# Xerox' Interests in ALICA

- Extending its offer in Information Access Tools (mostly ML-based)
  - Use the crowd in the most efficient way to get labelled data (when we directly "sell" the models)
  - Alleviate the need of a costly training phases
  - Optimize user's comfort when interacting with the tools
  - Provide efficient solutions to the cold-start problem (new users / new items)

- Deepening the methodologies around Intelligent Crowdsourcing
  - From Business Process Outsourcing (BPO) to Business Process Crowdsourcing (BPC) : How does Xerox move from BPO to BPC? What type of tasks?
  - New labour market, models and relations

- Develop new tools and new services
  - Customer Care (partly automated)
  - Expert Sourcing

# Previous Work related to ALICIA scope

- Information Access in Social Network [FRAGRANCES]
  - Mixed Community-Topic Detection
  - Expert Retrieval (Topic & community Relevance + Centrality)
- Multi-faceted Search & Diversity
  - diversity-ensuring retrieval
  - recall-based active learning
  - entity-based facets (objects are decomposed in "entities/features" of different types)
- Large-scale recommendation systems
  - item/user or expert/task
  - Based on (NN) matrix factorization
- Mechanism Design
- Crowdsourcing:
  - Ethnomethodological studies (XRCE)
  - Multi-platform variability studies (XRCI)
  - On-line optimization and MAB for setting parameter selection
  - Crowdsourcing for image annotation
  - Crowdsourcing for machine translation

# Examples of Crowdsourcing problems

- Model the performance variability
  - Dependency of KPIs (accuracy, time to completion, …) wrt some setting parameters (platform, day/time, price, …)

- How to present the task in order to minimize the error rate:
  - Categorization task: giving examples for guiding workers

- How to split a "complex" task into sub-tasks
  - Eg. Sentence translation (three steps): initially too complex

# Ethnomethdological Study of Turker Nation

- *Study:* extensive reading, collection and analysis of posts and threads

- *Key Findings:*
  - Turking as work, primary motivation of earning
  - Earnings vary but Turking is low wage work
  - Workers aspire to earn at least $7-10/hr
  - Many Turkers choose AMT because they cannot find a good ‘regular’ job or need other income.
  - Turker Nation provides information and support on tools, techniques, tricks of the trade, earning, and learning
  - *Relationships are key:* anonymity and flexibility but want good working relationships, courteous communication and fair pay for fair work
  - Members mostly behave ethically.
  - Preference for regulating the market by sharing information and acting cooperatively

# Contributions to ALICIA (I)

- Task 1 (T1.2)
  - Modelling Higher Order Data
    - User "profiles", item/task profiles, utility metrics, …
- Task 2 (T2.2 & T2.3)
  - Dedicated bandit models (e.g. hidden/latent contextual bandits, special bandits with budget constraints, restless bandits for non-stationary environment, …)
  - Algorithms for learning "teams" (non-compositional set of experts or workers)

# Contributions to ALICIA (II)

- Task 4 (T4.3) : Intelligent Crowdsourcing
  - Targets both "standard" at-home crowds and in-house crowds
  - [For standard crowds] On top of existing platforms (AMT, Crowdflower, MobileWorks, …)
  - Sub-tasks:
    - Gather real-world data (task, worker, context, interactions) . Non-obvious task –"out-of-policy evaluation" problem
    - Assess ALICIA's algo for 5 application domains:
      - Problem resolution in Contact Centers and help-desk (in-house crowds)
      - Document Translation
      - Generating training data (sentiment classification, image categ.)
      - Citizen applications (traffic/parking management)

# A short Overview of Using MAB for Efficient Crowdsourcing

- 5 Papers:
  - L. Tran-Than et al.(2012), *Efficient Crowdsourcing of Unknown Experts using MAB*
  - C. Ho and J.W. Vaughan (2012), *Online Task Assignment in Crowdsourcing Markets*
  - E. Celis et al.(2013), *Adaptive Crowdsourcing for Temporal Crowds*
  - V. Rajan et al. (2013), *CrowdControl: online learning for optimal task scheduling in a dynamic crowd platform*
  - I. Abraham et al., *Adaptive Crowdsourcing Algorithms for Bandit Survey Problem*

# Issues

- Is MAB meaningful for Crowdsourcing ??
  - Not necessarily repeated interactions
  - No tracking of Worker
  - Non-stationary environment
  - A worker does not behave as an arm (not always available on demand)

- How to avoid these issues?
  - Working at an aggregated level
  - Design “ad-hoc” MAB algo (special constraints, restless bandits, …)

- Remember most recent approaches in “learning from crowds” (EM, latent traits, …) – very often deal with redundancy in order to extract the most plausible answer to a task.

# L. Tran-Than et al.(2012), *Efficient Crowdsourcing of Unknown Experts using MAB*

- Arms = workers
- Assumption: stationary environment (single “task”, workers) with individual worker selection
- Ideas:
  - Bounded bandits: global budget B(constraints on the total number of pulled arms weighted by their cost) + individual worker constraint (max number of times on individual arm can be chosen); workers with different costs; minimize regret
  - Two-phase (bounded epsilon-first algo)
    - Exploration : budget = epsilon*B (uniform choice of arms)→ $\mu_a$ /
    - Exploitation: budget = (1-epsilon)*B, based on bounded knapsack problem (greedy approximation, based on “density” $\mu_a / c_a$ )
  - Bounds: cumulative expected regret at most $O(B^{2/3})$→no-regret with suitable choice of epsilon (depending on the max diff. between density values)

# C. Ho and J.W. Vaughan (2012), *Online Task Assignment in Crowdsourcing Markets*

- Arms = task (type); Context= worker; Special Budget Constraint (on individual arms)
- Assumptions:
  - stationary environment (multiple tasks of different types, workers with latent skills wrt each task type) with tracking of individual worker. Workers arrive one at a time randomly: stochastic arrival order (some kind of random permutation). Price $c_i$ depends on the task (type) *i*, not on the worker.
- Ideas:
  - There are $n_i$ tasks instances for each task type *i*. No redundancy: each arm should be pulled $n_i$ times ➔ Budget B ($n_i$ x $c_i$ ) is also fixed (but not a standard budgeted-MAB). If skill-levels known, standard assignment problem (LP), but on-line; Use of dual formulation → gives an optimal threshold $x_a$ for each task type, such that the task assignment decision is argmax( $\mu_{a,t}$ -$x_a$)
  - Two-phases (bounded epsilon-first)
    - Exploration : budget = epsilon*B (approximately uniform choice of arms and contexts $\mu_{a,t}$
    - Exploitation: budget = (1-epsilon)*B, based on the Dual Task assigner: when worker *a* arrives, give her the task i = argmax($\mu_{a,t}$ - $x_a$) (among tasks whose budget is not yet exhausted) and decrement the task - *i* counter.
  - Bounds: there are some bounds on the "ratio of performance" (wrt to the offline optimal) with suitable choice of epsilon

# E. Celis et al.(2013), *Adaptive Crowdsourcing for Temporal Crowds*

- Arms = job posting parameter configuration (platform, deadline, …). Single task. Reward (accuracy, completion time)=function(job posting parameters). No context.
- Assumption: non-stationary environment (time-varying reward function) – Workers are aggregated into "platforms".
- Ideas:
  - Continuously balance exploration/exploitation (but no need to be as defensive as against an malicious adversary)
  - Random walk model of the reward distribution, defined by a sequence of means $\mu_{0,a}$, $\mu_{1,a}$, $\mu_{2,a}$,… for each arm a: lazy random walk with reflection boundaries on $[\mu_{min},\mu_{max}]$ and step Δ (→Markov chain whose stationary distribution has mean $\mu max+\mu min/2$ (used for simulation and for a UCB-like strategy)
  - Must adapt some MAB strategies for this non-stationary environment:
    - Windowed ε-greedy strategy (estimates $\mu\downarrow a$ on the *m* last observations)
    - EXP3-m strategy: reinitialise arm weights to uniform every *m* iterations
    - Windowed ε-greedy strategy with (temporary) arm elimination, based an extra "pseudo" confidence interval on the variation of $\mu\downarrow a$ since last update ($\gamma\sqrt{t-\tau a}$ where *t* is the current time and $\tau_a$ is the time where a was pulled for the last time
  - No bounds. Anyway, these aren't no-regret algo. Last algo gives best strong regret for different rewards (accuracy and response time). EXP3m gave poor performance.

# V. Rajan et al. (2013), *CrowdControl: online learning for optimal task scheduling in a dynamic crowd platform*

- Arms = job posting parameter configuration (platform, deadline, ...) ***x***.
  Reward (accuracy, completion time)=function(job posting parameters).
  No context. Single task.
- Assumption: stationary environment (stationary reward function) – Workers are aggregated into "platforms".
- Ideas:
  - Discretize the domain of x into K cells and use K-armed bandits
    → Use Thompson sampling (TS) on "bernouilli-sed" reward (1 if better than previous performance; else, 0)
  - On-line stochastic optimisation: simultaneously learn the (noisy) reward function and optimize it (actually its accumulated version, ie minimize the regret)
    → Use GP-UCB [the reward function is modeled a a Gaussian Process; choose ***x*** =argmax $\mu_{t-1}(x)+\sqrt{\beta_t}\,\sigma_{t-1}(x)$and update $\mu_t(x),\sigma_t(x)$ ]. Virtually infinite number of arms (i.e. continuous arms)

- Experiments with real and simulated datasets and one single parameter (the batch size) : TS better than GP-UCB

# I. Abraham et al., *Adaptive Crowdsourcing Algorithms for Bandit Survey Problem (I)*

- Arms: crowd “segments” (between crowd platform and single worker); different costs for each segment. Single task instance (!), with *O* possible options.

  Rewards: a distribution over possible options (‘bandit survey’), then summarized as a single “gap” value: “gap” is the difference between the largest and the second largest values in the distribution (noted $\varepsilon_a$ for the unknown distribution; $\varepsilon_a$ from the sample)
- Assumptions: stationarity, segments (arms) identifiable and available “on demand”
- Ideas
  - Solve the problem by specifying independently)
    - a stopping criterion (we have enough observations to take a decision on the correct option with bounded guaranteed accuracy)
    - a crowd segment selection criterion (based on a “virtual reward” which is the estimated time before “stopping” if we always choose this arm
  - Composite stopping criterions:
    - We have a stopping criterion for all arms individually, but also a global stopping criterion (as if all crowd segments were merged and we consider the merged distribution of options also as a new arm)
    - Once any arm (including the new merged arm) has reached the stopping criterion , we stop the process and take the highest probability options as the final answer

## I. Abraham et al., *Adaptive Crowdsourcing Algorithms for Bandit Survey Problem (II)*

- Stopping Criterion:
  - Stop if $N_{i,t} > C.1/\varepsilon_{i,t}^{2}$ ($N_{i,t}$: number of times arm *i* was chosen up to time t)
    - For a specified C formula (depending on δ), the error rate is at most $O(\delta)$
    - The expected stopping time is at most $O(\varepsilon_{i}^{-2})$
- Arm Selection Criterion:
  - Virtual reward: $r_i = \varepsilon_{i} / \sqrt{c_i}$
  - UCB-like criterion: argmax $\varepsilon_{i,t} / \sqrt{c_i} + C / \sqrt{c_i * N_{i,t}}$