony in which responses are pre-programmed in a server as a workflow component. Still, speech recognition and speech synthesis applications (“speech applications”) have lagged in the use of IP telephony.

In particular, speech applications operate on real-time audio signals which cannot tolerate latencies associated with traditional data communications. As such, where speech applications have been incorporated in an IP telephony topology, the speech applications have been closely integrated with IP telephony server in order to preclude a negative impact from network based latencies. Accordingly, the design and development of such IP telephony enabled speech applications have been closely linked to the proprietary nature of the IP telephony server.

The tight linkage between the speech application and the IP telephony server substantially limits both the design and the extensibility of the speech application. Specifically, in the present paradigm the speech application design must incorporate functionality directly related to the chosen protocol for transporting packetized voice data to a speech recognition system and from a speech synthesis system. The development of a superior voice transport protocol, by nature of the tight linkage between the IP telephony server and the speech application, can compel the redesign of the speech application. Accordingly, there exists a need for a speech VoIP-based speech system in which the design and implementation of the speech application remains separate from the design and implementation of the IP telephony system.

SUMMARY OF THE INVENTION

It is an object of the present invention to provide a VoIP-based speech system in which the design and implementation of the speech application remains separate from the design and implementation of the IP telephony system. It is a further object of the present invention to provide a VoIP-enabled speech server which can receive audio input from the IP telephony system over a VoIP network. It is yet another object of the present invention to provide a method for coupling a speech application to a telephony gateway server in a VoIP network. Finally, it is an object of the present invention to provide each of the VoIP-based speech system, the VoIP-enabled speech server and the method for coupling the speech application to the telephony gateway server using standards-based interfaces to the VoIP network, the server and the speech application.

These and other objects of the present invention are accomplished in a VoIP-based speech system including: a VoIP telephony gateway server; at least one speech server, each speech server containing a VoIP-enabled speech application; a VoIP-compliant call control interface between the VoIP telephony gateway server and the speech server; and, a VoIP communications path between the VoIP telephony gateway-server and the speech application in the at least one speech server. In the VoIP-based speech system, the VoIP telephony gateway server and the speech application can establish the VoIP communications path through the VoIP-compliant call control interface.

In operation, the VoIP telephony gateway server can receive audio signals from a telephony interface, digitize the audio signals into digitized audio data, compress the digitized audio data into VoIP-compliant packets, and transmit the VoIP-compliant packets to the speech application in the at least one speech server through the VoIP communications path using the VoIP protocol. Correspondingly, the speech application can receive the VoIP-compliant packets, reconstruct the digitized audio data from the VoIP-compliant packets, and speech-to-text converting the digitized audio data. In addition, the speech application can synthesize text into digitized audio data, encapsulate the digitized audio data in VoIP-compliant packets and transmit the VoIP-compliant packets through the VoIP communications path to the VoIP telephony gateway server. Subsequently, the VoIP telephony-gateway server can receive the VoIP-compliant packets, reconstruct the digitized audio data from the VoIP-compliant packets, and transmit the digitized audio data through the telephony interface.

In one aspect of the present invention, the VoIP telephony server can include a telephony interface and a VoIP Gatekeeper. The VoIP Gatekeeper can receive a voice call

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.