US 20030182132A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2003/0182132 A1**

Niemoeller (43) **Pub. Date: Sep. 25, 2003**

(54) **VOICE-CONTROLLED ARRANGEMENT AND METHOD FOR VOICE DATA ENTRY AND VOICE RECOGNITION**

(76) Inventor: **Meinrad Niemoeller**, Holzkirchen (DE)

Correspondence Address:
BELL, BOYD & LLOYD, LLC
P. O. BOX 1135
CHICAGO, IL 60690-1135 (US)

(21) Appl. No.: **10/363,121**

(22) PCT Filed: **Aug. 16, 2001**

(86) PCT No.: **PCT/EP01/09475**

(30) **Foreign Application Priority Data**

Aug. 31, 2000 (EP) 00118895.2

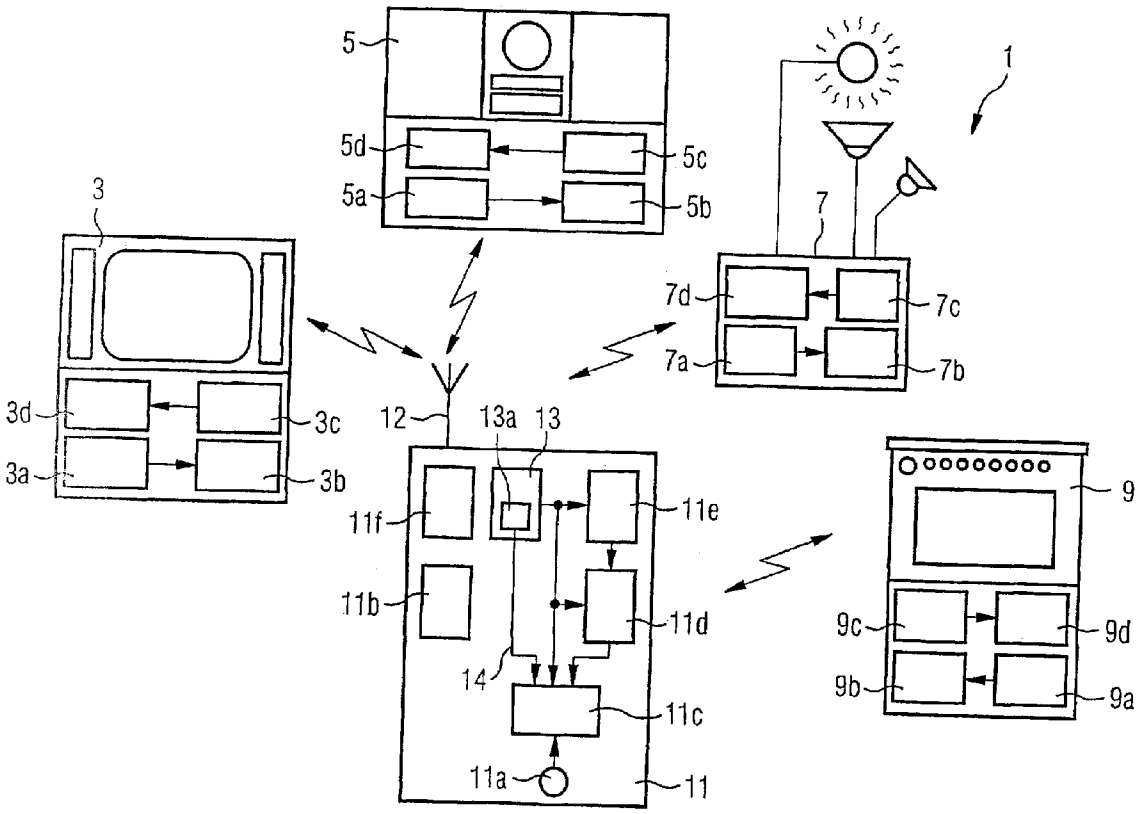
Publication Classification

(51) **Int. Cl.**[7] **G10L 11/00**

(52) **U.S. Cl.** **704/275**

(57) **ABSTRACT**

The invention relates to a voice-controlled arrangement (**1**) comprising a plurality of devices to be controlled (**3** to **9**) and a mobile voice data entry unit (**11**) which is connected to said devices by a wireless communication link. At least some of the devices each have a device vocabulary memory (**3***a* to **9***a*) and a vocabulary transmission unit (**3***b* to **9***b*), and the voice data entry unit has selection means for selecting the vocabularies to he loaded according to the route destination.



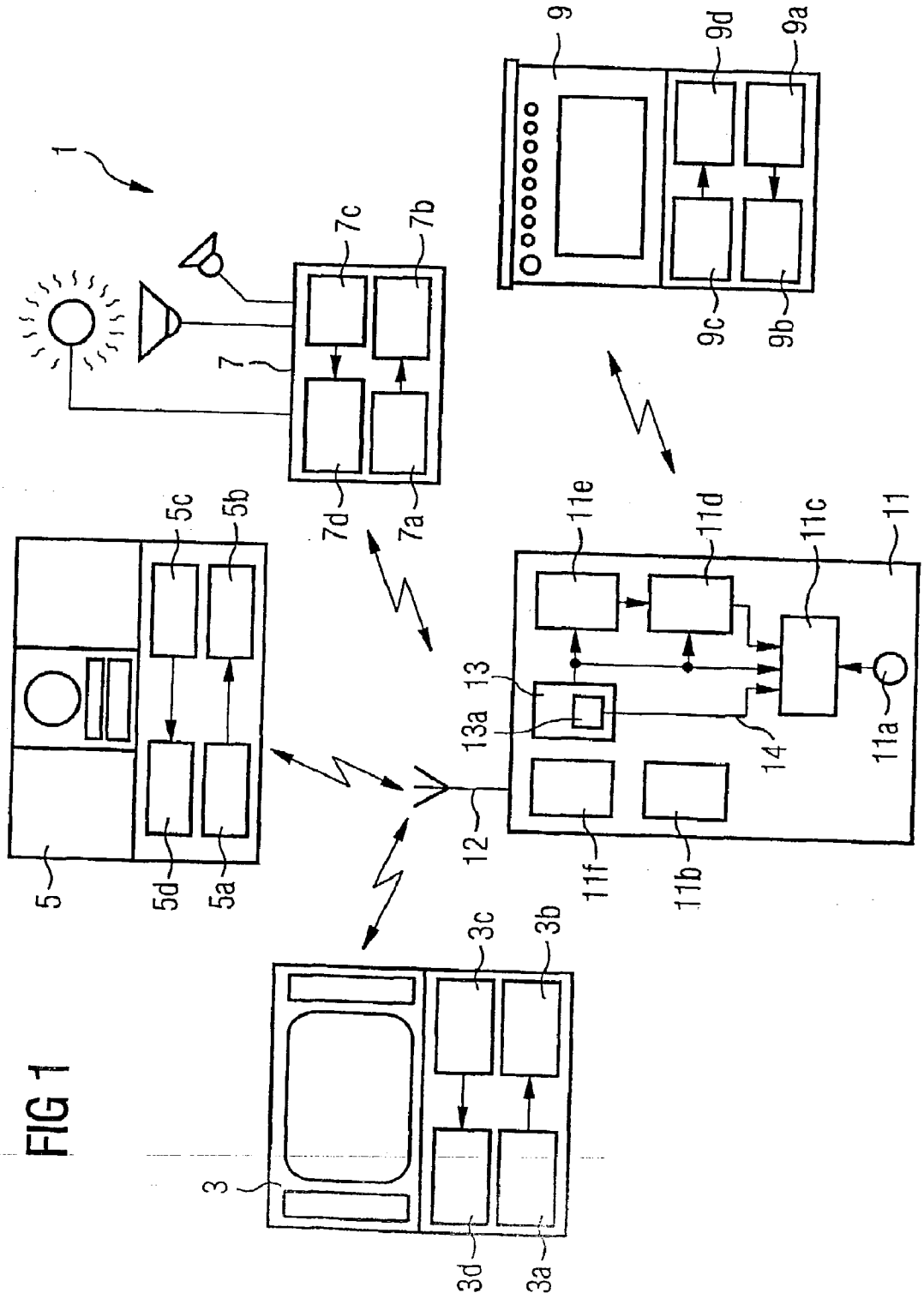
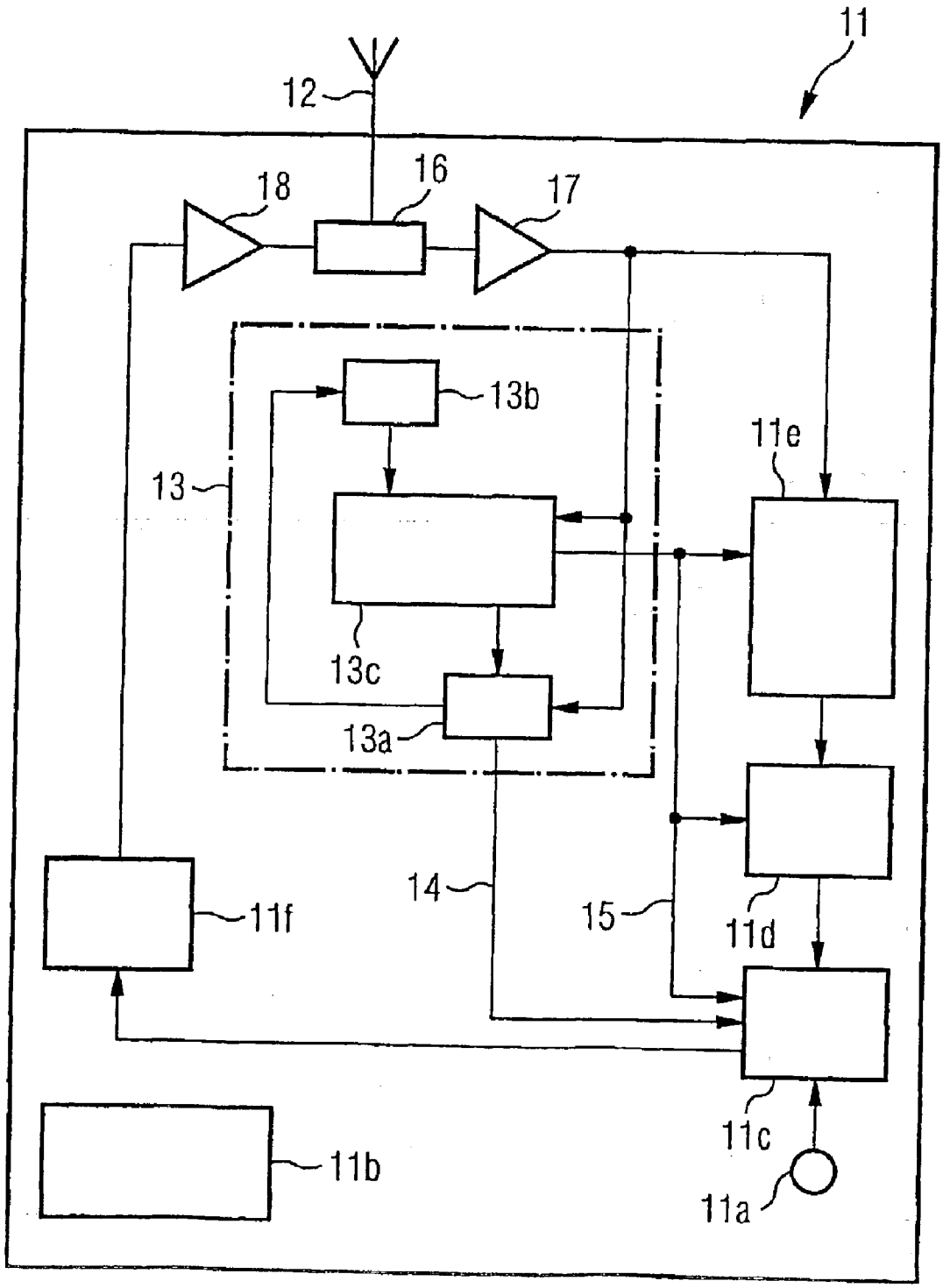


FIG 1

FIG 2



VOICE-CONTROLLED ARRANGEMENT AND METHOD FOR VOICE DATA ENTRY AND VOICE RECOGNITION

[0001] The invention relates to a voice-controlled arrangement comprising a plurality of devices according to the preamble of claim 1, and to a method for inputting and recognizing a voice, which can be applied in such an arrangement.

[0002] Since voice recognition systems have increasingly developed into a standard component in powerful computers for professional and private use, including PCs and Notebooks in the medium and lower price ranges, more and more work is being carried out on the possibilities of applying such systems in devices which are used in everyday life. Electronic devices such as mobile phones, cordless phones, PDAs and remote controls for audio systems and video systems etc. usually have an input keypad which comprises at least one numerical input array and a series of functional keys.

[0003] Some of these devices—in particular of course the various kinds of telephones, but also increasingly remote controls and other devices—are increasingly equipped with microphones and possibly also headphones for inputting and outputting voice. Devices of this type (for example some types of mobile phones) in which a simple voice recognition procedure is implemented for control functions on the device itself are already known. One example of this is the voice-controlled setting up of links by a voice input of a name into a mobile phone, said name being stored in an electronic telephone directory of the telephone. Furthermore, primitive to simple voice controls are also known for other devices which are used in everyday life, for example in remote controls for audio systems or lighting systems. All known devices of this type each have a separate dedicated voice recognition system.

[0004] It is possible to envisage a development which will entail an increasing number of technical devices and systems from everyday life, in particular in the domestic sphere and in motor vehicles, being equipped with their own respective voice recognition systems. As such systems are relatively complex in terms of hardware and software, and thus expensive, if they are to provide an acceptable level of operator convenience and sufficient recognition reliability, this development is a fundamental factor which drives costs higher and is thus welcomed by consumers only to a limited degree. For this reason, the primary goal is to reduce the expenditure on hardware and software further in order to be able to make available the most cost-effective solutions possible.

[0005] Arrangements have already been proposed in which a plurality of technical devices are assigned an individual voice input unit via which various functions of these devices are controlled by voice control. The control information is preferably transmitted here in a wire-free fashion to terminals (fixed or even mobile). However, the technical problem arises here that the voice input unit has to store a very large vocabulary for the voice recognition in order to be able to control various terminals. However, handling a large vocabulary involves adverse effects on the speed and precision of the recognition processes. In addition, such an arrangement has the disadvantage that it is not readily possible to make later updates with additional devices, which may not have been envisaged when the voice input unit was implemented. Last but not least, such a solution is still always very expensive, in particular due to the high memory requirements owing to the very large vocabulary.

[0006] In a German patent application which was not published before the priority date and which originates from the applicant, a voice-controlled arrangement comprising a plurality of devices to be controlled and a mobile voice input unit which is connected to the devices via an, in particular, wire-free telecommunications link is disclosed in which a device-specific vocabulary, but no processing means for the voice recognition, are respectively provided in the individual devices of the arrangement. On the other hand, the processing components of a voice recognition system are implemented in the voice input unit (in addition to the voice input means).

[0007] At least some of the devices each have a device vocabulary memory for storing a device-specific vocabulary and a vocabulary transmission unit for transmitting the stored vocabulary to the voice input unit. In contrast, the voice input unit comprises a vocabulary reception unit for receiving the vocabulary transmitted by a device or the vocabularies transmitted by devices. If the voice input unit is placed in the spatial vicinity of one or more devices, so that a telecommunications link is set up between the voice input unit and devices, the devices transmit their vocabularies to the voice input unit which buffers them. As soon as the telecommunications link between one or more devices and the voice input unit is broken, for example if the spatial distance becomes too large, the voice input unit can reject one or more buffered vocabularies again. The voice input unit accordingly administers the vocabularies of the terminals in a dynamic fashion.

[0008] The advantage of this arrangement is principally the fact that means with a relatively small storage capacity are sufficient to store the vocabularies in the voice input unit as, owing to the spatial separation of the vocabularies from the actual voice recognition capacity, the vocabularies do not need to be continuously stored in the voice input unit. This also increases the recognition rate in the voice input unit as fewer vocabularies are to be processed. However, when there is a plurality of spatially closely adjacent devices, in particular if their transmission ranges overlap, the voice input unit may nevertheless have to store and process a large number of vocabularies or may not be able to serve all the terminals given a limited storage capacity. Particularly the latter case is inconvenient for a user as he has no influence on which vocabularies are loaded into the voice input unit by terminals and which are rejected. Even if the transmission ranges of the terminals are comparatively small—for example have diameters of only a few meters—it is possible, particularly given a concentration of a large number of different terminals in a small space as in the domestic sphere or in an office, for the user to be able to carry out voice control on only some of these terminals owing to the abovementioned problems.

[0009] The invention is therefore based on the object of proposing an arrangement of this type which in particular avoids the abovementioned problems and especially develops the selection of the terminals to be controlled by voice.

The arrangement is also intended to be distinguished by low costs and an efficient method for inputting and recognizing voice.

[0010] This object is achieved by means of an arrangement having the features of patent claim 1 and by means of a method having the features of patent claim 13.

[0011] The invention develops the voice-controlled arrangement mentioned at the beginning having a plurality of devices and a mobile voice input unit connected to the devices via a wire-free telecommunications link in particular by virtue of the fact that selection means for selecting vocabularies to be loaded into the voice input unit are provided in the voice input unit. For this purpose, the selection means evaluate a directional information item of received signals which have been transmitted by the devices. The principle applied here originates from human communication: one person communicates with another by directing his attention at the person. Conversations in the surroundings of the two communicating people are "blanked out". Other people to whom the communicating people do not direct their attention therefore also feel that they are not being addressed.

[0012] The invention ensures that only specific vocabularies are loaded by devices which have been selected by the selection means. As a result, the recognition rate is significantly improved with spatially closely adjacent terminals as, owing to the directionally dependent selection, fewer vocabularies are loaded into the voice input unit, and therefore fewer vocabularies have to be processed. For example, radio or else infrared transmission links are possible as wire-free transmission methods between the devices and the voice input unit.

[0013] The selection means preferably comprise a detector, in particular an antenna, with a directional characteristic. The directionally dependent selection takes place by orienting the detector with the devices to be controlled as the level of a received signal of a device changes with the orientation of the detector with respect to a device transmitting the signal. In the case of an infrared transmission link, the selection means comprise an infrared detector which has a limited detection range, for example by virtue of a lens placed in front of it, so that infrared signals outside the detection range do not cause a corresponding vocabulary to be loaded.

[0014] In order to be able to evaluate the level of received signals, the voice input unit preferably has a level evaluation and control device. The latter determines the level of at least one received signal and controls, as a function thereof, the loading of a vocabulary into the vocabulary buffer or buffers by means of the vocabulary reception unit, said vocabulary being transmitted by means of the signal. The level evaluation and control device is preferably designed in such a way that it does not load a vocabulary transmitted by a received signal until a specific level is exceeded.

[0015] In one preferred embodiment, a plurality of vocabularies of devices are loaded simultaneously into the voice input unit. The level evaluation and control device is expediently constructed in this embodiment in such a way that the vocabulary of a further device is loaded into the voice input unit and replaces a vocabulary loaded there as soon as the received signal of the further device exceeds a predefined level and/or the levels of the signal which transmits the vocabulary to be replaced and/or is assigned to it. A plurality of vocabularies are thus stored in the voice input unit so that even a corresponding multiplicity of devices can be controlled. However, this gives rise to a corresponding need for storage in the voice input unit.

[0016] In one development, precisely one vocabulary of a device, which is replaced by the vocabulary of another device, can then be loaded into the voice input unit as soon as a received signal of the other device exceeds a predefined level and/or the level of the signal which transmits the vocabulary to be replaced and/or is assigned thereto. Therefore, as soon as the voice input unit is directed to another device so that its transmitted signal fulfils the criteria for loading into the voice input unit, the vocabulary which has already been loaded is replaced. The advantage of this embodiment is in particular the low storage requirement in the voice input unit as only one vocabulary is ever loaded.

[0017] In the preceding embodiment, the level evaluation and control device is expediently also designed to allocate different priorities to the vocabularies loaded into the voice input unit. If a new vocabulary is loaded, the vocabulary to be replaced can be determined by reference to the priorities. A vocabulary to be loaded will usually replace the loaded vocabulary with the lowest priority. The priorities can be allocated as a function of various criteria such as for example prioritization of the devices, the frequency of control of the devices, the time for which the vocabularies remain in the voice input unit, etc. The prioritization will appropriately be allocated as a function of the frequency with which the devices are controlled, i.e. devices which are controlled very often have a higher priority than devices which, in comparison, are controlled rarely. However, the assignment of priorities preferably takes place as a function of the conditions of the levels of the signals which transmit the vocabularies and/or are assigned to them. A relatively high level brings about a higher priority than a relatively low level here.

[0018] In one particularly preferred embodiment, the level evaluation and control device generates at least one control signal which can control or influence the recognition function of the voice recognition stage, specifically as a function of the evaluated level of a received signal. The influencing or control is advantageously carried out by raising or lowering the probabilities of the occurrence of a word or a plurality of words and/or the probabilities of a boundary between words of a vocabulary which is in particular proportional to the level.

[0019] By influencing the probabilities during recognition, use is made of the fact that a plurality of terminals have the same instructions and, when such an instruction is input, the probability is used to decide which device is to be controlled. In other words, various devices can be controlled with identical instructions, which of the devices is addressed being determined by the user by the orientation of the voice input unit.

[0020] The communication between the voice input unit and the devices preferably takes place according to the Bluetooth standard. For this purpose, the vocabulary transmission unit or vocabulary transmission units and vocabulary reception unit are embodied as a radio transceiver unit according to the Bluetooth standard. The Bluetooth standard

DOCKET ALARM

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.