



(12) **United States Patent**
Bennett

(10) **Patent No.:** **US 7,203,646 B2**
(45) **Date of Patent:** **Apr. 10, 2007**

(54) **DISTRIBUTED INTERNET BASED SPEECH RECOGNITION SYSTEM WITH NATURAL LANGUAGE SUPPORT**

5,068,789 A 11/1991 van Vliembergen
5,146,405 A 9/1992 Church
5,157,727 A 10/1992 Schloss
5,231,670 A 7/1993 Goldhor et al.
5,278,980 A * 1/1994 Pedersen et al. 707/4
5,293,584 A 3/1994 Brown et al.

(75) Inventor: **Ian M. Bennett**, Palo Alto, CA (US)
(73) Assignee: **Phoenix Solutions, Inc.**, Palo Alto, CA (US)
(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(Continued)

FOREIGN PATENT DOCUMENTS

(21) Appl. No.: **11/419,736** EP 1094388 4/2001

(22) Filed: **May 22, 2006**

(Continued)

(65) **Prior Publication Data**
US 2006/0200353 A1 Sep. 7, 2006

OTHER PUBLICATIONS

Related U.S. Application Data

Coffman, Daniel et la., Provisional Application for Patent, U.S. Appl. No. 60/117,595, filed Jan. 27, 1999, 111 pages.

(63) Continuation of application No. 09/439,174, filed on Nov. 12, 1999, now Pat. No. 7,050,977.

(Continued)

(51) **Int. Cl.**
G10L 15/18 (2006.01)
G06F 17/20 (2006.01)

Primary Examiner—Martin Lerner
(74) *Attorney, Agent, or Firm*—J. Nicholas Gross

(52) **U.S. Cl.** **704/257; 704/270.1; 707/5**
(58) **Field of Classification Search** **704/251, 704/252, 255, 257, 270, 270.1, 275; 707/3, 707/4, 5**

(57) **ABSTRACT**

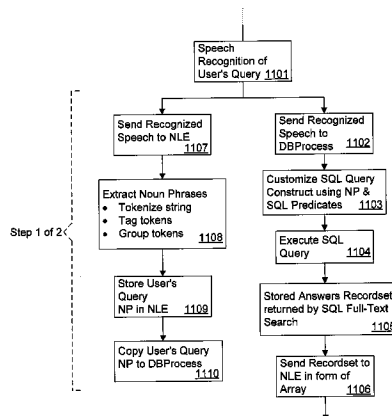
See application file for complete search history.

A speech-enabled internet based computing system includes a configurable speech recognition engine used for interacting with content on a web accessible page. The speech recognition engine is distributed across a client and server architecture, and is adaptive so that speech processing operations can be allocated as needed between the two. This allows for support for client devices having differing computing capabilities. Natural language operations can also be supported as desired. A user can thus interact with a web page and select items of interest using speech as a mode of input. Dynamic grammars can assist in the recognition operations to improve speed and comprehension.

(56) **References Cited**
U.S. PATENT DOCUMENTS

10 Claims, 31 Drawing Sheets

- 4,473,904 A 9/1984 Suehiro et al.
- 4,587,670 A 5/1986 Levinson et al.
- 4,783,803 A 11/1988 Baker et al.
- 4,785,408 A 11/1988 Britton et al.
- 4,852,170 A 7/1989 Bordeaux
- 4,914,590 A 4/1990 Loatman et al.
- 4,991,094 A 2/1991 Fagan et al.
- 4,991,217 A 2/1991 Garrett et al.



U.S. PATENT DOCUMENTS

5,371,901	A	12/1994	Reed et al.	
5,384,892	A	1/1995	Strong	
5,454,106	A *	9/1995	Burns et al.	707/4
5,475,792	A	12/1995	Stanford et al.	
5,509,104	A	4/1996	Lee et al.	
5,513,298	A	4/1996	Stanford et al.	
5,553,119	A	9/1996	McAlliser et al.	
5,602,963	A	2/1997	Bissonnette et al.	
5,625,814	A *	4/1997	Luciw	707/5
5,652,897	A	7/1997	Linebarger et al.	
5,668,854	A	9/1997	Minakami et al.	
5,675,707	A	10/1997	Gorin et al.	
5,680,511	A	10/1997	Baker et al.	
5,680,628	A	10/1997	Carus et al.	
5,694,592	A *	12/1997	Driscoll	707/3
5,727,950	A	3/1998	Cook et al.	
5,758,322	A	5/1998	Rongley	
5,802,526	A	9/1998	Fawcett et al.	
5,819,220	A	10/1998	Sarukkai et al.	
5,836,771	A	11/1998	Ho et al.	
5,860,063	A	1/1999	Gorin et al.	
5,867,817	A	2/1999	Catallo et al.	
5,873,062	A	2/1999	Hansen et al.	
5,884,302	A	3/1999	Ho	
5,915,236	A	6/1999	Gould et al.	
5,934,910	A	8/1999	Ho et al.	
5,956,683	A	9/1999	Jacobs et al.	
5,960,394	A	9/1999	Gould et al.	
5,960,399	A	9/1999	Barclay et al.	
5,978,756	A	11/1999	Walker et al.	
5,987,410	A	11/1999	Kellner et al.	
5,995,918	A	11/1999	Kendall et al.	
5,995,928	A	11/1999	Nguyen et al.	
6,009,387	A	12/1999	Ramaswamy et al.	
6,023,697	A *	2/2000	Bates et al.	707/4
6,029,124	A	2/2000	Gillick et al.	
6,032,111	A	2/2000	Mohri	
6,035,275	A	3/2000	Brode et al.	
6,044,266	A	3/2000	Kato	
6,044,337	A	3/2000	Gorin et al.	
6,078,914	A *	6/2000	Redfern	707/3
6,081,774	A *	6/2000	de Hita et al.	704/9
6,088,692	A *	7/2000	Driscoll	707/5
6,101,472	A	8/2000	Giangarra et al.	
6,105,023	A *	8/2000	Callan	707/5
6,112,176	A	8/2000	Goldenthal et al.	
6,119,087	A	9/2000	Kuhn et al.	
6,125,284	A	9/2000	Moore et al.	
6,125,341	A	9/2000	Raud et al.	
6,138,089	A	10/2000	Guberman	
6,141,640	A	10/2000	Moo	
6,144,848	A	11/2000	Walsh et al.	
6,144,938	A	11/2000	Surace et al.	
6,173,279	B1 *	1/2001	Levin et al.	707/5
6,178,404	B1	1/2001	Hambleton et al.	
6,182,038	B1	1/2001	Balakrishnan et al.	
6,182,068	B1	1/2001	Culliss	
6,185,535	B1	2/2001	Hedin et al.	
6,192,110	B1	2/2001	Abella et al.	
6,195,636	B1	2/2001	Crupi et al.	
6,226,610	B1	5/2001	Keiller et al.	
6,233,559	B1	5/2001	Balakrishnan	
6,243,679	B1	6/2001	Mohri et al.	
6,246,986	B1	6/2001	Ammicht et al.	
6,246,989	B1	6/2001	Polcyn	
6,256,607	B1	7/2001	Digalakis et al.	
6,269,336	B1	7/2001	Ladd et al.	
6,278,973	B1	8/2001	Chung et al.	
6,292,767	B1	9/2001	Jackson et al.	

6,327,568	B1	12/2001	Joost	
6,330,530	B1	12/2001	Horiguchi et al.	
6,363,349	B1	3/2002	Urs et al.	
6,374,219	B1	4/2002	Jiang	
6,374,226	B1	4/2002	Hunt et al.	
6,381,594	B1	4/2002	Eichstaedt et al.	
6,389,389	B1	5/2002	Meunier et al.	
6,408,272	B1	6/2002	White et al.	
6,411,926	B1	6/2002	Chang	
6,418,199	B1	7/2002	Perrone	
6,427,063	B1	7/2002	Cook et al.	
6,434,524	B1 *	8/2002	Weber	704/257
6,434,529	B1	8/2002	Walker et al.	
6,446,064	B1 *	9/2002	Livowsky	707/5
6,453,020	B1	9/2002	Hughes et al.	
6,499,011	B1	12/2002	Souvignier et al.	
6,499,013	B1	12/2002	Weber	
6,510,411	B1	1/2003	Norton et al.	
6,513,037	B1	1/2003	Ruber et al.	
6,522,725	B2	2/2003	Kato	
6,532,444	B1	3/2003	Weber	
6,539,359	B1	3/2003	Ladd et al.	
6,567,778	B1	5/2003	Chao Chang et al.	
6,574,597	B1	6/2003	Mohri et al.	
6,584,464	B1 *	6/2003	Warthen	707/4
6,594,269	B1	7/2003	Polcyn	
6,594,348	B1	7/2003	Bjurstrom et al.	
6,601,026	B2 *	7/2003	Appelt et al.	704/9
6,614,885	B2	9/2003	Polcyn	
6,618,726	B1	9/2003	Colbath et al.	
6,633,846	B1	10/2003	Bennett et al.	
6,681,206	B1	1/2004	Gorin et al.	
6,697,780	B1	2/2004	Beutnagel et al.	
6,742,021	B1	5/2004	Halverson et al.	
6,823,308	B2	11/2004	Keiller et al.	
6,862,713	B1 *	3/2005	Kraft et al.	715/728
6,871,179	B1	3/2005	Kist et al.	
6,901,366	B1	5/2005	Kuhn et al.	
6,922,733	B1	7/2005	Kuiken et al.	
6,940,953	B1	9/2005	Eberle et al.	
6,941,273	B1	9/2005	Loghmani et al.	
6,961,954	B1	11/2005	Maybury et al.	
6,964,012	B1	11/2005	Zirngibl et al.	
6,965,864	B1	11/2005	Thrift et al.	
6,965,890	B1	11/2005	Dey et al.	
7,058,573	B1	6/2006	Murveit et al.	
2001/0016813	A1	8/2001	Brown et al.	
2001/0032083	A1	10/2001	Van Cleven	
2001/0056346	A1	12/2001	Ueyama et al.	
2002/0032566	A1	3/2002	Tzirkel-Hancock et al.	
2002/0046023	A1	4/2002	Fujii et al.	
2002/0059068	A1	5/2002	Rose et al.	
2002/0059069	A1	5/2002	Hsu et al.	
2002/0086269	A1	7/2002	Shpiro	
2002/0087325	A1	7/2002	Lee et al.	
2002/0087655	A1	7/2002	Bridgman et al.	
2002/0091527	A1	7/2002	Shiau	
2003/0191625	A1	10/2003	Gorin et al.	
2005/0091056	A1	4/2005	Surface et al.	
2005/0131704	A1	6/2005	Dragosh et al.	

FOREIGN PATENT DOCUMENTS

EP	1096471	5/2001
WO	9811534	3/1998
WO	9948011	9/1999
WO	9950830	10/1999
WO	0014727	3/2000
WO	0017854	3/2000
WO	0020962	4/2000
WO	0021075	4/2000

WO	0030072	5/2000
WO	0030287	5/2000
WO	0068823	11/2000
WO	0116936	3/2001
WO	0118693	3/2001
WO	0126093	4/2001
WO	0178065	10/2001
WO	0195312	12/2001
WO	0203380	1/2002

OTHER PUBLICATIONS

Fourney, G.D., "The Viterbi Algorithm," Proc. IEEE, vol. 73, pp. 268-278, Mar. 1973.

Baker, J.H., "The dragon system—An Overview," IEEE Trans. on ASSP Proc., ASSP-23(1):Feb. 24-29, 1975.

Bennett, I., "A Study of Speech Compression Using Analog Time Domain Sampling techniques," A Dissertation Submitted to the Dept. Of Electrical Engineering and the Committee on Graduate Studies of Stanford University, May 1975, pp. 16-32; 76-111.

Rabiner, L.R., "Digital Processing of Speech Signals," Prentice Hall, 1978, pp. 116-171; 355-395.

Jelinek, F. et al., "Continuous Speech Recognition: Statistical methods" in Handbook of Statistics, II, P.R. Krishtnaiad, Ed. Amsterdam, The Netherlands, North-Holland, 1982.

Bahl, L.R. et al., "A maximum likelihood approach to continuous speech recognition," IEEE Trans. Pattern Anal. Mach. Intell., PAMI-5: 179-190, 1983.

Hudson, R.A., "Word Grammar," Blackman Inc., Cambridge, MA, 1984, pp. 1-14; 41-42; 76-90; 94-98; 106-109; 211-220.

Quirk, R. et al., "A Comprehensive Grammar of English Language", Longman, London and New York, 1985, pp. 245-331.

Makhoul, J. et al., "Vector Quantization in Speech Coding," Proceedings of the IEEE, vol. 73, No. 11, Nov. 1985, pp. 1551-1588.

Rabiner, L.R., "A Tutorial on Hidden Markov Models and Selected Applications ins Speech Recognition," Proc. IEEE, vol. 77, No. 2, Feb. 1989, pp. 257-286.

Gersho, A. et al., "Vector Quantization and Compression," Kluwer Academic Publishers, 1991, p. 309-340.

Rabiner, L.R. et al., "Fundamentals of Speech Recognition," Prentice Hall, 1993, pp. 11-68.

Morgan, N. et al., "Hybrid Neural Network/Hidden Markov Model Systems for Continuous Speech Recognition," Journal of Pattern Recognition and Artificial Intelligence, vol. 7, No. 4 pp. 899-916. (1993).

Lieberman, P., "Intonation, Perception and Language," Research Monograph No. 38, MIT Press, Cambridge, Mass., 1967, pp. 5-37.

Unisys Corp., "Natural Language Speech Assistant (NLSA) Capabilities Overview," NLR 3.0, Aug. 1998, Malvern, PA, 27 pages.

Baum, L.E. et al., "A Maximum Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains," The Annals of Mathematical Statistics, 1970, vol. 41, No. 1, pp. 164-171.

Arons, B., "The Design of Audio Servers and Toolkits for Supporting Speech in the User Interface," believed to be published in: Journal of the American Voice I/O Society, pp. 27-41, Mar. 1991.

Hazen, T et al., "Recent Improvements in an Approach to Segment-Based Automatic Language Identification," believed to be published in: Proceedings of the 1994 International Conference on Spoken Language Processing, Yokohama, Japan, pp. 1883-1886, Sep. 1994.

House, D., "Spoken-Language Access to Multimedia (SLAM): A Multimodal Interface to the World-Wide Web," Masters Thesis, Oregon Graduate Institute, Department of Computer Science & Engineering, 59 pages, Apr. 1995.

Julia, L. et al., "http://www.speech.sri.com/demos/ atis.html," believed to be published in: Proceedings AAAI'97: Stanford, pp. 72-76, Jul. 1997.

Lau, R. et al, "Webgalaxy-Integrating Spoken Language and Hypertext Navigation," believed to be published in: in Kokkinakis, G. et al., (Eds.) Eurospeech '97, Proceedings of the 5th European

Digalakis, V. et al., "Product-Code Vector Quantization of Cepstral Parameters for Speech Recognition over the WWW," believed to be published in: Proc. ICSLP '98, 4 pages. 1998.

Melin, H., "On Word Boundary Detection in Digit-Based Speaker Verification," believed to be published in: Workshop on Speaker Recognition and Its Commercial and Forensic Applications (RLA2C), Avignon, France, Apr. 20-23, pp. 46-49, 1998.

Ramaswamy, G. et al., "Compression of Acoustic Features for Speech Recognition in Network Environments," believed to be published in: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 977-980, Jun. 1998.

Lu, B. et al., "Scalability Issues in the Real Time Protocol (RTP)," Project Report for CPSC 663 (Real Time Systems), Dept. of Computer Science, Texas A & M University, 19 pages, 1999.

Giuliani, D. et al., "Training of HMM with Filtered Speech Material for Hands-Free Recognition," believed to be published in: Proceedings of ICASSP '99, Phoenix, USA, 4 pages, 1999.

Digilakis, V. et al., "Quantization of Cepstral Parameters for Speech Recognition over the World Wide Web," believed to be published in: IEEE Journal on Selected Areas of Communications, 22 pages, 1999.

Tsakalides, S. et al., "Efficient Speech Recognition Using Subvector Quantization and Discrete-Mixture HMMs," believed to be published in: Proc. ICASSP '99, 4 pages, 1999.

Lin, B. et al., "A Distributed Architecture for Cooperative Spoken Dialogue Agents with Coherent Dialogue State and History," believed to be published in: IEEE Automatic Speech Recognition and Understanding Workshop, Keystone, Colorado, USA, 4 pages, Dec. 1999.

Meunier, J., "RTP Payload Format for Distributed Speech Recognition," 48th IETF AVT WG—Aug. 3, 2000, 10 pages, 2000.

Sand Cherry Networks, SoftServer product literature, 2 pages, 2001.

Kim, H. et al., "A Bitstream-Based Front-End for Wireless Speech Recognition on IS-136 Communications System," IEEE Transactions on Speech and Audio Processing, vol. 9, No. 5, pp. 558-568, Jul. 2001. (11 pages).

Agarwal, R., Towards a PURE Spoken Dialogue System for Information Access, believed to be published in Proceedings of the ACL/EACL Workshop on Interactive Spoken Dialog Systems: Bringing Speech and NLP Together in Real Applications, Madrid, Spain, 1997, 9 pages.

Ammicht, Egbert et al., "Knowledge Collection for Natural Language Spoken Dialog Systems," believed to be published in Proc. Eurospeech, vol. 3, p. 1375-1378, Budapest, Hungary, Sep. 1999, 4 pages.

AT&T Corp., "AT&T Watson Advanced Speech Applications Platform," 1996, 3 pages.

AT&T Corp., "AT&T Watson Advanced Speech Application Platform Version 2.0," 1996, 8 pages.

AT&T Corp., "AT&T Watson Advanced Speech Applications Platform Version 2.0," 1996, 3 pages.

Gorin, Allen, "Processing of Semantic Information in Fluently Spoken Language," believed to be published in Proc. ICSLP, Philadelphia, PA, Oct. 1996, 4 pages.

Gorin, Allen et al., "How May I Help You," believed to be published in Proc. IVTTA, Basking Ridge, NJ, Oct. 1996, 32 pages.

Mohri, Mehryar, "String Matching With Automata," Nordic Journal of Computing, 1997, 15 pages.

Prudential News, "Prudential Pilots Revolutionary New Speech-Based Telephone Customer Service System Developed by AT&T Labs—Company Business and Marketing," Dec. 6, 1999, 3 pages.

Riccardi, Giuseppe et al., "A spoken language system for automated call routing," believed to be published in Proc. ICASSP '97, 1997, 4 pages.

Sharp, Douglas, et al., "The Watson Speech Recognition Engine," accepted by ICASSP, 1997, 9 pages.

European Patent Office search report for EP Application No. 00977144, dated Mar. 30, 2005, 5 pages.

Burstein, A. et al. "Using Speech Recognition In A Personal Communications System," Proceedings of the International Con-

Kuhn, T. et al., "Hybrid In-Car Speech Recognition For Mobile Multimedia Applications," Vehicular Technology Conference, Houston, Texas, May 1999, pp. 2009-2013.

Travis, L., "Handbook of Speech Pathology", Appleton-Century-Crofts, Inc., 1957, pp. 91-124.

Baum, L.E., et al., "Statistical inference for probabilistic functions for finite state Markov chains," Ann. Math. Stat., 37: 1554-1563, 1966.

Flanagan, J.L., "Speech Analysis Synthesis and Perception", 2nd edition, Springer-Verlag Berlin, 1972, pp. 1-53.

Baum, L.E., "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," Inequalities 3: 1-8, 1972.

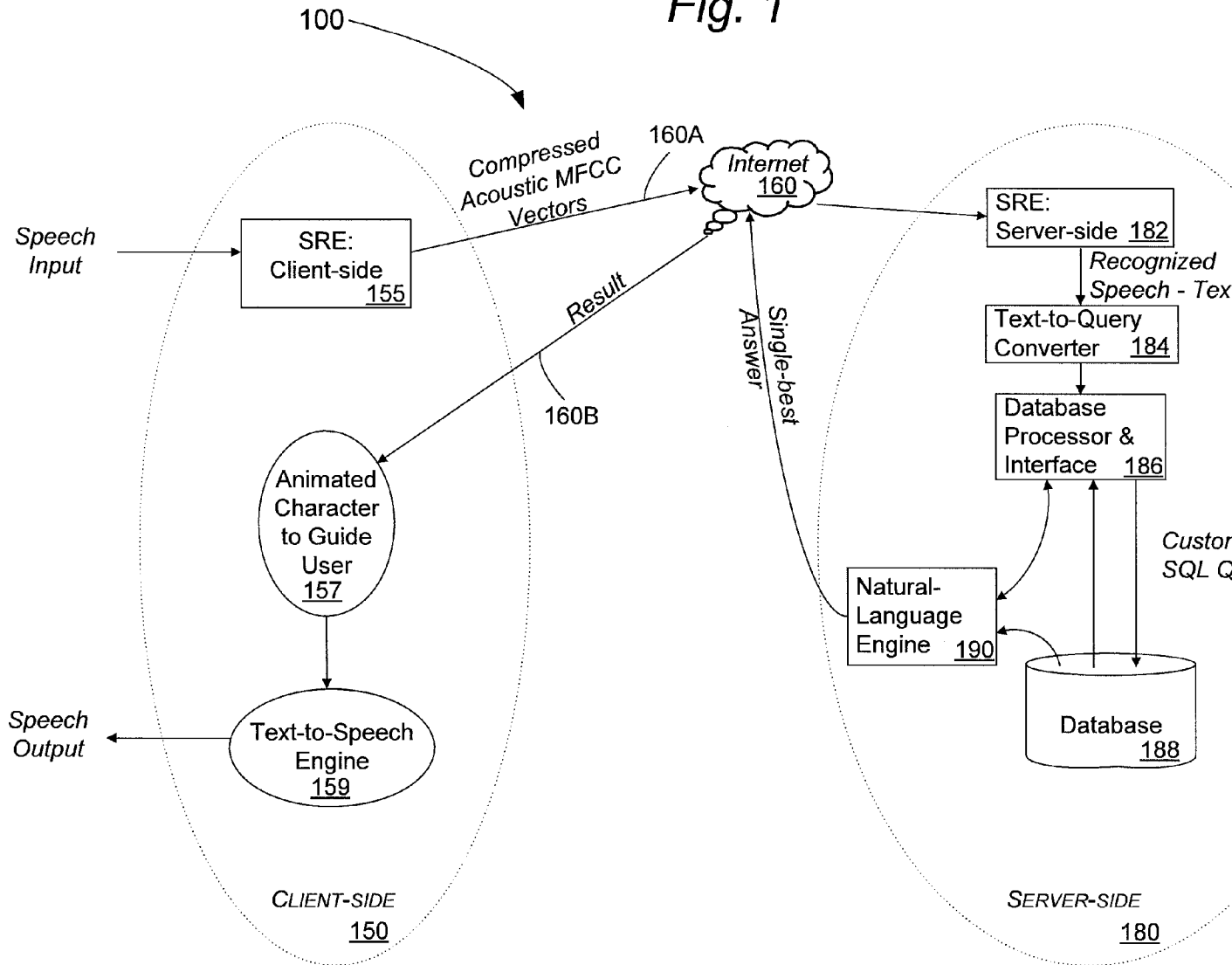
Cox, Richard V. et al., "Speech and Language Processing for Next-Millennium Communications Services," Proceedings of the IEEE, vol. 88, No. 8, Aug. 2000, pp. 1314-1337.

Kuhn, Roland, et al., "The Application of Semantic Classification Trees to Natural Language Understanding," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, No. 5, May 1995, pp. 449-460.

AT&T Corp., "Network Watson 1.0 System Overview," 1998, 4 pages.

* cited by examiner

Fig. 1



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.