package com.twitter.home\_mixer.functional\_component.filter

import com.twitter.home\_mixer.util.CandidatesUtil

import com.twitter.home\_mixer.util.TweetImpressionsHelper

import com.twitter.product\_mixer.component\_library.model.candidate.TweetCandidate

import com.twitter.product\_mixer.core.functional\_component.filter.Filter

import com.twitter.product\_mixer.core.functional\_component.filter.FilterResult

import com.twitter.product\_mixer.core.model.common.CandidateWithFeatures

import com.twitter.product\_mixer.core.model.common.identifier.FilterIdentifier

import com.twitter.product\_mixer.core.pipeline.PipelineQuery

import com.twitter.stitch.Stitch

/\*\*

\* Filter out users' previously seen tweets from 2 sources:

\* 1. Heron Topology Impression Store in Memcache;

\* 2. Manhattan Impression Store;

\*/

object PreviouslySeenTweetsFilter extends Filter[PipelineQuery, TweetCandidate] {

override val identifier: FilterIdentifier = FilterIdentifier("PreviouslySeenTweets")

override def apply(

query: PipelineQuery,

candidates: Seq[CandidateWithFeatures[TweetCandidate]]

): Stitch[FilterResult[TweetCandidate]] = {

val seenTweetIds =

query.features.map(TweetImpressionsHelper.tweetImpressions).getOrElse(Set.empty)

val (removed, kept) = candidates.partition { candidate =>

val tweetIdAndSourceId = CandidatesUtil.getTweetIdAndSourceId(candidate)

tweetIdAndSourceId.exists(seenTweetIds.contains)

}

Stitch.value(FilterResult(kept = kept.map(\_.candidate), removed = removed.map(\_.candidate)))

}

}