

Artificial Intelligence for Human Resources

Creating and Coping
With Disruption Across
HR Functions at TCS

Girish Keshav Palshikar
Amol Khanapurkar

IN BRIEF

Uncertainties in the business environment present multiple challenges to HR managers. The good news is that HR can leverage technology to bring in agility and efficiency across functions. Data analytics can provide people insights, that are critical for large enterprises, while, machine learning can automate mundane tasks to free up time for HR managers to focus on strategic and value added work. The caveat is that this brings changes to HR's perception of itself and changes to its functions. It eliminates many well understood roles, creates new possibilities and challenges.

TCS Research has initiatives that support each HR function, from talent acquisition, through employee engagement, allocation, learning and development, to performance evaluation, attrition and retention. The research-based models underlying each tool are robust. They have been tested and validated with field data. Several of the tools are currently being used by TCS HR. These mature tools have become a part of CHROMA™—the TCS HR Management Platform.

Many top global companies place uncertainty within the business environment as their topmost risk. Volatility in the operating environment leads to unexpected changes in regulations, technology, market conditions, and competitors' strategies. These keep Human Resources (HR) continually in a fire-fighting mode. Mergers and acquisitions (M&As), abrupt changes in visa rules, phasing out of technology

(that was the buzz term two years ago), new and different types of competitors, big layoffs, or high volume hiring at short notice in a skills-short world shake up HR's best planned strategies and operations. HR faces its own existential problems—low employee to HR ratio, manual processes, automation vs people threat, discordance over the Bell Curve, and poor perception of HR as a function.

Fact File

TCS Research: HR Analytics

Outcomes: QUEST, HiSPEED, vTrain, HCV, RINX, TEAC, iWFM, iSupport

Principal Investigators: Girish K. Palshikar and Rajiv Srivasav

Techniques Used: Machine Learning, Natural Language Processing

Industries Benefited: All functions within HR Managements

Patents: 12 filed, 2 granted (US, EPO)

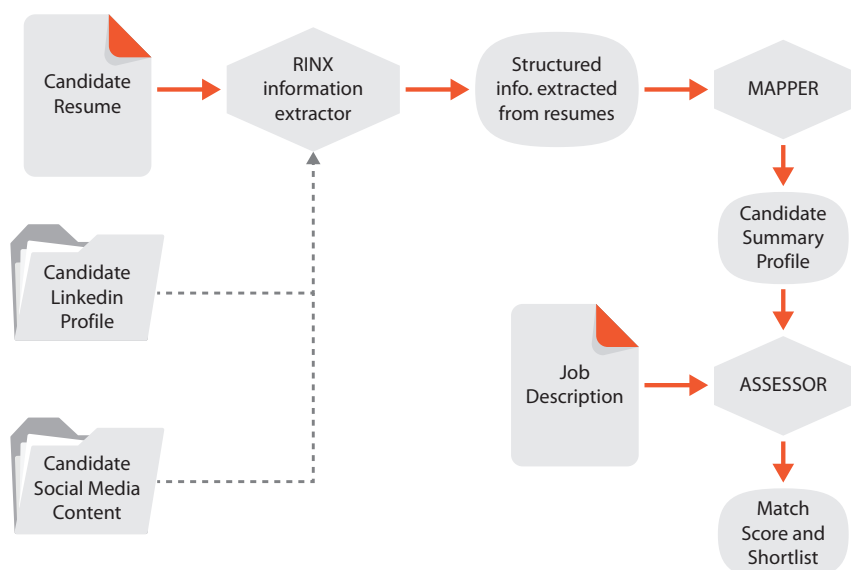
Papers: 4 journals, 3 book chapters, 27 conference presentations

How can technology help HR stay agile in the volatility, uncertainty, complexity, and ambiguity (VUCA) world and be a strategic business partner in large enterprises? The short answer: leverage artificial intelligent (AI).

How can technology help HR stay agile in the volatility, uncertainty, complexity, and ambiguity (VUCA) world and be a strategic business partner in large enterprises? The short answer: Leverage artificial intelligent (AI).

The HR function is increasingly using AI to automate drudge work, act on data-driven insights, minimize bad decisions, and save more time for strategic planning as well as personal interactions that can provide value for stakeholders.

A number of analytics-based tools are available across HR functions—from talent acquisition, employee engagement, allocation across a large set of job profiles, learning and development to performance evaluation, attrition and retention. The key is to have models that have been validated in the field and are robust; models that can be tweaked in relation to the changing business requirements. As every role takes in more automation, HR is forced to reinvent its roles and processes.



RINX Input: Unstructured resume documents, in format of doc, docx, pdf, rtf, or html

Figure 1: Overview of the TEAC system for automatic candidate shortlisting

TCS can be seen as an archetype of a large, global enterprise that pushes itself to innovate and stay ahead. By presenting several research interventions made for TCS' HR department, we hope readers can get a picture of the dramatic changes in HR's perception of its functions and the resulting changes to operations.

Talent acquisition

With the introduction of cloud-based platforms for automated testing and evaluation, the *operations* of talent acquisition takes a fraction of the time and effort it used to five years ago. This offers greater bandwidth for HR to use novel devices for better hiring: mining databases for top talent or gamify hiring. This is, in the bargain, creating a disruption of sorts. The manual test administrator's role has vanished (travelling to colleges, carrying test sheets and stationery, checking ids, supervision, and so on) and the role of the digital test administrator has been created. The latter requires vastly different skills (testing the digital platform, controlling access, generating evaluation reports, mining such data for insights, working with technical teams and maybe even bug fixing!). Such are the transformations in HR, requiring continual reskilling and adaptability to new roles.

Toward off-campus hiring, large enterprises receive thousands of resumes every month and have to resort to some amount of automation to pick the right curriculum vitae (CVs). Off-the-shelf resume screening tools have long been suspect for picking some keywords (if CVs are submitted in a particular format) and rejecting the rest, missing the essence of an application. Our research has




The most awaited
CODING
EXTRAVAGANZA
is back!

<C*deVita/>
The TCS Global Coding Contest

SEASON VII

Opportunity to match your coding talent with the top coders across the globe

9 Coding Languages
1 Winner across the Globe

Participate and WIN!

1st Prize: **Cash Prize USD* 10,000**

2nd Prize: **Cash Prize USD* 7,000**

3rd Prize: **Cash Prize USD* 3,000**

To register log on to:
www.tcscodevita.com



Figure 2: CodeVita: The TCS Coding Competition—Over 100,000 Participants Across Continents

focused on creating an accurate profile of the candidates from multiple public sources. We also have a tool to extract entities from submitted resumes. TEAC, a system currently under trial in the Talent Acquisition Group in TCS HR, builds a standardized model of candidates from their resumes (using a resume miner) and automatically prepares a ranked short-list with matching scores for a given job description. This is being validated in the field by TCS HR.

Some skills need Gen-Y and Gen-Z thinking. At TCS, such hiring is gamified, attracting only those who love to solve challenges and problems. The gamified initiatives for specific skills use our social platform Campus Commune (See Figures: 2 and 3).

Several aspects of the recruiting process (exams, evaluation, CV mining, matching to skills) are digitized in TCS. This data is available for analysis. Even in terms of the personal interview, optimal interview teams can be proposed algorithms to make the interview engaging for the candidates, and productive for the company.

TCS Research continues to work in the areas, such as:

- Predictive models for candidate selection/rejection, and time to join
- Recommending optimal interview panels

Talent engagement

Understanding of the Workforce: Digitization and automation of a number of HR processes ensure more and diverse data about employees' work, activities, events, operations, transactions as well as their interactions

within the organization and with the outside world is available. Analytics can break silos among these diverse data sources and build a knowledge profile of the workforce from different perspectives, such as technical knowledge, domain knowledge, quality assurance knowledge, project management knowledge, products and services knowledge, leadership, and innovation. Leaders can use these knowledge profiles in innovative ways: identify experts for specific tasks; analyze strengths and weaknesses; compare knowledge profiles of different units; and identify opportunities for improvements, identify high potential employees, help in succession planning, predict readiness to handle new businesses and service, understand diversity and culture fitments, and so on.

In the realm of employee engagement, we have created a tool for mining structured and textual responses of TCS' annual employee satisfaction survey PULSE [1].



Figure 3: TCS Campus Commune—An Onboarding Platform that Runs Contests to Hire Specific for Skills

The tool is also integrated in a TCS offering called TRAPEZE™, which is a solution accelerator for business process services (BPS) customers. The tool provides standard reports, charts, and drill-down views, and also includes specially designed analytics algorithms to answer “business questions” related to identification of areas of concern, “interesting subsets” of unusually unhappy employees in terms of shared characteristics, root causes of unhappiness, “what-if” analysis, and generate an optimal plan to achieve desired increase in employee satisfaction under specific constraints. A related system SurveyCoder provides a multiclass multilabel classification of textual responses by matching them with a user-given template of response categories.

Work Design: The traditional and more manual ways of internal staffing and allocation are evolving into more complex models that are more suited for a diverse, virtual, distributed workforce. Allocations will not be piecemeal and manual. Analytics and AI techniques for team formation will help to optimize global organizational objectives, such as optimal utilization of different types of knowledge and experience across projects or services, optimal build-up of employee knowledge profiles, meeting employee career objectives, minimizing costs and maximizing chances of success. Newer AI-driven methods in team formations will bring in more transparency, allow competitive (auction-based) and freelance styles of work, permit more elastic sharing



of employees across projects or services, encourage rotations and role changes, improve behavioral dynamics, and ensure employee career growth.

Best fit engine?

A “BestFit engine” is a TCS initiative driven by its Resource Management Group (RMG) in collaboration with TCS Research. The purpose of BestFit engine is to suggest right openings for people and right people for a given job at times when people are looking for openings or vice-versa. In order to achieve this objective it essentially does three things:

1. Establish a common currency between openings and people’s skillsets
2. Use formal mathematical literature to do a *match* and perform *global optimization*
3. Manage churn: handle obsolescence; emergence of new skills. Deal with imperfect data relating to attrition and recruitment, updating or deleting of skills, allocation or deallocation status, location movement, visa readiness, and so on

Each of these includes numerous sub-items. However, the technology, the math, and the people-domain information blend in well to help TCS keep its allocation rates above 85% and attrition the lowest in the industry [2].

More work is underway along the lines of:

- Predictive models for project team member release
- Predictive models for transitions of employees from one role to another
- Recommending right or left shifts among L1/L2 support teams

- Discovering experts, experienced persons, and specialists for IT infrastructure support

Learning and development: Due to the shorter utility life of technical skills, there is a need for more frequent and effective training plans that are customized to suit aptitudes and learning styles of different employees. The training plans have to be well-aligned with the individual career goals as well as consistent with long-term business projections and target knowledge profile of the organization. Analytics can help significantly here by learning models from past data on these trainings and career evolution of employees.

Skills recommendation engine: TCS is currently using a recommendation engine (see Box on pg.138) based on our research to guide the choice of training from a bewildering array of courses [7].

Automated homework evaluation: TCS conducts an Initial Learning Program (ILP) for campus recruits. The duration of ILP changes according to many factors, primarily the business macro environment. ILP is a residential program wherein participants have to submit homework and undergo assessments every week. The faculty had to set question papers and also correct homework, manually. This was too monotonous both for the faculty and students. Hence, an automatic homework evaluation solution was designed and implemented. This brought in many benefits—reduction in evaluation time, increased self-learning, improvement in scores, cost savings, and more.

Performance appraisals: Performance appraisal (PA) has always been a vital need and a thorny issue for HR, with known

Training Recommendation Engine

TCS easily qualifies as a large “university” since its workforce spends approximately 3.8 million days per year in formal internal trainings. TCS offers a repertoire of approximately 15,000 web-based, and in-classroom courses, covering technical, domain, management, behavioral and soft skills. Given this bewildering variety of options, a Training Recommendation Engine fills up an urgent need by helping employees make informed choices of courses that match their interests, career goals and skills needed in their current role and project.

The Training Recommendation Engine has multiple specially-designed unsupervised machine learning algorithms that ‘learn’ from historical data of the course sequences of employees. As an example, one business unit had 118587 distinct training sequences (one per employee) of at least length 2. There are 5019 unique courses (unigrams), each of which was present in at least one sequence, and which occurred 934167 times, giving an average of 186.13 occurrences per course. Likewise, there were 192168 unique bigrams (ordered pairs of consecutive trainings), which occurred 815580 times, giving an average of 4.24 occurrences per bigram. Using such historical data, the recommendation engine generates a personalized recommendation for next courses for any employee, taking into account her current and past projects and past courses. This tool is integrated into a service called “Ultimatix Recommends.” Ongoing research focuses on extending this tool to a full-fledged personalized career advisory system, that will help employees achieve their career goals, such as becoming a GUI designer or solution architect.

side-effects such as perceptions of unfairness and subjectivity, attrition, and loss of employee morale. No single model of PA is widely accepted, and most suffer from issues, such as halo effect, rater leniency, interpersonal affect, and inaccurate judgments of performance. Analytics can help in improving the quality and effectiveness of the PA process by mining the data and text generated during past and current PA.

AI Tool for PA is a system delivered to TCS HR and uses specially designed data and text-mining techniques on performance appraisal data or text to generate novel and actionable insights or patterns, and to help in improving the quality and effectiveness of the PA process [6].

It generates insights pertaining to:

- Improving the quality of the goal setting process
- Improving the quality of the self-appraisal comments and supervisor feedback comments,
- Discovering actionable supervisor suggestions for performance improvements,
- Discovering evidence provided by employees to support their self-assessments,
- Measuring the quality of supervisor assessments,
- Understanding the root-causes of poor and exceptional performances. (see Box 2)

Analytics can help in improving the quality and effectiveness of the PA process by mining the data and text generated during past and current PA.

AI for Performance Appraisal

We briefly discuss the use of some of the functionalities of TCS' AI Tool for PA towards the appraisal of 156,904 employees in one appraisal period in one TCS business unit. The total number of goals assigned were 2,176,974 (39.66% selected from templates and 60.34% manually created), and the total number of attributes (soft skills) assigned were 1,442,557. The number of words and sentences were 68,345,028 and 4,118,312 for self-appraisal comments and 23,944,158 and 2,903,844 for supervisor comments respectively.

Out of 311,371 manually created goals for employees in DEVELOPER role, HiSPEED automatically matched 210,841 (67.71%) to goals in the template. HiSPEED grouped the remaining 100,530 manually created goals into 100 clusters. Among the 868,813 sentences in supervisor comments (for the DEVELOPER role), HiSPEED found 310,417 (35.7%) as SUGGESTION, 203,741 (23.4%) as APPRECIATION and 28,919 (3.3%) as COMPLAINT. In 1,189,761 sentences in the self-appraisal comments (DEVELOPER role), HiSPEED found 713,900 (60.0%) as EVIDENCE sentences. After identifying specific suggestions, HiSPEED groups them into clusters, such as learn technical skills, improve code quality, improve meetings, improve compliance to schedule, and so on. Most of the 736 supervisors for these employees in DEVELOPER role were highly realistic, the average realistic score being 87.0.

Economists measure country's prosperity by its gross domestic product (GDP). This single number is an aggregation of many other metrics that indicate various factors about the economy in general. HCV in TCS is a number analogous to GDP. It answers a very important question, "Is the organizational capability improving or degrading over time?"

The quest for the ideal performance appraisal system

Human capital valuation (HCV):

Economists measure country's prosperity by its gross domestic product (GDP). This single number is an aggregation of many other metrics that indicate various factors about the economy in general. HCV in TCS is a number analogous to GDP. It answers a very important question, "Is the organizational capability improving or degrading over time?"

At an employee level, TCS defines something called as a Human Development Index (HDI). The HDI comprises the weighted averages of ~200 attributes that are captured in the system per employee. The employees are aware of the attributes considered.

In fact, the whole data processing is completely transparent. The data is picked up from systems where employees enter details such as timesheets, leave requests and update competencies in learning management systems, or go outside the call of their duty to win GEMS (TCS Rewards and Recognition Currency), or file intellectual property rights or update project plans. Formally speaking, there are six dimensions of HCV: technical knowledge, domain knowledge, project management, organization values, growth, and innovation. The ~200 attributes lie within these six dimensions. A custom formula is used to compute HDI. Aggregation of HDIs at sub-unit, unit, ISU, or Geo levels form the HCV for those units.

HCV is a number in the range of 0 to 100. A value of 100 is roughly equivalent to all employees performing at their personal best at all times. The math behind HCV computation is pretty flexible. If and when, any of the six dimensions gains more importance, HDIs and HCVs can be recomputed to obtain new scores. Similarly, over time some attributes may be obsolete and some new may appear. HCV computation can accommodate those too. Finally, HCV computation takes year 2015 as

the base data. All improvements (or degradations) are with respect to this base year. The base year can be changed, if required, to reindex HCV computations. Most works in HCV focus on estimating the monetary value of the workforce (in dollars), particularly the potential for revenue or profit. Our focus is on estimating the more intangible “value” of the workforce in terms of skills, expertise, and experience in technology, business domains, project management, and leadership. Our HCV model allows us to

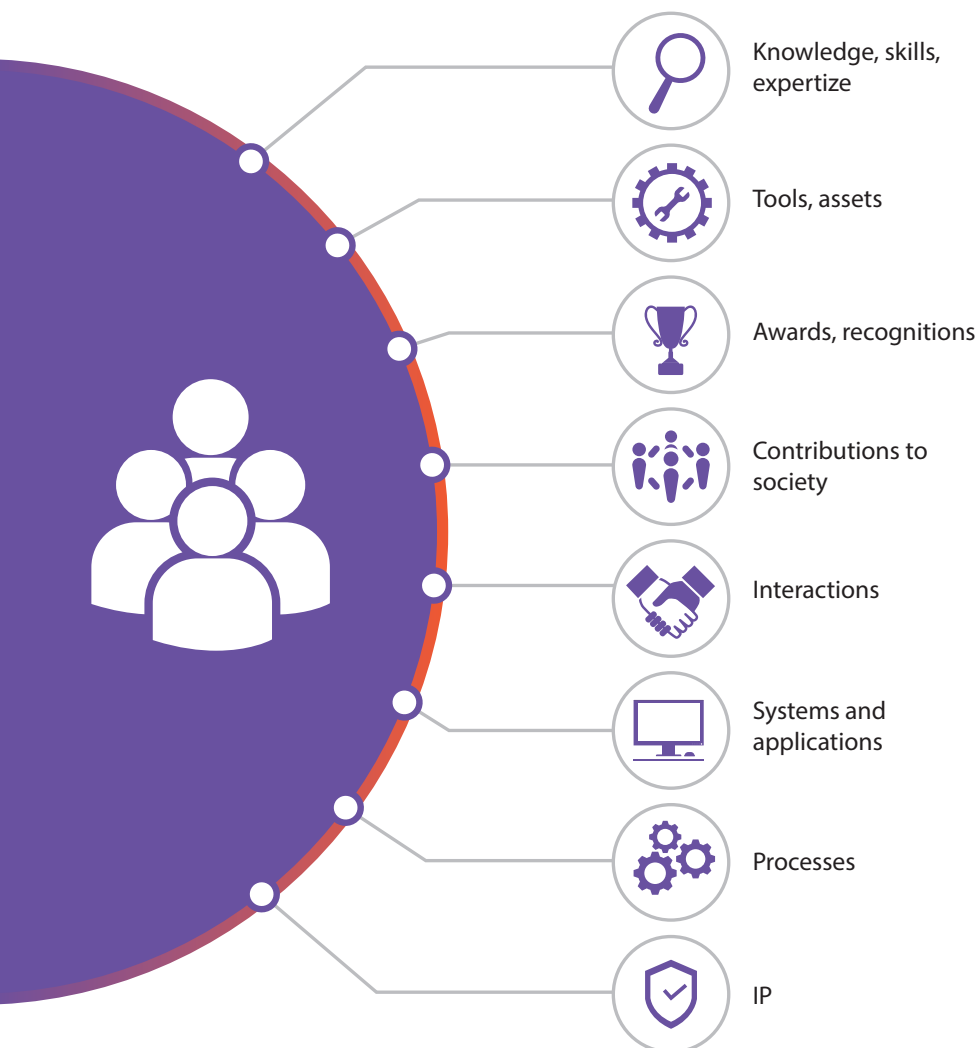


Figure 5 HCV: Many sources of gathering intangible value.

integrate HCV into many other HR use-cases, such as identifying high potential employees [5] and succession planning [3,4].

Work also progresses along the following dimension:

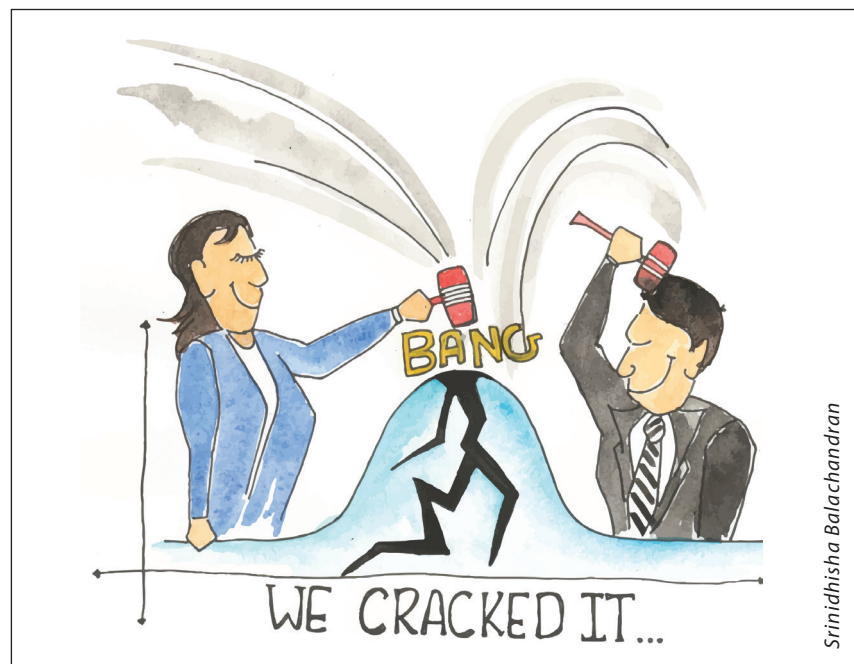
- Attrition prediction, and root-cause analysis
- Predictive models for candidate selection/rejection, and time to join

The research team has built several chat-bots that are in use within TCS to answer user questions related to HR policies (see article 'Ask Cara' in the Data section). This is another example of complete recast of a specific HR role. The bot Cara is trained once on health

insurance (HIS) policy and after that Cara answers most of the mundane, repeat queries on HIS, and learns along the way. It takes fewer associates to train and work alongside Cara, than originally deployed to support HIS—freeing up associates to take up value adding work.

TCS Research and Innovation (R&I) is deeply engaged with TCS HR. It has already delivered several solutions to TCS HR and to HR management systems platform CHROMA.

Enterprises are bracing themselves for robots in the workplace and the disruption that will cause. With rapid digitization, HR is getting a foretaste of it, having to recalibrate itself at every level.



Knell for the Bell curve ?

References

- 1 G.K. Palshikar, S. Deshpande, S. Bhat, *QUEST: Discovering Insights from Survey Responses*, Proc. **8th Australasian Data Mining Conf. (AusDM 2009)**, Dec. 1–4, 2009, Melbourne, Australia, P.J. Kennedy, K.-L. Ong, P. Christen (Ed.s), CRPIT, vol. 101, published by Australian Computer Society, pp. 83–92, 2009.
- 2 V. Vijaya Saradhi, G.K. Palshikar, *Employee Churn Prediction*, **Expert Systems with Applications**, Vol. 38, No. 3, March 2011, pp. 1999–2006.
- 3 M. S. Gharote, A. K. Sodani, G. K. Palshikar, R.R. Srivastava, *Comparison of IT Services*

Organizational Units on the basis of Human Capital Measures, Proc. **IEEE International Conference on Industrial Engineering and Engineering Management (IEEM2015)**, 69 December 2015, Singapore.

- 4 G. K. Palshikar, K. Sahu, R. Srivastava, *After You, Who? Data Mining for Identification of Replacements*, Proc. **3rd International Conference on Mining Intelligence and Knowledge Exploration (MIKE 2015)**, 911 December 2015, Hyderabad, India, R. Prasath, A.K. Vuppala, T. Kathirvalavakumar (ed.s), LNAI 9468, Springer, pp. 543–552.
- 5 G.K. Palshikar, K. Sahu, R. Srivastava, *Ensembles of Interesting Subgroups for Discovering High Potential Employees*, Proc.

Pacific Asia Knowledge Discovery and Data Mining Conference (PAKDD 2016) Auckland, New Zealand, 1922 April 2016, LNCS 9652, J. Bailey, L. Khan, T. Washio, G. Dobbie, J. Z. Huang, R. Wang (Ed.s), Springer, pp. 208–220.

- 6 G.K. Palshikar, M. Apte, S. Pawar, N. Ramrakhiani, *HisPEED: A System for Mining Performance Appraisal Data and Text*, Proc. **4th IEEE International Conference on Data Science and Advanced Analytics (DSAA2017)**, 19–21 October 2017, Tokyo, Japan, pp. 477–486.
- 7 R. Srivastava, G.K. Palshikar, S. Chaurasia, A. Dixit, *What's Next? A Recommendation System for Industrial Training*, **Data Science and Engineering** (Springer), Vol. 3, No. 3, Sep. 2018, pp. 232–247.



Girish Keshav Palshikar

Girish Palshikar is a Principal Scientist at TCS Research and Innovation. His areas of research include machine learning, data mining, text mining, natural language processing and their applications to various domains, including fraud detection and HR management. Since 1992, he has been working in TCS Research, TRDDC, Pune, India. In 2012, he was honored with the title of TCS Distinguished Scientist.

Girish has ~120 publications in international journals and conferences. He is also a visiting lecturer at the Computer Science Department of University of Pune and Government College of Engineering, Pune (GCOEP). Girish has, with his team, published 20-plus papers on the subjects discussed here and these are referenced in his google scholar page linked by QR code below. Girish Keshav Palshikar obtained M.Sc. (Physics) from IIT, Bombay in 1985 and M.S. (Computer Science and Engineering) from IIT, Madras in Jan. 1988.



Amol Khanapurkar

Amol Khanapurkar is Chief Technology Officer for TCS HR. He is responsible for developing a Code evaluation platform for various technologies for both internal and external purposes. His research interest lies in the area of Application Performance engineering. In his current role, he is focused on building scalable HR technologies using well researched HR theories and mapping suitable IT technologies to solve People centric problem statements. As a Researcher he has published several papers and filed patents.

Amol has completed his Masters in Computer Application (MCA 1998 - 2001) from Nagpur University.



All content / information in Artificial Intelligence for Human Resources is the exclusive property of Tata Consultancy Services Limited (TCS) and/or its licensors. This publication is made available to you for your personal, non-commercial use for reference purposes only; any other use of the work is strictly prohibited. Except as permitted under the Copyright law, this publication or any part or portion thereof may not be copied, modified, adapted, translated, reproduced, republished, uploaded, transmitted, posted, created as derivative work, sold, distributed or communicated in any form or by any means without prior written permission from TCS. Unauthorized use of the content/information appearing here may violate copyright, trademark and other applicable laws, and could result in criminal or civil penalties.

TCS attempts to be as accurate as possible in providing information and insights through this publication, however, TCS and its licensors do not warrant that the content/information of this publication, including any information that can be accessed via QR codes, links, references or otherwise is accurate, adequate, complete, reliable, current, or error-free and expressly disclaim any warranty, express or implied, including but not limited to implied warranties of merchantability or fitness for a particular purpose. In no event shall TCS and/or its licensors be liable for any direct, indirect, punitive, incidental, special, consequential damages or any damages whatsoever including, without limitation, damages for loss of use, data or profits, arising out of or in any way connected with the use of this publication or any information contained herein.

©2019 Tata Consultancy Services Limited. All Rights Reserved.

Tata Consultancy Services (name and logo), TCS (name and logo), and related trade dress used in this publication are the trademarks or registered trademarks of TCS and its affiliates in India and other countries and may not be used without express written consent of TCS. All other trademarks used in this publication are property of their respective owners and are not associated with any of TCS' products or services. Rather than put a trademark or registered trademark symbol after every occurrence of a trademarked name, names are used in an editorial fashion only, and to the benefit of the trademark owner, with no intention of infringement of the trademark.