

Министерство образования и науки Республики Казахстан
АО «Национальный центр государственной научно-технической экспертизы»

Дорожная карта
по направлению «Информационно-коммуникационные
технологии»
под-направлению «Базовые прикладные технологии в ИКТ»

Астана, 2013

Дорожная Карта «Базовые прикладные технологии и направления ИКТ»

1. Паспорт

Наименование Дорожной Карты	Развитие базовых прикладных технологий и направлений в ИКТ в Каза
Приоритетные тематики	<ul style="list-style-type: none"> • Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык) • Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полу-автоматической поддержки решений • Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS) • Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data
Основание для разработки, цели и задачи	
Основной результат (продукты/услуги, технологии)	Основными результатами является появление новых продуктов и технологий, развитие и обновление существующих.
Этапы реализации	2014-2020 2020-2025 2025-2030
Основные ресурсы и участники процесса реализации Дорожной карты	МОН РК, МТК РК, АО «Зерде», АО «НЦГНТП» университеты и НИИ

2. SWOT-анализ поднаправления «Базовые прикладные технологии»

Сильные стороны	Слабые стороны
<ul style="list-style-type: none"> • Наличие государственной программы развития отрасли («Информационный Казахстан 2020» утвержденной Указом Президента Республики Казахстан №464 от 8 января 2013 года); • Наличие опытных отраслевых институтов развития ИКТ, операторов «в сфере информатизации» и «единой транспортной среды государственных органов» • Наличие государственной ИК-инфраструктуры; 	<ul style="list-style-type: none"> • Сравнительно небольшой потребительский рынок и небольшая плотность населения, в связи с которой увеличиваются расходы на построение и поддержание телекоммуникационной инфраструктуры; • Потеря отечественной школы по ряду направлений исследований в базовых прикладных технологиях • Отсутствие квалифицированных специалистов и недостаток практики в области облачных и мобильных

<ul style="list-style-type: none"> • Наличие данного направления в перечне критических технологий определенных решением Совета по технологической политике под председательством Премьер – Министра Республики Казахстан от 20 декабря 2012 года. • Наличие и возможность подготовки молодых квалифицированных специалистов в базовых технологиях ИКТ. • Ускорение темпов развития государственной экономики, увеличение инвестиционной привлекательности РК на международном уровне • Высокий научный потенциал в отдельных направлениях науки • Сохраняющийся достаточно высокий образовательный уровень населения 	<p>технологий;</p> <ul style="list-style-type: none"> • Низкая конкурентоспособность отечественной ИКТ отрасли • Каналы связи и слабая пропускная способность «Единой транспортной среды государственных органов» • Низкий уровень публикаций по ИКТ в целом и по базовым прикладным технологиям • Отсутствие стандартов и политик по информационной безопасности «облачных» решений. • Превалирование импорта над технологическими разработками и трансфертом технологий • Низкая культура лицензирования, нарушение отечественными компаниями авторских прав,
<p>Возможности</p>	<p>Угрозы</p>
<ul style="list-style-type: none"> • Рост спроса на ИКТ продукцию в мире Экономическая и технологическая эффективность от внедрения облачных технологий (по данным IDC, увеличение мобильности (46%), продуктивности (41%), стандартизации (35%) и снижение затрат (20%) по данным «Даталайн», вывод серверной инфраструктуры в облако обойдется организациям на 15–70% дешевле размещения серверов на собственной площадке в зависимости от конфигураций); • Создание и реализация целевых технологических программ развития критических технологий; • Множество вариантов внедрения облачных и мобильных технологий в повседневную жизнь. • Выход на региональные рынки (Таможенный союз, Центральная и Восточная Азия, Прикаспийский регион, включая Закавказье, Западный Китай) Стабильная макроэкономическая среда, благоприятная политическая обстановка в стране для привлечения иностранных инвесторов в ИКТ • Низкий порог вхождения на мировой рынок индустрии ИКТ • Начало перехода промышленных предприятий на интенсивные методы роста эффективности с использованием 	<ul style="list-style-type: none"> • Отсутствие системных мер по организации НИОКР • Высокий уровень мировой конкуренции • Преобладание конкурентных импортных решений других стран (России, Китай); • Отказ от широкого внедрения ввиду угроз информационной безопасности; • Перегрузка каналов связи в связи с увеличением трафика (неспособность инфраструктуры удовлетворить потребности в каналах связи). Ограниченность доступа к инвестиционным средствам, неразвитость финансовых инструментов, отсутствие венчурных фондов • Утечка высококвалифицированных специалистов

Основные контуры развития ИКТ к 2030

1. К этому времени в Казахстане будут построены высокоскоростные каналы связи до каждого населенного пункта, с шириной канала по требованию.

Причем технологии построения коммуникаций будут гибридными, учитывая геоэкономику региона. Возможно, для передачи данных будут использоваться и электрические сети.

Все организации и домохозяйства будут подключены одному из Дата-центров, государственных или частных.

Сами Дата Центры могут динамически объединяться в сети через шлюзы для оптимальной балансировки нагрузки и распределенной обработки Big Data.

При этом услуги оплачиваются «по факту», то есть только те, которые потребили клиенты. Информация становится таким ресурсом как тепло, вода и электроэнергия.

2. Все предприятия, в зависимости от масштаба и специфики технологического процесса имеют четкую политику создания, потребления, переработки и передачи данных как составная часть основного технологического процесса.

При этом большая часть процесса обработки данных находится в услугах аутсорсинга.

Это экономически выгодно, поскольку на рынке есть крупные ИКТ-компании, специализированные на отраслевых решениях и их услуги дешевле и качественнее, чем создание собственного ИКТ. При этом такие задачи как учет, финансы, планирование бюджета, логистика, управление активами, закупки и прочие уже не разрабатываются самими предприятиями. Они предприятиями покупаются, именно как услуга, у поставщиков.

Технологический процесс предприятий обеспечивается ИКТ инструментами на всех этапах.

3. В социальной сфере (государственные услуги, образование, здравоохранение), торговле, СМИ практически вся работа перейдет на электронные форматы и в сеть.

Это обеспечит индивидуализацию сервиса для каждого клиента, сервисы будут динамически адаптированы к психофизическим особенностям учащихся, состоянию и диагнозам пациентов, интересам и потребностям покупателей.

Чтобы перейти к такому уровню необходимо выполнить главные задачи:

1. Подготовить необходимые компетенции как у производителей, так и потребителей.

Должны быть разработаны новые профессиональные стандарты для активного использования ИКТ во всех отраслях экономики, учебные программы в применении ИКТ в учебном процессе.

Разработка таких стандартов требует проведение предварительных исследований и трансферт готовых технологий

2. Необходимо разработать нормативные и проектно-сметные документы, где предусмотрены ИКТ-емкие технологии, которые позволят обязать применение ИКТ начиная от этапа строительства промышленных и социальных объектов

Для этого требуется проведение исследований и анализа существующих международных стандартов и нормативов, их адаптацию под казахстанские условия.

Если еще не существует международных стандартов и нормативов, то необходимо разработать новые.

3. Необходимо создавать крупные отечественные центры R&D по основным технологическим направлениям исследований и интегрировать их с мировыми центрами для первоначального трансферта технологий и стажировки отечественных молодых исследователей через программу Болашак.

При этом, ВУЗы обязаны включаться в сеть исследовательских организаций для координации и обмена результатами исследований.

Всемерно поощрять ВУЗы для кооперации с ведущими мировыми университетами и исследовательскими центрами и давать им академическую свободу.

4. Создавать R&D центры для смежных дисциплин и открыть новые специальности и учебные программы для продвижения профессий будущего.

3. Перечень тематик исследований

При отборе перспективных тематик для под-направления «Базовые прикладные технологии», были отобраны наиболее актуальные для Казахстана, с учетом мировых трендов и с учетом, имеющейся в стране научной школы.

Соединение в один блок задач искусственного интеллекта и моделирования сложных объектов и систем в данном случае вызвано необходимостью комплексной поддержки прикладных задач на системной уровне.

Практическими потребностями, прежде всего стало растущее развитие новых направлений в промышленности.

- Переход на новые методы информатизации, особенно в геологии и геофизики
- Переход мировой и казахстанской индустрии на новые стандарты промышленной интеграции (Индустрия 4.0)
- Цифровые технологии в проектировании и в промышленном производстве
- Развитие 3D принтинг и умное производство
- Роботизация производства (интеллектуальные роботизированные системы)
- Повсеместное комплексное внедрение информационных технологий в биотехнологии, аграрном секторе и в медицине.
- Вовлечение все более широких слоев населения в электронные услуги и социальные сети

Общими системными основаниями во всех этих направлениях и тенденциях являются системы искусственного интеллекта, поддержки принятия решения, вопросы оценки и моделирования в условиях неопределенности и близко сопряженные с ними задачи обработки огромного числа неструктурированных данных. Новый поток таких порождается развитие новых факторов как

- Цифровое производство и полный мониторинг его циклов
- Социализация контента и бизнеса (социальные сети)
- Мобильные системы и приложения
- Лавинообразный рост сенсорных датчиков и устройств

По данному поднаправлению в качестве приоритетных тематик выбраны следующие:

- Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)
- Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полу-автоматической поддержки решений
- Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)

- Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data

Отбор проводился из так называемого длинного списка тематик, продуктов и технологий В длинный список технологий Дорожной Карты «Базовые прикладные технологии» входят следующие:

Продукты	Технологии	Тематики научных исследований
Системы искусственного интеллекта и моделирование сложных систем	Технологии распознавания образов и зависимостей (Pattern recognition)	<ul style="list-style-type: none"> • Разработка и использование алгоритмов компьютерного зрения для распознавания образов, реконструкция 3D изображения и приложения стерео зрения, автоматический анализ видео. • Анализ мультимодальных сигналов и сенсоров включая невидимый спектр • Исследования по автоматическому анализу текста, распознанию и генерации речи (с фокусом на казахский язык)
	Экспертные системы и модели принятия решений	<ul style="list-style-type: none"> • Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений • Исследования по разработке новых моделей эффективного управления и прогнозирование рисков • Исследования по разработке новых моделей извлечения (knowledge elicitation) знаний • Исследования по разработке архитектуры коллективного интеллекта и эффективный краудсорсинг включая новые методы организации сетевой экспертизы
	Технологии моделирования и анализа сложных процессов	<ul style="list-style-type: none"> • Исследования по разработке прикладного математического моделирования с неполными и гетерогенными данными включая новых методов приближенного вычисления (numerical approximation) • Исследования по разработке новых математических методов моделирования, распознания закономерностей и классификации сложных и стохастических процессов, включая обучения без присмотра (unsupervised learning), не линейную оптимизацию и приложения квантовой теории вычислений • Исследования по прикладному применению существующих методов искусственного интеллекта в

		неисследованных доменах включая применение многоступенчатого анализа
Новый интерфейс взаимодействия «человек-машина»	Технологии взаимодействия человека и компьютера (Human-Computer Interaction)	<ul style="list-style-type: none"> • Исследования по разработке интерфейсов для натурального взаимодействия и общения • Исследования по разработке адаптивных (к пользователю) приложений и информационной архитектуры • Исследования по разработке новых методов, практик и нормативов для удобного пользования (User Experience)
	Машина –машина (M2M)	<ul style="list-style-type: none"> • Архитектура межмашинного взаимодействия
BigData и аналитика	Аналитика в реальном времени	<ul style="list-style-type: none"> • Исследования по разработке новых методов потоковой обработки большого количества данных в реальном времени • Исследования по оптимизации вычислительной нагрузки и новые модели вычисления в операционной памяти
	BigData и обработка неструктурированных данных	<ul style="list-style-type: none"> • Исследования технологий BigData для обработки неструктурированных (полуструктурированных) данных • Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS) • Исследования и разработка новых методов для параллелизации сложных вычислений для обработки больших объемов данных включая усовершенствование алгоритмов на основе Map Reduce • Новые модели и методы сжатия и обобщения неструктурированных данных • Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data
	Технологии совместного использования контента (Content Sharing) платформы для менеджмента и организации знаний	<ul style="list-style-type: none"> • Исследования по разработке новых моделей и организации коллаборативной деятельности и краудсорсинга (crowdsourcing) • Исследования по разработке новых методов синхронных и асинхронных коммуникаций • Исследования по разработке новых методов обеспечения авторского права и интеллектуальной собственности

Краткое резюме тематик и технологий

Системы искусственного интеллекта и моделирование сложных систем

Объединение искусственного интеллекта и моделирования сложных систем в одно направление создает базис для развития многих прикладных направлений.

Тенденцией и ключевым трендом будущего является широкое проникновение технологий искусственного интеллекта во все отрасли промышленности, государственного управления и социальной жизни.

Уже в ближайшем будущем многие повседневные операции начального и среднего уровня будут автоматизированы. Причем, это не простое автоматическое выполнение рутинных процедур. Произойдет реальное замещение части функций, выполняемых человеком, интеллектуальными операциями на основе искусственного интеллекта.

Основной упор сделан на создание интеллектуальных компьютерных программ.

При этом, искусственный интеллект не обязательно ограничивается биологически правдоподобными методами.

Многогранность систем искусственного интеллекта дополняется сопряженными с ними задачами моделирования сложных систем. Данное объединенное направление научных исследований относится к наиболее сложному, состоящему из очень большого количества под-направлений и тематик, которые, в свою очередь, также состоят из своих составных компонентов.

Необходимость формирования такого объединенного направления вызвана следующим.

Определяющим фактором стало использование возможностей моделей искусственного интеллекта и прикладных систем на его основе для решения сложных прикладных задач. Эти прикладные задачи, в свою очередь, являются актуальными для конкурентоспособного развития промышленности и социальной сферы Казахстана.

К ним относятся инструментарий ИИ, встроенные интеллектуальные системы, обработка естественных языков, распознавание образов и зависимостей, компьютерное зрение и другие.

Технологии распознавания образов и зависимостей (Pattern recognition)

Применение моделей и методов искусственного интеллекта эффективно применяется для распознавания текстов и речи, с автоматическим переводом на разные языки.

Исследование состоит из нескольких взаимосвязанных компонентов – распознавания речи (текстов), перевод текста и речи на основные мировые языки, включая казахский, реализация в виде программно-аппаратных и комплексов и компьютерных программ, использование облачных технологий.

Необходимо использование передовых методов семантического и морфологического анализа (расознавания) с речи и текстов, применение облачных технологий хранения и распределенных вычислений и обработки Bigdata. Математические исследования по формализации структурного (морфологического) анализа, сжатия информации, применения математических методов обработки цифровых и аналоговых сигналов, выделения устойчивых последовательностей и построение морфологической интерпретации.

Результаты исследований применяются для всех направлений и отраслей экономики и социальной жизни, где необходим скоростной перевод текста и речи для лингвистического обеспечения. Это касается прежде всего сферы образования, перевода художественной и технической литературы, использования опыта и знаний иностранных консультантов и инвесторов в Казахстане, распространения казахстанских технологий / Исследования предполагается высокий уровень новизны научных и опытно-конструкторских результатов.

В них предполагается использование самых современных моделей и методов не детерминистического моделирования (на основе вероятностей), включая лингвистическую обработку, распознавания образов, речевого синтеза.

Значимость проекта заключается в вовлечении и доступе всех слоев общества, включая жителей удаленных населенных пунктов и не владеющих основными мировыми языками к мировым базам знаний, библиотекам и огромным сетевым интернет ресурсам.

Приоритет и фокус на лингвистические и фонетические особенности казахского языка обеспечат вовлечение казахского языка в мировое лингвистическое пространство.

Актуальность и необходимость качественного перевода и синтеза казахского языка и речи на казахском языке особенно велика в связи с широкой интеграции нашей страны в мировое экономическое и культурное пространство.

Отмечается высокая важность и актуальность исследований в данном направлении уровень научной новизны и значимость направления высокие

В этой связи, в число приоритетных выбрана тематика «Исследования по автоматическому анализу текста, распознанию и генерации речи (с фокусом на казахский язык)»

Экспертные системы и модели принятия решений

Исследования направлены на разработку современных методов принятия решения с использованием методов искусственного интеллекта. Они являются основой экспертных систем широкого профиля. Практически во всех отраслях экономики и социальной жизни существует проблемы обоснованного выбора вариантов, на основе сложившихся процедур принятия решения. Выбор решения всегда производится в условиях неопределенности.

В целом, исследование относится к классу систем поддержки принятия решений и интеллектуальных систем (Decision Support and Intelligent Systems) и включает в себя обычные и групповые системы поддержки принятия решений, прикладные методы искусственного интеллекта, экспертные системы и предиктивную аналитику.

Применение результатов данной научной тематики позволит автоматизировать процесс принятия решения в различных отраслях, в ситуациях связанных с принятием решения из многих вариантов развития. Математические основы и компьютерные технологии позволят создать оболочку и среду функционирования при которых выбор является обоснованное и доказательным. Применяются различные методы многокритериального анализа, оптимизационный алгоритмы, ускоряющих процедуру принятия решения. При этом, часто, применяются технологии с участием человека-эксперта. Участвующего в процессе обоснованного и согласованного выбора.

Исследования предполагают применение всех современных научно-инженерных подходов по теории выбора обоснованного решения. Предполагается создание программного комплекса, на основе математических моделей принятия решения.

В связи с развитием самих компьютерных технологий, вычислительных мощностей и средств отображения, возрастает значимость системы для применения в самых отраслях человеческой деятельности. Вопросы обоснованного и квалифицированного выбора решений (вариантов) в сложных инженерных и управленческих задачах являются актуальными

В этой связи, в число приоритетных выбрана тематика «Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений»

BigData и аналитика

BigData и обработка неструктурированных данных

Сущность исследования заключается в использовании большие вычислительных и аналитических возможностей технологии BigData (больших данных) для использования в различных сферах промышленности, науки и социальной жизни, использующих обработку больших массивов неструктурированной и слабосвязанной информации.

Ранее бизнес-анализ строился на основе структурированных данных, но с появлением проблемы Больших Данных обострилась необходимость в работе с неструктурированными данными, составляющими свыше 80% всех хранимых данных, причем их количество нарастает на порядок быстрее по сравнению со структурированными данными.

Особенно это касается обработки геофизических данных, данных аэрокосмического мониторинга, биотехнологий, селекционной работы, обработки социальных сетей, семантического анализа текстов и многих других.

Уровень предлагаемых решений достаточно высок, потому что исследование находится в русле мировых трендов и имеет большую значимость в части прикладных применений.

Масштаб приложений охватывает практически все отрасли экономики, так как практически везде в настоящее время используется интенсивная и распределенная обработка данных.

Согласно обновленному в январе 2013 года прогнозу IDC, мировой рынок технологий Bigdata будет расти в среднем на 31,7% в год и достигнет к 2016 году \$23,8 млрд. Аналитики отмечают бурный тем роста соответствующих технологий и сервисов, а также активное формирование подсегментов этого рынка, в том числе появление новых. На протяжении ближайших четырех лет эти высокие темпы прироста рынок сохранит, полагают эксперты.

Рост отдельных сегментов мирового рынка Big Data будет варьироваться при этом от 21,1% для сервисов до 53,4% для систем хранения. Высокий темп роста в сегменте систем хранения больших данных связан с особенностями их востребованности на настоящий момент.

Недостаток специалистов на рынке с компетенцией в области аналитики и управления большими данными приведет к тому, что все больше покупателей будут выбирать облачные решения, а также аутсорсинг аналитических сервисов, не только непосредственно технологий.

Решения Big Data для обработки данных, сгенерированных машинами, начнут набирать обороты ближе к концу прогнозного периода. Таковы основные тренды, ожидаемые IDC на этом рынке.

Мировой объем рынка Big Data и аналитических решений на ее основе по итогам 2013 года вырастет почти на 50%. При этом распределение рынка Big Data на сегменты было следующим: 41% пришелся на услуги, 30% - на программное обеспечение, 14% - на серверные технологии, 12% - на технологии хранения данных, 3% - на сетевые технологии.

Большая часть рынка BigData (но не вся) принадлежит рынку бизнес аналитики, который по прогнозам IDC в 2012-м году составил около \$100 млрд.

Согласно прогнозам IDC - рынок связанный с технологией BigData вырастет до \$16,9 млрд. в 2015 году. То есть продемонстрирует совокупный среднегодовой темп роста (CAGR) равный 39,4%, или будет расти примерно в семь раз быстрее, чем ИКТ рынок в целом.

Другие технологические области, на которые оказывают существенное влияние большие данные, это инфраструктура приложений, промежуточное ПО, ПО хранения информации, системы управления БД, интеграция данных, BI и SCM.

Технологические платформы BigData

Исследования предполагают применение новых и эффективных моделей и алгоритмов. В том числе, специализированные алгоритмы управления и синхронизации процессов обработки данных, применения технологии Hadoop, MapReduce и других технологических решений.

В настоящее время возникает большая необходимость в построении технологической платформы для работы с Большими данными (BigData)

Такая платформа, полноценно решающая проблему Больших Данных, должна обеспечивать возможность работы с данными всех типов, представленных в произвольных форматах, а для этого она должна включать:

- средства визуализации и обнаружения — данные должны быть представлены пользователю в виде, удобном для восприятия, понимания содержания и поиска в различных источниках;
- инструменты аналитики цифровых и текстовых данных в пакетном режиме и режиме реального времени;
- системы обработки потоковых данных;
- средства поддержки традиционных хранилищ данных и систем оперативной обработки транзакций.

Сегодня пока нет таких полнофункциональных платформ, и под платформами для Больших Данных понимают самые обычные традиционные компьютерные платформы, адаптированные для работы с большими объемами данных.

В этой связи, в число приоритетных были выбраны следующие тематики:

- **«Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)»**
- **«Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data»**

4. Этапы реализации Дорожной карты

Краткое видение по развитию этапов

На начальном этапе проводится детальный анализ целевых рынков, а также анализ и измерение спроса промышленности на результаты ИКТ в целом и тематик в частности.

При этом определяется потребность в приложениях с глубоким уровнем исследований и разработок, таких как приоритетное применение мобильных приложений, сенсоров, искусственного интеллекта, моделирования сложных систем и других.

Далее, происходит формирование знаний, компетенций и подготовка кадров.

В этот же период первоначального накопления сил и ресурсов, основным инструментом должен стать трансферт технологий, совместные исследования и совместные R&D.

Ориентировочно, ко второму этапу – до 2020 года, должно произойти появление наших, казахстанских школ и направлений.

К этому же времени должна завершиться подготовка всех необходимых стандартов, требований и сертификатов по ИКТ.

Учитывая полноценное и глубокое присутствие ИКТ по всех отраслях промышленности, науки, социальной жизни, особое развитие получают прикладные технологии и тематики, находящиеся на стыке ИКТ и отраслей.

Более того, во многих отраслях индустрии ИКТ должны стать составной и обязательной частью промышленных и социальных технологий, встроенные в технологические карты процессов.

Тематики приоритетных исследований

1. Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)
2. Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений
3. Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)
4. Разработка новых технологических платформ для создания прикладных аналитических решений по BigData

Операционный план – краткосрочный период

Общая часть для всех тематик

• В рамках стратегических программных документов подготовка государственных решений по развитию приоритетных исследований, как базовой основы для остальных направлений в ИКТ и в промышленности в целом

Поддержка и развитие исследований

- Создание Центров R&D
- Формирование специализированных центров при университетах и научно-исследовательских центрах
- Разработка совместных детальных исследований по тематикам
- Подготовка профессиональных кадров
- Подготовка учебных процессов и курсов у университетах
- Создание Центров разработки технологических платформ
- Интеграция тематик
 - применение единых или близко сопряженных моделей и методов на основе искусственного интеллекта и системного анализа
 - использование технологических платформ обработки bigdata

Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)

- Изучение технологий, моделей, методов и мировых тенденций по тематике и отбор наиболее перспективных из них
- Разработка моделей и технологий математической лингвистике и распознаванию обменов
- Разработка и внедрение компьютерных программ по анализу и распознаванию текстов
- Разработка технологии Text mining
- Интеграция с технологическими платформами BigData

Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений

- Разработка математических моделей и технологий по тематике
- Классификация методов, алгоритмов и различных систем поддержки принятия решений

- Отбор, формирование и развитие наиболее актуальных систем
- Разработка и внедрение промышленных систем компьютерных систем принятия решения

Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)

- Формирование моделей и методов работы с неструктурированными данными
- Отработка технологий Big Data
- Отработка систем хранения и обработки структурированных и неструктурированных данные из различных источников, включая правительственные фонды данных, базы данных компаний и общедоступные данные социальных сетей.
- Интеграция с облачными решениями
 - Подключение к казахстанским облачным сервисам
- Разработка программных комплексов, платформ и программных сред (framework) для распределенной обработки данных различных типов

Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data

- Разработка архитектуры и спецификаций по нескольким базовым технологическим платформам Big Data
- Развития прикладных решений на основе зарекомендовавших себя технологических платформ (например Hadoop)
- Применение платформы Big Data для сбора и анализа данных в промышленности, преимущественно в сфере ГМК и нефти-газовых отраслей
- Применение платформы Big Data для исследований в сферах биотехнологий и медицины
- Интеграция технологических платформ с моделями искусственного интеллекта для изучения сложных зависимостей

Стратегический план – среднесрочный период

Особенностью этого периода станет повсеместное развитие Big Data

Применение технологий Big Data генерируемая и используемая в областях

- Сенсорные датчики
- Гибкая и носимая электроника
- Новые производственные процессы
- Нанотехнологии
- Искусственный интеллект
- 3D-визуализация и рендеринг
- Цифровые и беспроводные сети
- Комплексное моделирование
- Пользовательский контент
- Архивы, метаданные, поиск
- Безопасность и защита данных
- Взаимодействие человека и машины

Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)

- Автоматическое распознавание текста и автоматический перевод на мировые языки
- Создать системы искусственного интеллекта, понимающей речь и управляющей знаниями.
- Создание системы ИИ для перевода, распознавания и анализа текстов, вычленение необходимых кусков текста из огромных массивов данных. При этом разрабатываемая технология не зависит от выбора языка

Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений

- Разработка новых алгоритмов на базе интегрированных технологических платформ,
- Внедрение систем принятия решения в различных отраслях, прежде всего энергетика, диспетчерские центры, телемедицина, транспорт, беспилотный автоматический транспорт, инвестиционные и страховые компании
- Массовое внедрение роботизированных систем
- Автоматическая обработка сигналов от сенсорных устройств
- Интеграция в системы промышленной автоматизации и Scada систем
- Пространственный анализ, использующих топологическую, геометрическую и географическую информацию в данных;
- Переход на повсеместную визуализацию аналитических данных и визуальную аналитику

Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)

- Интеграция и обработка всех типов информации
- Реализация в сложных сетях и компьютерных кластерах, с применением высокопроизводительных компьютерных систем

Разработка новых технологических платформ для создания прикладных аналитических решений по BigData

- Создание и промышленная эксплуатация платформ для поддержки *BDAP (BigData application platform)*
- Создание специальных систем и машин поддержки аналитики.
- Появление платформ нового поколения, интегрированные с суперкомпьютерами и сетями нового поколения, со скоростными методами обработки в узлах
- Когнитивные компьютерные системы

Долгосрочное видение – долгосрочный период

Общим явлением для всех тематик является новое развитие искусственного интеллекта и его повсеместное применение

Параллельно получают развития такие направления, как: «мягкие вычисления» (softcomputing), включающие нечеткую логику, нейронные сети, эволюционное моделирование и др.; эвристический поиск, экспертные системы; машинное зрение и машинное обучение; обработка текстов на естественном языке (Natural Language Processing); раскопки, или добыча, данных (Data Mining). Новые приложения такие как Text Mining, обобщение контента и другие

Среди многочисленных приложений - робототехника и Big Data, обработка массивов разнообразных данных, которые по своим размерам недоступны человеку. Важным свойством этого периода является переход, аналогичный переходу от детерминированной ньютоновской физики к релятивистской физике. На смену расчетам, занимавшим монопольное положение в ИКТ, приходят анализ и добыча данных.

Важной тенденцией станет краудсорсинг и вовлечение широкой экспертной аудитории в решение целого ряда задач. Это позволит провести категоризацию и обогащение данных силами широкого, неопределённого круга лиц, привлечённых на основании публичного заявления, без вступления в какие-либо трудовые отношения. Другим устойчивым явлением долгосрочного периода станет смешение и интеграция данных. Возникнет целый класс техник, позволяющих интегрировать разнородные данные из разнообразных источников для возможности глубокого анализа.

К примеру, цифровая обработка сигналов и обработка естественного языка (включая тональный анализ) в распределенной вычислительной среде, с обработкой большого количества разнообразной информации от самых разных источников, включая аудиофайлы.

Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)

- Создание объемного корпуса для исследования онтологических связей и зависимостей, позволяющую программам решать сложные задачи из области ИИ на основе логического вывода и привлечения здравого смысла
- Изучение модели когнитивной архитектуры.
- Создание систем, с помощью которого конструируются рефлексивные предположения о процессах человеческого мышления, то есть модели, описывающие различные виды человеческой деятельности (чтение текста, понимание языка, контроль полета).
- Моделирование головного мозга, в том числе, понимание работы человеческого мозга, в части накопления и обработки информации.

Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений

- Интегрированные экспертные системы класса PureSystems
- Широкое проникновение технологий искусственного интеллекта во все отрасли промышленности, государственного управления и социальной жизни.
- Автоматизация повседневных операций начального и среднего уровня
- Замещение части функций, выполняемых человеком, интеллектуальными операциями на основе искусственного интеллекта
- Интеграция различных архитектур обучения, целеполагания и сохранения знаний внутри единой системы для разработки систем, которые демонстрируют разумное поведение
- Создание систем общего назначения, способных к логическим рассуждениям.
- Умные здания, умные светофоры и другие городские технологии, объединенные в «Интернет вещей» благодаря технологиям беспроводного доступа

Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)

- Создание платформы, на которой разные группы ученых и инженеров смогут применять свои ИИ-алгоритмы, пользуясь единой базой данных о реальности и единой формой представления

- Максимальная интеграция человека в компьютерные сети в самом широком смысле.
Человек, с сенсорными датчиками и носимыми устройствами становится активным элементом сети
- Разработка новых систем обработки данных
 - аналитические СУБД, адаптированные к работе на масштабируемых кластерах новой архитектуры;
 - распределенные файловые системы;
 - нереляционные СУБД;
 - СУБД с поколоночным хранением
 - технологии для обработки сложных событий (Complex Event Processing, CEP).

Разработка новых технологических платформ для создания прикладных аналитических решений по Big Data

Появление платформ нового поколения, интегрированные с суперкомпьютерами и сетями нового поколения. Со скоростными методами обработки в узлах

- Когнитивные компьютерные системы
- Создание когнитивных вычислений и обработки данных, не использующих традиционного «неймановского» принципа организации вычислительного процесса
- Новое поколение нейроморфных компьютеров, то есть имитирующим деятельность мозга
- Развитие когнитивного подхода с сохранением существующего технологического базиса
- Обработка и интеграция больших потоков как структурированных, так и неструктурированных данных, порожденных датчиками и устройствами, образующих «Интернет вещей»
- Обработка и анализ массивов данных порожденных цифровой экономикой
- Модели и прототипы нового поколения суперкомпьютеров развитие новых способов обработки данных.
- Комплексный подход к обработке и анализу данных в сфере здравоохранения
- Применение технологических платформ в проактивной медицине, с использованием геномного секвенирования
- Интегрированные сети городов будущего и сети обмена данными, предоставляющие предоставлять различную информацию об окружающей среде.

5. Целевые индикаторы

Целевые индикаторы

- Количество R&D лабораторий - 10 единиц;
- Перевод 30% существующих и новых данных на технологии Big Data для ГМК и нефтегазовой отрасли
- Учеба и тренинг ведущих специалистов за рубежом, участие в зарубежных международных конференциях - до 200 специалистов ежегодно
- Внедрение в учебные программы специализации по вопросам базовых прикладных технологий - по 2 специальности в 10 Вузах
- Подготовка кадров по программе стажировка в области базовых прикладных технологий - до 30 стажеров ежегодно
- Подготовка кадров по программе зарубежная магистратура в области базовых прикладных технологий - до 150 магистров каждые два года
- Подготовка кадров по программе зарубежное РНД в области базовых прикладных технологий - до 50 докторантов каждые три года

- Количество специализированных курсов в сфере базовых прикладных технологий, преподаваемых в ВУЗах РК - 26 единиц;
- Количество специализированных курсов по тематикам под-направления - 40 единиц;
- Количество собственных дата-центров, серверных комнат, серверного оборудования государственных органов в 2017 году должно быть сокращено на 80% по сравнению с 2012 годом, в 2020 году - на 50 %;
- Сокращение затрат государственных органов на закуп собственного лицензионного программного обеспечения, услуг создания и обслуживания собственных информационных систем в 2017 году должно составить 40 % по сравнению с 2012 годом, в 2020 году - 50 %;

6. Научно-технологические разработки в отрасли (базовые технологии)

В отрасли информационных технологий к числу базовых относятся следующие:

- Облачные технологии
- Мобильные технологии
- Телекоммуникации
- Системы управления и менеджмента данных и информационными системами
- Информационная безопасность
- Алгоритмы и методы вычислений
- Технологии Big Data
 - Платформы типа Hadoop, архитектура распределённых данных (HDFS)
 - Кластерные вычисления с использованием Map Reduce (например Mahout)

К этим инфраструктурным технологиям примыкают технологии системного анализа

- Теория принятия решения (Decision Support)
 - Графические сети (Байесовские сети)
 - Нейронные сети
 - Генетические алгоритмы
 - Метод анализа иерархий
- Теория распознавания образов
- Математическая лингвистика
- Статистические и методы оптимизации (линейной и нелинейной)
- Прикладные системы разработок искусственного интеллекта (класса Prolog)
- Математические модели и методы распознавания образов
- Методы системного структурно-функционального анализа
- Методы синтеза речи
- Теория игр
- Семантический анализ.
- Морфологический анализ
- Стохастический анализ
- Теория решения изобретательских задач (ТРИЗ)
- Исследование операций

На основе этих базовых технологий строятся прикладные технологии
К ним относятся

- Разработка и использование алгоритмов компьютерного зрения
- Исследования по разработке новых моделей эффективного управления и прогнозирование рисков
- Исследования по разработке архитектуры коллективного интеллекта и эффективный краудсорсинг включая новые методы организации сетевой экспертизы
- Исследования по разработке прикладного математического моделирования с неполными и гетерогенными параметрами
- Исследования по разработке новых математических методов моделирования, распознавания закономерностей и классификации сложных и стохастических процессов
- Исследования по прикладному применению существующих методов искусственного интеллекта
- Исследования по разработке интерфейсов для естественного взаимодействия и общения
- Архитектура межмашинного взаимодействия
- Исследования и разработка новых методов для параллелизации сложных вычислений для обработки больших объемов данных включая усовершенствование алгоритмов на основе MapReduce
- Исследования в области визуализации большого количества данных для аналитики
- Новые модели и методы сжатия и обобщения неструктурированных данных
- Исследования по разработке новых моделей и организации коллаборативной деятельности и краудсорсинга (crowdsourcing)
- Исследования по разработке новых методов синхронных и асинхронных коммуникаций

Все вышеуказанные технологии имеют явно выраженный прикладной характер

На их основе разрабатываются и формируются различные приложения в различных отраслях науки, индустрии и социальной жизни

- Биомедицина и фармакология
- Медицинские технологии
- Цифровая экономика
- Робототехника
- Мобильный мир
- Интернет вещей
- Новая энергетики
- Интеллектуальные транспортные системы
- Нанотехнологии и новые материалы
- Умный город
- Потребительский рынок
- Носимые сенсорные датчики
- Smart R&D в ИКТ
- Конвергенция ИКТ с другими отраслями

Для получения указанных продуктов были выделены приоритетные направления научных исследований. Исследования по ним, наряду с другими исследованиями и технологиями, позволят достичь практического результата и конкретных прикладных решений, на основе казахстанских научных исследований и разработок.

В настоящее время уровень публикаций и компетенций казахстанских ученых и инженеров в вышеперечисленных областях невелик. Тем не менее. Высокий уровень казахстанской прикладной математической школы, кадры подготовленные по программе

«Болашак», современные системы ИТ образования позволяют надеяться на получение хорошего результата.

Тематика

«Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)»

Основные исследования в Казахстане и в мире

У данной тематики научных исследований высокий уровень готовности к применению в самых различных отраслях экономики, социальной сферы и жизни человека. Особенность всех языков, как самой универсальной модели общения в обществе обеспечит высокую применимость предложенной тематики. Также значительный уровень патентоспособности исследований, потому что они ведутся в направлениях, которые являются новыми, в целом в мировой научной практике и в Казахстан

Схожие казахстанские исследования

- Нейросетевой метод семантического разбора текста в глобальной информационной сети (на примере казахского языка),
Институт проблем информатики и управления
- Разработка математической модели распознавания рукописного казахского текста и ее программная реализация.
Евразийский национальный университет им. Л.Н.Гумилева
- Отображение алгоритмов семантической обработки информации на архитектурах компьютеров с массовым параллелизмом,
Евразийский национальный университет им. .Н.Гумилева
- Разработка эффективных технологии компьютерного перевода казахского языка на английский и русский языки (и обратно) на основе методов формальных грамматик и статистических методов,
Республиканское государственное предприятие "Казахский национальный университет имени Аль-Фараби" Дочернее государственное предприятие "Научно-исследовательский институт математики и механики"
- Разработка системы смыслового анализа поиска текстов нового поколения ориентированной на казахский язык,
ДГП "Институт проблем информатики и управления" РГП "Институт математики, информатики и механики" КН МОН РК
- Автоматизация распознавания и порождения письменной и устной речи казахского языка, Евразийский национальный университет имени Л.Н.Гумилева
- Разработка информационной системы "Антиплагиат" для проверки текстовых материалов на казахском языке на уникальность (на полное соответствие и парафразы),
ПГУ им. С. Торайгырова
- Создание прототипа интеллектуальной голосовой системы с интерактивным диалоговым режимом на государственном языке
ЧУ "Центр энергетических исследований"
- Разработка защищенной операционной системы с поддержкой казахского языка на основе Linux-платформ
Республиканское государственное предприятие "Казахский национальный университет имени Аль-Фараби" Дочернее государственное предприятие "Научно-исследовательский институт математики и механики"

Мировые исследования

Мировые исследования в этом направлении отличаются чрезвычайной активностью.

Необходимо отметить, что ведущие корпорации ИТ индустрии продолжают разрабатывать собственные технологии автоматического распознавания и перевода текстов и речи.

Подготовленные алгоритмы семантического и синтаксического анализа могут использоваться и для других приложений распознавания образов в целом.

Тематика

«Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полуавтоматической поддержки решений»

Основные исследования в Казахстане и в мире

- Разработка ИС поддержки принятия решений при формировании эффективных технологических схем стабилизации качества полезного ископаемого, Рудненский индустриальный институт
- Разработка стохастических моделей, методов и алгоритмов распределения и размещения ресурсов и объектов
Казахский национальный технический университет имени К.И. Сатпаева
- Разработка моделей, методов и алгоритмов статистического анализа и синтеза, динамических частотно-импульсных систем автоматического управления объектами с запаздыванием,
Казахский национальный технический университет имени К.И. Сатпаева
- Разработка методов и алгоритмов для оптимизационных задач в интеллектуальных системах автоматического контроля и управления
Научно-исследовательский институт математики и механики, Казахский национальный университет им. Аль-Фараби
- Разработка интеллектуальных динамических систем прогнозирования и управления сложными объектами,
ДГП "Институт проблем информатики и управления" РГП "Институт математики, информатики и механики" КН МОН РК
- Разработка и исследование логических методов моделирования поведения сложных систем,
ДГП "Институт проблем информатики и управления" РГП "Институт математики, информатики и механики" КН МОН РК
- Методы и математические модели оценки научно-инновационных проектов,
РГП "Казахский Национальный университет имени Аль-Фараби"
- «Создание математической модели «риск-анализа» возникновения аварийных ситуаций на магистральных трубопроводах с использованием геоинформационных технологий и математических методов распознавания и классификации».
НЦ космических исследований и технологий
- Совершенствование методологии и разработка инструмента технологии моделирования и мониторинга состояний сложных систем ТОФИ
Компания системных исследований «Фактор»
- Разработка ИС поддержки принятия решений при формировании эффективных технологических схем стабилизации качества полезного ископаемого
Рудненский индустриальный институт

Тематики, связанные с использованием технологии Bigdata

- Тематика «Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)»

- Тематика «Разработка новых технологических платформ для создания прикладных аналитических решений по BigData»

Схожие казахстанские исследования

- Организация распределенных и облачных вычислений для решения ресурсоемких задач
- РГП "Казахский Национальный университет им. Аль-Фараби" ДГП "НИИ математики и механики"
- Разработка моделей и приложений высокопроизводительной распределенной обработки данных на основе технологии MapReduce – Hadoop для задач нефтедобычи
- Научно-исследовательский институт математики и механики
- Исследование механизмов обнаружения экстремистского контента в закрытой части сети Интернет на основе методов интеллектуального анализа Data Mining
- Восточно-казахстанский государственный технический университет им. Д. Серикбаева
- Инновационный прототип системы наиболее надежного хранения данных на базе технологии распределенного хранения
- Казахский гуманитарно-юридический инновационный университет

Мировой рынок

Мировой рынок по вышеперечисленным тематикам относится к одному из самых быстроразвивающихся

По прогнозам ведущих аналитических групп, суммируя мнения множества аналитиков ИТ-индустрии, то общими технологическими тенденциями наступающей эпохи являются следующее. (<http://www.gartner.com/newsroom/id/2209615>)

К 2020 году, отрасль ИКТ достигнет объемов финансирования свыше \$5 триллионов — что на \$1.7 триллионов больше, чем сегодня. По крайней мере 70% роста будет обеспечено технологиями так называемой Третьей Платформы.

Под термином «Третья платформа», IDC называют следующую связку трендов и технологий, значительно трансформирующих отрасль:

- облачные сервисы,
- мобильность,
- социальные сети
- большие данные/аналитика

Также, по мнению IDC, которое разделяет все ведущие мировые компании ИТ отрасли, ожидается взрывной рост новых решений на базе указанных направлений.

По сути, уже в 2012 году началась борьба за лидерство на рынке 2020 года.

Эти мировые тренды, поддержанные основными участниками рынка и правительствами многих стран заметно изменяют как структуру рынка, так и будущий ландшафт ИК технологий и казахстанскому сектору ИКТ необходимо участвовать в этом общемировом направлении.

7. Стратегии и программы развития направления

К основным стратегическим и программным документам государственного управления и регулирования ИКТ Казахстана, утвержденные Президентом и Правительством РК, относятся следующие:

- Стратегия развития Республики Казахстан до 2050 года;
- Стратегия развития Республики Казахстан до 2030 года;

- Стратегический план развития Республики Казахстан до 2020 года;
- Государственная программа форсированного индустриально-инновационного развития на 2010 - 2014 годы;
- Государственная программа «Информационный Казахстан -2030»;
- Закон Республики Казахстан «Об информатизации» (Проект)

К основным стратегическим и программным документам государственного управления и регулирования ИКТ Казахстана, утвержденные Президентом и Правительством РК, относятся следующие:

- Стратегия развития Республики Казахстан до 2050 года;
- Стратегия развития Республики Казахстан до 2030 года;
- Стратегический план развития Республики Казахстан до 2020 года;
- Государственная программа форсированного индустриально-инновационного развития на 2010 - 2014 годы;
- Государственная программа «Информационный Казахстан -2030»;

8. Идентификация рынков

Первая очередь:

- Государственные и квазигосударственные организации
- Крупный бизнес
- Малый и средний бизнес

Вторая очередь:

- Страны Таможенного союза
- Ближнее и дальнее зарубежье

9. Основные акторы

Отечественные университеты и научно-исследовательские

- Карагандинский государственный университет имени академика Е.А. Букетова
- Карагандинский государственный технический университет
- Казахский национальный технический университет имени К.И. Сатпаева
- Алматинский университет энергетики и связи
- Евразийский национальный университет им. Л.Н.Гумилева
- Казахстанско-Британский Технический университет
- Казахский национальный университет имени Аль-Фараби
- Восточно-Казахстанский государственный технический университет им. Д. Серикбаева
- Рудненский индустриальный институт
- Южно Казахстанский Государственный Университет имени М. Ауезова
- Казахский агротехнический университет им. С. Сейфуллина
- Таразский государственный Университет им. М.Х. Дулати
- Международный университет информационных технологий
- Каспийский государственный университет технологии и инжиниринга им. Ш. Есенова
- Университет имени Сулеймана Демиреля
- Nazarbayev University Research and Innovation System, Назарбаев Университет
- Кокшетауский государственный университет им. Ш. Уалиханова
- Костанайский государственный университет им. А. Байтурсынова
- Казахский научно-исследовательский институт водного хозяйства

- Северо-Казахстанский Государственный Университет

НИИ

- Институт проблем информатики и управления
- Казчерметавтоматика
- Институт ядерной физики, Национальный ядерный центр Республики Казахстан
- Научно-исследовательский институт транспорта и коммуникаций
- РГП "Институт математики, информатики и механики"
- Институт математики и математического моделирования
- Национальная компания «Казахстан Ғарыш Сапары»
- Национальный центр космических исследований и технологий
-

Государственные органы

- МОН РК
- МТК РК
- МИНТ РК
- МНГ РК

Холдинги и национальные компании

- АО «Самрук-Казына»
- АО «КМГ»
- АО «Зерде»
- АО «Казахтелеком»
- АО «КТЖ»
- РГП «ГТС»

10. Необходимые ресурсы для реализации Дорожной карты

- Финансовые ресурсы:
 - Инвесторы (венчурные фонды, бизнес-ангелы, краудфандинг)
 - Льготное кредитование финансовыми институтами
 - Государственно-частное партнерство на основе отраслевых фондов развития;
 - Программы поддержки и акселерации технологических стартапов
 - Грантовые программы поддержки рискованных исследований
 - 1% отчисления предприятий на R&D
- Ведущие НИИ и ВУЗы:
 - Центр исследовательских лабораторий при АО НИТ;
 - НИИ и ВУЗы (МУИТ, НУ, КазНУ, ЕНУ и др.);
- Кадровый потенциал:
 - Разработка и совершенствование образовательных программ бакалавриата, магистратуры и докторантуры, увеличение количества грантов (стипендия “Болашак”, институты Fraunhofer, ETRI, Инновационный центр Сколково, Исследовательский парк Stanford);
 - Развитие стандартизации, сертификации CMMI (Capability Maturity Model) и PMI (Project Management Institute) и специализированных программ с международной аккредитацией (включая Microsoft, CISCO, ORACLE)
 - Решающим условием успешного выполнения Дорожной карты является высококвалифицированный кадровый состав. За весь период действия Дорожной карты необходимо подготовить не менее 1000 высококвалифицированных специалистов в различных областях ИКТ.

- Основные специальности
- комплексный анализ данных (data scientist)
- теория и методы искусственного интеллекта и принятия решений
- биоинформатика и компьютерная геномика
- системный анализ и исследование операций
- статистика, предсказательная статистика
- геоинформатика и геоинформационные системы, включая космическую геодезию
- робототехника

- Инфраструктура:
 - Развитие программно-аппаратный комплекса для разработки программных решений и среды предпроизводственного тестирования (тестовая площадка) при Центр исследовательских лабораторий АО НИТ;
- Трансферт технологий

11. Риски и ограничения

Риски, возникающие при реализации Дорожной Карты: научно-технологические, социально-экономические и политические.

Наименование
Научно-технологические риски
Отсутствие системного подхода к технологическому развитию базовых прикладных отраслей
Слабая заинтересованность предприятий во внедрении новых технологий Дорожной Карты
Утечка высококвалифицированных специалистов
Социально-экономические и финансовые риски
Снижение доли расходов на науку в ВВП.
Недостаток финансовых средств на НИОКР по Дорожной Карте
Низкие затраты бизнеса на науку (преимущественная ориентация на адаптацию импортируемых технологий)
Неэффективность форм финансирования фундаментальной и прикладной науки
Слабая реальная защита прав интеллектуальной собственности авторские права
Риски нормативной базы и инфраструктуры
Риски, связанные с неэффективным управлением реализацией Дорожной Карты
Риски негативного отношения к реализации Дорожной карты со стороны представителей органов управления, общественности и СМИ
Риски некачественного проведения конкурса для участия в реализации Сценария и экспертизы результатов проведенного конкурса
Политические риски
Невостребованность научных достижений в экономике, патентование за рубежом.
Низкая эффективность мер государственной поддержки инновационной деятельности
Поступательная инновационная политика развитых стран
Несовершенство форм взаимодействия государства, промышленности и научных центров

12. Мониторинг реализации Дорожной карты

Контроль над ходом реализации Дорожной карты будет осуществлять Комитет науки МОН РК непосредственно, а также через посредство Национального центра научно-технической экспертизы (далее – Центр). Для более эффективного управления программой Дорожной карты предлагается в рамках Центра создать офис управления Дорожными картами, в функцию которого будет входить обязанность организации и контроля выполнения всех запланированных этапов, в том числе выполнения плана управления рисками программы.

Создание Проектного офиса позволит проводить детальный мониторинг выполнения Дорожных карт, включая анализ отклонений от плановых показателей. Кроме того. Создание Проектного офиса позволит упорядочить организационно-документарный аспект управления научным процессом.

13. План мероприятий по реализации Дорожной карты

№ п/п	Наименование мероприятия	Ожидаемый результат	Сроки исполнения	Ответственные исполнители
Нормативно-правовое обеспечение Дорожной Карты				
1	Разработка и обсуждение Дорожной карты по базовым прикладным технологиям	Проект Дорожной карты	I квартал 2014 года	Национальный центр научно-технической экспертизы
2	Исследование современного состояния уровня прикладных технологий в отраслях экономики и их потенциал на трансферт новых технологий	Аналитический отчет и обсуждение на межведомственной конференции	II квартал 2014 года	Комитет науки МОН РК МТК РК
3	Изучение международного опыта и международных регламентов и стандартов	Аналитический отчет и обсуждение на межведомственной конференции	II квартал 2014 года	Комитет науки МОН РК МТК РК
4	Подготовка необходимых национальных законодательно-нормативных актов по базовым прикладным технологиям, технических стандартов, единых требований, адаптация международных стандартов и требований	Проекты законодательных документов	IV квартал 2020 года	Комитет науки МОН РК МТК РК

Методическое обеспечение Дорожной Карты				
5	Учеба и тренинг ведущих специалистов за рубежом, участие в зарубежных международных конференциях (6-7 командировок), приглашение ведущих зарубежных ученых для проведения консультаций (2-3 приезда)	200 – специалистов ежегодно	IV квартал 2030 года	Комитет науки МОН РК МОН РК МТК РК
6	Разработка авторских курсов в тематикам под-направления	Учебно-методические комплексы	IV квартал 2030 года	МТК МОН РК
7	Внедрение в учебные программы специализации по вопросам базовых прикладных технологий	По 2 специальности в 10 Вузах	IV квартал 2030 года	МОН РК
8	Подготовка кадров по программе стажировка	до 30 стажеров ежегодно	IV квартал 2030 года	МОН РК
9	Подготовка кадров в зарубежной магистратуре	до 150 магистров каждые два года по под-направлению	IV квартал 2030 года	МОН РК
10	Подготовка кадров по программе PhD за рубежом	до 50 докторантов каждые три года	IV квартал 2030 года	МОН РК
Институциональное обеспечение Дорожной Карты				
11	Модернизация структуры и функций технического комитета по стандартизации и сертификации по тематикам под-направления	Технический комитет	IV квартал 2019 года	МТК РК МИНТ РК
12	Создание R&D на базе существующих ВУЗов и НИИ	до 10 лабораторий	IV квартал 2030 года	МТК РК МИНТ РК
13	Создание центра обучения и сертификации по тематикам под-направления	до 10 центров	IV квартал 2030 года	МТК РК МИНТ РК
14	НИР и НИОКР в рамках Дорожной Карты	Заключительные научные отчеты, 5 патентов РК, 10 научных публикаций в высокорейтинговых международных журналах.	IV квартал 2030 года	МОН РК МТК РК

ПАСПОРТ ТЕМАТИКИ НАУЧНОГО ИССЛЕДОВАНИЯ

	Классификационный код тематики научного исследования	
Наименование тематики научного исследования	Исследования по автоматическому анализу текста, распознаванию и генерации речи (с фокусом на казахский язык)	
<u>Сущность исследования, уровень предлагаемых решения/ожидаемых результатов, масштабы применимости ожидаемых результатов</u>		
<p>Применение моделей и методов искусственного интеллекта для распознавания текстов и речи, с автоматическим переводом на разные языки. Исследование состоит из нескольких взаимосвязанных компонентов – распознавания речи (текстов), перевод текста и речи на основные мировые языки, включая казахский, реализация в виде программно-аппаратных комплексов и компьютерных программ, использование облачных технологий.</p> <p>Необходимо использование передовых методов семантического и морфологического анализа (распознавания) с речи и текстов, применение облачных технологий хранения и распределенных вычислений и обработки Bigdata. Математические исследования по формализации структурного (морфологического) анализа, сжатия информации, применения математических методов обработки цифровых и аналоговых сигналов, выделения устойчивых последовательностей и построение морфологической интерпретации.</p> <p>Результаты исследований применяются для всех направлений и отраслей экономики и социальной жизни, где необходим скоростной перевод текста и речи для лингвистического обеспечения. Это касается прежде всего сферы образования, перевода художественной и технической литературы, использования опыта и знаний иностранных консультантов и инвесторов в Казахстане, распространения казахстанских технологий</p>		
<u>Уровень научной новизны и значимости</u>		
<p>Исследования предполагается высокий уровень новизны научных и опытно-конструкторских результатов. Предполагается использование самых современных моделей и методов математической лингвистики, распознавания образов, речевого синтеза..</p> <p>Значимость проекта заключается в вовлечении и доступе всех слоев общества, включая жителей удаленных населенных пунктов и не владеющих основными мировыми языками к мировым базам знаний, библиотекам и огромным сетевым интернет ресурсам. Приоритет и фокус на лингвистические и фонетические особенности казахского языка обеспечат вовлечение казахского языка в мировое лингвистическое пространство.</p> <p>Актуальность и необходимость качественного перевода и синтеза казахского языка и речи на казахском языке особенно велика в связи с широкой интеграции нашей страны в мировое экономическое и культурное пространство.</p> <p>Отмечается высокая важность и актуальность исследований в данном направлении. уровень научной новизны и значимость направления высокие</p>		
<u>Прикладная важность исследования (в том числе патентоспособность и возможности для коммерциализации результатов)</u>		
<p>У данного тематики научных исследований высокий уровень готовности к применению в самых различных отраслях экономики, социальной сферы и жизни человека. Особенность всех языков, как самой универсальной модели общения в обществе обеспечит высокую применимость предложенной тематики. Также значительный уровень патентоспособности исследований, потому что они ведутся в направлениях, которые являются новыми, в целом в мировой научной практике и в Казахстан</p>		

<u>Дескриптор и технические характеристики (в том числе индикаторы ожидаемой завершающей стадии исследований)</u>		
<p>Исследования предполагают применение новых и эффективных моделей и алгоритмов. В том числе, специализированные алгоритмы синтаксического, морфологического и семантического анализа позволят проводить скоростное распознавание текстов и речи в реальном режиме времени.</p> <p>С учетом активности исследований и публикаций в этом направлении, ожидаемое завершение исследований до уровня проектной и технологической готовности высокое.</p>		
<u>Базовые технологии</u>		
Графематический анализ	Морфологический анализ	Семантический анализ, с построением семантических графов
Прикладные системы разработок искусственного интеллекта (класса Prolog)	Математические модели и методы распознавания образов	Методы оптимизации
Методы системного структурно-функционального анализа	Методы синтеза речи	Облачные технологии
Уровень разработок по теме исследования (%), либо граница (в годах) в сравнении с наиболее развитой в данной сфере страной		Возможность самостоятельной разработки (высокая, средняя, низкая)
50		высокая
Профессиональные научно-исследовательские группы в области	Отечественные	<ul style="list-style-type: none"> • Евразийский национальный университет им.Л.Н.Гумилева • Институт математики и математического моделирования • Научно-исследовательский институт математики и механики, • Казахский национальный университет им. аль-Фараби • Университет имени Сулеймана Демиреля • Институт проблем информатики и управления
	Международные	Большое количество университетов, государственных и корпоративных исследовательских лабораторий.

ПАСПОРТ ТЕМАТИКИ НАУЧНОГО ИССЛЕДОВАНИЯ

	Классификационный код тематики научного исследования	
Наименование тематики научного исследования	Исследования новых методов сбора, анализа, классификации и обобщения информации для автоматической или полу-автоматической поддержки решений	
<u>Сущность исследования, уровень предлагаемых решения/ожидаемых результатов, масштабы применимости ожидаемых результатов</u>		
<p>Исследования направлены на разработку современных методов принятия решения с использованием методов искусственного интеллекта. Они являются основой экспертных систем широкого профиля. Практически во всех отраслях экономики и социальной жизни существует проблемы обоснованного выбора вариантов, на основе сложившихся процедур принятия решения. Выбор решения всегда производится в условиях неопределенности. В целом, исследование относится к классу систем поддержки принятия решений и интеллектуальных систем (Decision Support and Inrelligent Systems) и включает в себя обычные и групповые системы поддержки принятия решений, прикладные методы искусственного интеллекта, экспертные системы и предиктивную аналитику. Применение результатов данной научной тематики позволит автоматизировать процесс принятия решения в различных отраслях, в ситуациях связанных с принятием решения из многих вариантов развития. Математические основы и компьютерные технологии позволят создать оболочку и среду функционирования при которых выбор является обоснование и доказательным. Применяются различные методы многокритериального анализа, оптимизационный алгоритмы, ускоряющих проедуру приянтия решения. При этом, часто, применяются технологии с участием человека-эксперта. Участвующего в процессе обоснованного и согласованного выбора.</p> <p>Результат исследования применим во всех отраслях экономики. Особенно актуальным он является при принятии решения в сложных динамических системах нефтегазовой промышленности, обоснования геологических и геофизических исследований, выбор инвестиционных проектов, обоснование выбора технологий.</p>		
<u>Уровень научной новизны и значимости</u>		
<p>Исследования предполагают применение всех современных научно-инженерных подходов по теории выбора обоснованного решения. Предполагается создание программного комплекса, на основе математических моделей принятия решения.</p> <p>В связи с развитием самих компьютерных технологий, вычислительных мощностей и средств отображения, возрастает значимость системы для применения в самых отраслях человеческой деятельности. Вопросы обоснованного и квалифицированного выбора решений (вариантов) в сложных инженерных и управленческих задачах являются актуальными</p> <p>Отмечается высокая важность и актуальность исследований в данном направлении. уровень научной новизны и значимость направления высокие</p>		
<u>Прикладная важность исследования (в том числе патентоспособность и возможности для коммерциализации результатов)</u>		

Исследования относятся к классу с высоким уровнем коммерциализации и применимости к использованию в самых разных отраслях экономики. Прикладное применение очень высокое. С каждым годом повышается требования к уровню и качеству принимаемых решений. Высока ответственность за неверные решения, особенно при принятии финансовых, инвестиционных решений, вопросов выделения мер государственной поддержки и инновационных грантов. Выбор правильного комплекта оборудования, принятие решение по освоению месторождений, автоматизированный выбор оптимальных маршрутов, поддержка переговорных процессов – эти и многие другие задачи будут использовать модели, методы и алгоритмы данного исследования

Дескриптор и технические характеристики (в том числе индикаторы ожидаемой завершающей стадии исследований)

Формализация неопределенности и субъективности в математических моделях, использование субъективной вероятности, нечетких множеств, многокритериальных функций предпочтения, нейронных сетей, генетических алгоритмов и другие методы, Математическое и алгоритмическое обеспечение генерации вариантов управленческих решений, методы и алгоритмы ранжирования вариантов решений, математическое и алгоритмическое обеспечение поддержки проектного анализа в отраслях промышленности и финансовой сфере
Создание программных комплексов высокой готовности для использования в промышленности

Базовые технологии

Многокритериальный анализ	Исследование операций	Теория голосования и предпочтений
Методы моделирования	Системы алгебраических и дифференциальных уравнений	Семантический анализ
Проектный анализ	Системы программирования	Обработка неструктурированных данных

Уровень разработок по теме исследования (%), либо граница (в годах) в сравнении с наиболее развитой в данной сфере страной

Возможность самостоятельной разработки (высокая, средняя, низкая)

30%

высокая

Профессиональные научно-исследовательские группы в области

Отечественные

- Казахстанско-Британский технический университет
- Институт геологических наук им. К.И. Сатпаева,
- Национальный научно-технологический холдинг "Парасат"
- Евразийский национальный университет им.Л.Н.Гумилева
- Национальный центр

		<p>космических исследований и технологий</p> <ul style="list-style-type: none"> • Институт математики и математического моделирования • Научно-исследовательский институт математики и механики, • Казахский национальный университет им. аль-Фараби • Назарбаев Университет
	Международные	<p>Большое количество университетов, государственных и корпоративных исследовательских лабораторий.</p>

ПАСПОРТ ТЕМАТИКИ НАУЧНОГО ИССЛЕДОВАНИЯ

	Классификационный код тематики научного исследования	
Наименование тематики научного исследования	Разработка новых алгоритмов и абстракций для оптимизации вычислений, хранения и вывода данных на основе распределенной архитектуры файлов (HDFS)	
<p><u>Сущность исследования, уровень предлагаемых решения/ожидаемых результатов, масштабы применимости ожидаемых результатов</u></p> <p>Сущность исследования заключается в использовании большие вычислительных и аналитических возможностей технологии BigData (больших данных) для использования в различных сферах промышленности, науки и социальной жизни, использующих обработку больших массивов неструктурированной и слабосвязанной информации.</p> <p>Ранее бизнес-анализ строился на основе структурированных данных, но с появлением проблемы Больших Данных обострилась необходимость в работе с неструктурированными данными, составляющими свыше 80% всех хранимых данных, причем их количество нарастает на порядок быстрее по сравнению со структурированными данными.</p> <p>Особенно это касается обработки геофизических данных, данных аэрокосмического мониторинга, биотехнологий, селекционной работы, обработки социальных сетей, семантического анализа текстов и многих других.</p> <p>Уровень предлагаемых решений достаточно высок, потому что исследование находится в русле мировых трендов и имеет большую значимость в части прикладных применений.</p> <p>Масштаб приложений охватывает практически все отрасли экономики, так как практически везде в настоящее время используется интенсивная и распределенная обработка данных.</p>		
<p><u>Уровень научной новизны и значимости</u></p> <p>Исследования предполагается высокий уровень новизны научных и опытно-конструкторских результатов. Предполагается использование самых современных моделей и методов математической обработки данных, моделей распределенной (сетевой) асинхронной обработки данных, оптимальное управление процессами обработки в реальном режиме времени на большом количестве компьютерных узлов (процессоров).</p> <p>Отмечается высокая важность и актуальность исследований в данном направлении. уровень научной новизны и значимость направления высокие</p>		
<p><u>Прикладная важность исследования (в том числе патентоспособность и возможности для коммерциализации результатов)</u></p> <p>У данного тематики научных исследований высокий уровень готовности к применению в самых различных отраслях экономики, социальной сферы и жизни человека, требующих интенсивной и скоростной обработки разнообразных данных, в распределенном режиме..</p> <p>Также значительный уровень патентоспособности исследований, потому что они ведутся в направлениях, которые являются новыми, в целом в мировой научной практике и в Казахстан</p>		

Дескриптор и технические характеристики (в том числе индикаторы ожидаемой завершающей стадии исследований)

Исследования предполагают применение новых и эффективных моделей и алгоритмов. В том числе, специализированные алгоритмы управления и синхронизации процессов обработки данных, применения технологии Hadoop, MapReduce и других технологических решений.

Помимо этих методов возможно использование других новых методов распределенной обработки и логического синтеза распределенных данных.

Особенность исследования такова, что позволит казахстанским ученым и инженерам интегрироваться с мировыми исследовательскими центрами на кооперационной основе.

С учетом активности исследований и публикаций в этом направлении, ожидаемое завершение исследований до уровня проектной и технологической готовности высокое.

Базовые технологии

Технология Hadoop	Технология MapReduce	Методы стохастической оптимизации
Параллельные вычисления	Системы управления компьютерными сетями	Облачные технологии
Семантический анализ	Технологии сжатия (компрессии) данных	Обработка неструктурированной информации

Уровень разработок по теме исследования (%), либо граница (в годах) в сравнении с наиболее развитой в данной сфере страной

Возможность самостоятельной разработки (высокая, средняя, низкая)

20

средняя

Профессиональные научно-исследовательские группы в области

Отечественные

- Евразийский национальный университет им.Л.Н.Гумилева
- Институт математики и математического моделирования
- Научно-исследовательский институт математики и механики,
- Казахский национальный университет им. аль-Фараби
- Университет имени Сулеймана Демиреля
- Институт проблем информатики и управления

Международные

Большое количество университетов, государственных и корпоративных исследовательских лабораторий.

ПАСПОРТ ТЕМАТИКИ НАУЧНОГО ИССЛЕДОВАНИЯ

	Классификационный код тематики научного исследования	
Наименование тематики научного исследования	Разработка новых технологических платформ для создания прикладных аналитических решений по (Big Data)	
<u>Сущность исследования, уровень предлагаемых решения/ожидаемых результатов, масштабы применимости ожидаемых результатов</u>		
<p>В настоящее время нет таких полнофункциональных платформ, и под платформами для BigData понимают обычные традиционные компьютерные платформы, адаптированные для работы с большими объемами данных</p> <p>Новая технологическая платформа, полноценно решающая проблему BigData (Больших Данных), должна обеспечивать возможность работы с данными всех типов, представленных в произвольных форматах, а для этого она должна включать:</p> <ul style="list-style-type: none"> • средства визуализации и обнаружения — данные должны быть представлены пользователю в виде, удобном для восприятия, понимания содержания и поиска в различных источниках; • инструменты аналитики цифровых и текстовых данных в пакетном режиме и режиме реального времени; • системы обработки потоковых данных; • средства поддержки традиционных хранилищ данных и систем оперативной обработки транзакций. <p>Сегодня разработка и эксплуатация программного обеспечения невозможны без развитой платформы — среды, в которой работает ПО и под которую оно создается.</p> <p>С появлением облаков и сервисов PaaS (Platform as a Service) представление о платформах существенно расширилось, и сегодня можно говорить о различных категориях облачных платформ, а среди них есть и ориентированные на работу с BigData.</p> <p>Появление такой технологической платформы, в том числе, базирующейся на казахстанских решениях позволит значительно упростить и упорядочить работу по обработке информации и сконцентрироваться на аналитических расчетах решениях</p> <p>Уровень предлагаемых решений достаточно высок, потому что исследование находится в русле мировых трендов и имеет большую значимость в части прикладных применений.</p> <p>Масштаб приложений охватывает практически все отрасли экономики, так как практически везде в настоящее время используется интенсивная и распределенная обработка данных.</p>		
<u>Уровень научной новизны и значимости</u>		
<p>Исследования предполагается высокий уровень новизны научных и опытно-конструкторских результатов.</p> <p>Предполагается использование самых современных моделей и методов поиска скрытых закономерностей и знаний на основе визуального анализа.</p> <p>Результаты исследований используются в системы поддержки принятия решений и интеллектуальные системы (Decision Support and Intelligent Systems), в задачах искусственного интеллекта, экспертных системах, робототехнике и в других сферах, где необходим анализ данных и поиск скрытых закономерностей в них.</p> <p>Такое возможно на основе сочетания мощности компьютерных вычислений с аналитическими возможностями человека. Особенно это эффективно при использовании методов визуализации аналитических результатов.</p> <p>Отмечается высокая важность и актуальность исследований в данном направлении. уровень научной новизны и значимость направления высокие</p>		

Прикладная важность исследования (в том числе патентоспособность и возможности для коммерциализации результатов)

Отмечается высокая важность и актуальность исследований в данном направлении.

У данного тематики научных исследований высокий уровень готовности к применению в самых различных отраслях экономики, социальной сферы и жизни человека, требующих интенсивной и скоростной обработки разнообразных данных, в распределенном режиме.. Также значительный уровень патентоспособности исследований, потому что они ведутся в направлениях, которые являются новыми, в целом в мировой научной практике и в Казахстан

Дескриптор и технические характеристики (в том числе индикаторы ожидаемой завершающей стадии исследований)

- Платформы Hadoop/MapReduce
- In-Memory Data Grid
- Платформы с открытыми кодами (Проект Apache Ambari потенциально может стать основой для платформы BDAP с открытым кодом).
- Платформа Zettaset Orchestrator
- **IBM Platform Computing**
- Платформа Rock+

Исследования предполагают применение новых и эффективных моделей и алгоритмов.

Особенность исследования такова, что позволит казахстанским ученым и инженерам интегрироваться с мировыми исследовательскими центрами на кооперационной основе. С учетом активности исследований и публикаций в этом направлении, ожидаемое завершение исследований до уровня проектной и технологической готовности высокое

Базовые технологии

Облачные технологии и сервисы	Hadoop/MapReduce	реляционными и нереляционные СУБД
Поиск закономерностей в больших массивах данных (data mining)	BigData	Компьютерное моделирование

Уровень разработок по теме исследования (%), либо граница (в годах) в сравнении с наиболее развитой в данной сфере страной		Возможность самостоятельной разработки (высокая, средняя, низкая)
20		низкая
Профессиональные научно-исследовательские группы в области	Отечественные	
	Международные	Amazon Elastic Beanstalk, Heroku, EngineYard, Mendix, Google App Engine и Microsoft Azure IBM