

УДК 004.81

Антонов Сергей Георгиевич

доктор технических наук
ООО «Лингвистические
и информационные технологии»
г. Москва, Россия

Sergey Antonov

Doctor of Technical Sciences
«Linguistic and Information Technologies»
LLC
Moscow, Russia
Aserg2001@yandex.ru

Завалишин Даниил Евгеньевич

слушатель
Московский государственный
технический университет
им. Н. Э. Баумана
г. Москва, Россия

Daniil Zavalishin

Student
Bauman Moscow State
Technical University
Moscow, Russia
Zavalishin.DE@ya.ru

Крапивин Юрий Борисович

кандидат технических наук,
доцент кафедры теоретической
и прикладной лингвистики
Белорусский государственный
университет иностранных языков
г. Минск, Беларусь

Yury Krapivin

PhD in Technical Sciences,
Associate Professor of the Department
of Theoretical and Applied Linguistics
Belarusian State University
of Foreign Languages
Minsk, Belarus
ybox@list.ru

АНАЛИЗ SMS КАК СРЕДСТВА ЦЕЛЕВОЙ КОММУНИКАЦИИ

В статье предложено решение задачи автоматической обработки коротких текстовых сообщений, SMS сообщений, их анализа, основанное на лексико-грамматической модели представления текста и развитой лингвистической базе данных, что в совокупности обеспечивает качественное соотнесение выявляемых языковых признаков с классами речевых актов для пяти ситуаций коммуникации. Это позволяет определить, во-первых, в какой из них находятся корреспонденты и, во-вторых, присутствует ли в их сообщениях потенциальная угроза безопасности некоторому событию, объекту и т. п.

К л ю ч е в ы е с л о в а : *SMS коммуникация; анализ коротких текстов; лингвистическая база данных; типы коммуникации.*

ANALYSIS OF SMS AS A MEAN OF TARGETED COMMUNICATION

A solution to the problem of automatic processing and analysis of short text messages, SMS, based on a lexico-grammatical model of text representation and a comprehensive linguistic database is proposed. Together, these components ensure an accurate mapping of the identified linguistic features onto speech act classes for five communication situations and make it possible, first, to determine which of these situations the correspondents are in and, second, to identify whether their messages contain a potential security threat related to a particular event, object, or the like.

Key words: *SMS communication; short text analysis; linguistic database; types of communication.*

Неуклонный рост объемов передаваемой информации требует разработки новых, как по показателям их эффективности, так и по классу решаемых задач, методов и алгоритмов автоматической обработки текстовых данных, получаемых из разнообразных источников, в частности, таких как Интернет, сотовая связь.

К задачам автоматической обработки текстовых документов, представленных в электронном виде, можно отнести их классификацию, извлечение и последующий целевой анализ контента. При этом затруднения, возникающие при обработке некоторых классов документов, могут быть связаны с наличием в тексте различных типов информации и видов искажений.

Процесс коммуникации абонентов сотовых сетей предполагает обмен текстовыми документами специального типа – короткими текстовыми сообщениями, SMS (Short Message Service). Чем меньше объем текста сообщения, тем сложнее провести его автоматический анализ. К тому же тексты SMS, как правило, содержат много сокращений и искажений, нехарактерных для естественного языка. Зачастую авторы сообщений пренебрегают правописанием, правилами грамматики, используют разнообразные словесные конструкции и жаргонизмы, в особенности, если сами не являются носителями языка. Все это, безусловно, затрудняет автоматическую обработку и анализ коротких текстовых сообщений.

Элементы теории речевых действий в задачах обработки естественного языка

Как показывает проведенное исследование, при решении задач обработки коротких текстовых сообщений эффективным будет использование теории речевых действий. Это подход к формализации текста, основанный на представлении целей и мотивов участников речевого общения. Основной акцент в процессе общения делается на целях высказываний и на их интенсивности. Одним из главных компонентов речевого действия является иллокутивная цель, которая определяет коммуникативную цель высказывания и направлена на изменение какого-либо аспекта человеческого сознания: знания, намерения либо представления о текущей ситуации посредством композиции сообщений, заявлений и побуждений [1; 2].

С целью анализа достижения иллокутивной цели и в соответствии с ее видами выделяют классы речевых действий, например: репрезентативы, директивы, комиссивы, декларативы, экспрессивы, интеррогативы, перформативы [3; 4].

Инструментальные средства анализа текстовых сообщений

К числу наиболее распространенных программных средств базовой лингвистической обработки текста относятся морфологические анализаторы и многофункциональные программные библиотеки, обеспечивающие проведение важных с точки зрения автоматизации лингвистического анализа

этапов: установления границ слов и предложений, лемматизации, определения частей речи и, в ряде случаев, синтаксического анализа. Для русского языка интерес представляют MyStem, pymorphy2, UDPipe, Stanza [5; 6; 7; 8].

С целью устранения возможных ошибок ввода, а также случаев вариативности в написании и учета ситуаций использования нетекстовых элементов, которые непосредственно влияют на результаты морфологической разметки и последующей классификации, применяются алгоритмические или инструментальные средства и приемы автоматического исправления опечаток, нечеткого сопоставления словоформ, а также выделения специальных знаковых комплексов: телефонных номеров, адресов электронной почты, url-адресов, кодов, цифровых последовательностей и эмодзи. Однако для коротких текстовых сообщений такие элементы не являются исключительно служебными, поскольку могут выступать самостоятельными носителями значимой информации и, следовательно, должны учитываться при интерпретации коммуникативных целей.

В задачах классификации текстов часто используются программные библиотеки scikit-learn, fastText, Transformers, DeepPavlov [9; 10; 11; 12], с помощью которых создаются классификаторы на основе анализа символьных и словарных n-грамм, маркеров частеречной разметки, морфологических признаков и других формализуемых характеристик текста. К особенностям применения решений, построенных с их помощью, можно отнести необходимость обучения классификатора на предварительно подготовленном и размеченном репрезентативном обучающем корпусе текстовых сообщений, дополняемом при появлении новых словоформ, жаргонизмов и нестандартных синтаксических конструкций, а использование нейросетевых моделей, например, ruBERT [13], хоть и позволяет учитывать вариативность формулировок и выявлять скрытые семантические связи в текстовых сообщениях, что важно для решения целевой задачи, требует представительных объемов обучающих данных и не всегда обеспечивает требуемый уровень интерпретации получаемых результатов.

Таким образом, учитывая характерные особенности коротких текстовых сообщений, таких, как малый объем текста, ограниченность контекста, высокая вариативность графического оформления, наличие орфографических искажений, жаргонной лексики, а также смешение алфавитов наряду с включением нетекстовых элементов и связанных с ними ограничений в возможности применения в полном объеме стандартных средств обработки естественного языка, лексико-грамматическая модель анализа коротких текстовых сообщений, опирающаяся на современные средства базовой лингвистической обработки текста и развитую лингвистическую базу данных, способна обеспечить прозрачное соотнесение наблюдаемых языковых признаков с классами речевых актов.

Лингвистическая база данных

С целью автоматического лингвистического анализа коротких текстовых сообщений в контексте решаемой задачи в состав Лингвистической базы данных включен разработанный лексико-грамматический классификатор свойств русского языка и набор словарей, причем каждому слову из соответствующего словаря поставлен в соответствие, характеризующий его набор значений лексико-грамматических кодов классификатора.

Таким образом, помимо прочего, компонентами Лингвистической базы данных стали:

- Словарь правильных слов. Содержит общеупотребительные правильные (безошибочные) слова, используемые в различных текстах естественного языка.

- Тематический словарь. Предназначен для актуализации словаря в соответствии с интересующей тематикой SMS. Например, может включать слова: *иноагент, акбар, бисмилля, ваалейкум*.

- Словарь неправильных (искаженных) слов и жаргонизмов. Например: *уря, ахи, ббббб, мухаджиры*, слова с опечатками. Включение такого словаря в базу данных нецелесообразно, поскольку его содержание зависит от грамотности и фантазии авторов сообщений и не поддается адекватному описанию. При проведении лингвистического анализа текстов сообщений всем неизвестным словам ставится в соответствие маркер имени лексико-грамматического класса (ЛГК) – существительное, которое, как показали эксперименты, является наиболее частотным в проанализированных текстах, а пропуск глагола и вопросительных слов вследствие возможных искажений, фиксируемых в текстах сообщений, может приводить к некорректным результатам.

Содержащиеся в лингвистической базе данных сведения далее используются на этапе лексико-грамматического тегирования, или «раскраски» – приписывания соответствующих маркеров имен ЛГК лексическим единицам текста SMS, причем одной лексической единице может быть поставлено в соответствие несколько маркеров ЛГК [15], и процедуры редуцирования имен ЛГК, что, в свою очередь, позволит устранить возможную неоднозначность с целью обеспечения качества проведения анализа.

Состав лексико-грамматического классификатора свойств русского языка

Далее представлены фрагменты разработанного лексико-грамматического классификатора свойств русского языка с указанием имен ЛГК, и в скобках – буквенно-цифровых обозначений или маркеров, необходимых для проведения автоматического лексико-грамматического анализа и решения задачи автоматической обработки коротких текстовых сообщений:

- существительные (N);
- глаголы (V);
- местоимения (PP, P\$);
- предлоги (IN);
- наречия и частицы (EMPTY);
- вопросительные слова (PQ, RV);
- основные знаки пунктуации, обозначающие границы предложения: «.», «!», «?»;

– дополнительные знаки пунктуации для обозначения границы предложения, эквивалентные знаку «.»: «.», «:», «;», «–» после пробела, «#», «/», «+», «вертикальная черта», «одионочная скобка» (если автор не поставил вторую).

Что касается ЛГК, используемых с целью определения классов предложений, то в состав лексико-грамматического классификатора свойств русского языка вошли следующие:

- N – существительное (NNIMO – существительное нарицательное, неодушевленное, мужской род, именительный падеж).
- IN – предлог (ING – предлог, родительный падеж).
- JR – относительное прилагательное (JRMО – относительное прилагательное, мужской род, именительный падеж).
- JJ – качественное прилагательное (JJMO – качественное прилагательное, мужской род, именительный падеж).
- PP – личное местоимение (PP1O – личное местоимение, 1-е лицо, единственное число, именительный падеж).
- P\$ – притяжательное местоимение (P\$MO – притяжательное местоимение, мужской род, именительный падеж).
- PQ – вопросительно-относительное местоимение (PQMO – вопросительно-относительное местоимение, мужской род, именительный падеж).
- CEMPTY – подчинительный союз.
- RQEMPTY – наречие качества, образа действия
- RLEEMPTY – наречие места, направления
- MDEEMPTY – модальное слово.
- XXEMPTY – наречие.
- VPM – повелительный глагол совершенного вида (VPM2 – глагол совершенного вида, повелительное наклонение, 2-е лицо, единственное число).
- VIM – повелительный глагол несовершенного вида (VIM2 – глагол несовершенного вида, повелительное наклонение, 2-е лицо, единственное число).
- VII – глагол несовершенного вида, изъявительное наклонение (VIIR1 – глагол несовершенного вида, изъявительное наклонение, настоящее время, 1-е лицо, единственное число).

– VPI – глагол совершенного вида, изъявительное наклонение (VPIPM – глагол совершенного вида, изъявительное наклонение, прошедшее время, мужской род).

– VIL – глагол несовершенного вида, возвратный (VILIR1 – глагол несовершенного вида, возвратный, изъявительное наклонение, 1-е лицо, единственное число).

– VPC – глагол совершенного вида, инфинитив.

– VIC – глагол несовершенного вида, инфинитив.

– LPP – страдательное причастие (LPPMO – страдательное причастие, прошедшее время, мужской род, именительный падеж).

Проведение процедур лексико-грамматического тегирования и редуцирования имен ЛГК, необходимых для обеспечения качества проведения анализа предполагает необходимость осуществить следующее.

Неизвестные слова, имена собственные, количественные числительные, цифры и буквенно-цифровые комбинации маркируются как N (существительное), т. к. цифры и буквенно-цифровые комбинации могут указывать на номер или код ячейки камеры хранения, сотовый телефон, электронный адрес, что может являться значимой информацией.

В именах ЛГК маркеры CEMPTY, RQEMPTY, RLEEMPTY, MDEEMPTY остаются без изменений; для наречий сохраняются только EMPTY; для существительных – только N; для глаголов – 3 начальных знака ЛГК; для остальных частей речи – 2. Маркеры EMPTY в середине и конце предложения, а JJ в начале и середине предложения так же, как и вопросительные слова, после V – удаляются.

Пример результата лексико-грамматического тегирования на основании лексико-грамматического классификатора и соответствующих классов речевых действий

Ниже представлены пример результата лексико-грамматического тегирования предложений текстов SMS на основании лексико-грамматического классификатора и соответствующие классы речевых действий, важные с точки зрения решения задачи:

Репрезентативы /1/

Это может привести к конфликту.

TCEEMPTY VIIR3 VPC IND NNIMD.

EMPTY + VII + VPC + IN + N. /1/

Директивы /2/

Ты сделай эту работу.

PP2O VIM2 PDFA NNIFA.

PP + VIM + N. /2/

Комиссивы /3/

Я обещаю сделать эту работу.

PP1O VIIR1 VPC PDFA NNIFA.

PP+VII+VPC+N. /3/

Декларативы /4/

Я объявляю вас мужем и женой.

PP1O VIIR1 PP2G NNAMI CCEMPTY NNAFI.

PP + VII + PP + N + N. /4/

Экспрессивы /5/

Русские идут!

JRPO/JRPA VIIR3!

JR + VII /5/

Простите.

VPM /5/

Интеррогативы /6/

Что он может сказать?

RQINO PP3MO WDEMPTY/VIIR3 VPC?

RQ + PP + VII + VPC? /6/

Перформативы /7/

Прошу предъявить билеты.

VIIR1 VPC NNIMPA.

VII+VPC+N. /7/

Лингвистические правила анализа короткого текстового сообщения

Принимая во внимание вышеизложенное, процесс автоматического соотношения наблюдаемых лингвистических характеристик лексических единиц обрабатываемого короткого текстового сообщения с классами речевых действий с целью фиксации последних должен выполняться в границах анализируемого предложения с использованием ресурсов лингвистической базы данных и набора лингвистических правил анализа текста, фрагменты которого представлены ниже.

Если в начале предложения присутствует обращение (например, *уважаемый сэр*) и оно отделено знаком препинания, то его следует рассматривать как отдельное предложение класса /5/.

Предложение, содержащее нетекстовые знаки и заканчивающиеся эквивалентом знака «.», необходимо относить к классу /1/.

При наличии знаков эмодзи или их последовательности в отсутствии других слов (могут служить кратким согласием/подтверждением) предложения относить к классу /5/. Их, кроме обозначающих согласие или отказ, следует игнорировать.

Повествовательные предложения относить к классам /1/, /3/, /4/, /5/, /7/.

Принадлежность к классу /2/ повелительных предложений и к классу /6/ вопросительных – описывается следующим образом:

– *Повелительное предложение класса /2/*. Признаком класса являются повелительные глаголы, в основном с маркерами VPM, VIMI. Повелительный глагол должен находиться первым перед другими глаголами V (если они присутствуют). Признаком также может являться лексема с маркером MEMPTY, если после нее присутствует глагол V. Наличие знака «?» отменяет отнесение к классу /2/, в таком случае следует зафиксировать принадлежность классу /6/.

– *Вопросительное предложение класса /6/*. К признакам принадлежности предложения классу относится наличие в предложении вопросительных слов с маркерами: PQ, CEMPTY, RQEMPTY, RLEEMPTY, при этом само вопросительное слово должно находиться перед глаголом V. В вопросительном предложении класса /6/ должно быть вопросительное слово и глагол V, при этом знак «?» необязателен, или глагол V в сопровождении других лексических единиц и знака «?».

Если в предложении присутствуют 4 и более слов, знак препинания «?», но отсутствует глагол V, то установить принадлежность к классам /5,6/.

Если присутствует вопросительное слово, глагол V и знак препинания «!», то установить принадлежность к классам /5,7/.

Таким образом, при составлении правил учитывается количество лексических единиц в составе анализируемого предложения, их принадлежность ЛГК, местоположение в границах предложения, присутствие знаков препинания и т. п.

Ситуации коммуникации посредством SMS

Процесс целевой коммуникации посредством SMS можно разделить на 5 ситуаций (этапов):

- 1) знакомство;
- 2) вхождение в доверие;
- 3) определение слабых мест;
- 4) укрепление отношений;
- 5) финальный этап взаимодействия.

В таблице представлено процентное соотношение предложений классов, рассмотренных выше, в составе коротких текстовых сообщений, описывающих каждую из пяти ситуаций, полученную в результате автоматической обработки набора данных из 2 000 коротких текстовых сообщений. Очевидно, что каждая из ситуаций отличается количеством классов предложений, требуемым для своего корректного описания, в процентном соотношении, а в одном SMS могут присутствовать несколько предложений разных классов. Заметим, что проанализированные предложения, относящиеся к разным классам, рассматривались как несколько предложений каждого класса. Например, предложение, относящееся одновременно к двум классам /1,5/, рассматривалось как два предложения классов /1/ и /5/ соответственно.

Доля предложений, относящихся к соответствующим классам,
описывающим каждую из пяти ситуаций, в %

Классы речевых действий	Ситуация коммуникации				
	Знакомство	Вхождение в доверие	Определение слабых мест	Укрепление отношений	Финальный этап взаимодействия
/1/ Репрезентативы	22	30	65	41	60
/2/ Директивы	6	13	9	10	13
/3/ Комиссивы	0	8	2	6	4
/4/ Декларативы	0	0	2	6	0
/5/ Экспрессивы	38	6	9	13	3
/6/ Интеррогативы	34	43	12	15	17
/7/ Перформативы	0	0	1	9	3

Количественная оценка соотношения автоматически выявленных классов речевых действий, приходящихся на каждую из рассматриваемых типовых ситуаций целевой коммуникации, позволяет смоделировать наблюдаемую ситуацию при анализе группы коротких текстовых сообщений участников процесса коммуникации и установить на каком из этапов находятся отношения корреспондентов, а последующий анализ содержания сообщений позволит узнать коммуникативные цели корреспондентов, а также цели и средства их возможных реальных действий. Такая оценка будет тем точнее, чем больше выборка обрабатываемых сообщений.

В каждой из рассматриваемых ситуаций присутствуют составляющие сообщения, относящиеся к классу репрезентативов, как правило, это предложения, описывающие общие сведения или факты, причем в большей степени это характерно для ситуации определения слабых мест и финального этапа взаимодействия в ситуациях (65 %) и (60 %), соответственно. Для ситуации знакомства в текстах сообщений характерно наличие предложений, указывающих на преобладание классов речевых действий – экспрессивов (38 %) и интеррогативов (34 %), что предполагает наличие высказываний эмоционального характера и вопросов. В ситуации вхождения в доверие заметно преобладание вопросов (43 %) и в меньшей степени (13 %) повелительных предложений, в то время как их доля в ситуациях укрепления отношений составляет 15 % – интеррогативы, 10 % – директивы и финального этапа взаимодействия: 17 % – интеррогатив, 13 % – директивы соответственно.

В целом, полученные результаты отражают цели корреспондентов в рассматриваемых ситуациях коммуникации: эмоциональное привлечение объекта коммуникации к общению; задавание простых вопросов и озвучивание обязательств спрашивающего; осторожные вопросы и их эмоциональная окраска по набору фактов; вопросы, повеления и эмоции в комплексе для закрепления отношений корреспондентов; прямые рекомендации к действию на завершающем этапе коммуникации.

Приведем примеры полученных результатов решения рассматриваемой нами задачи для разных ситуаций коммуникации.

1. Знакомство

Как тебя зовут?

CSEMPTY + N + VII + ? /6/

Брат с вами все хорошо?

N + IN + PP + JJ + ? /5,6/

2. Вхождение в доверие

Где ты выучил русский язык брат?

RLEMPTY + PP + VPI + N + N + N+? /6/

На арабском плохо пишу пока что.

IN + VII + «.» /1,3/

3. Нащупывание слабых мест

СубханАллах тебя и никто не просит чтобы ты верил ему

N + PP + VII + PP + VII + PP /1,3/

Ты на стукача похож, /1,5/ причём сильно

PP + IN + N + JJ + «,» /1,5/ EMTY + JJ /5/

4. Укрепление отношений

Цель джихада ведь не заключается побольше убить кафиров

N + N + VIL + VPC + N /1,3/

Пусть Аллагь возвысит его степень в раю

EMPTY + N + VPL + PP + N + IN + N /1,3/

5. Финальный этап

ты маршрут более менее знаешь

PP + N + N /1,5/

Если взял намерение то уезжай черёд Беларусь нормально

EMPTY + VPI + N + VIM + N + N + JJ /7/

Заметим, что в приведенных примерах есть лексические единицы, которые не могут присутствовать в словаре правильных слов и, согласно описанным выше правилам, в процессе решения задачи считаются существительными.

Отметим, что процесс коммуникации может быть достаточно длительным, а корреспонденты не ограничиваются обменом несколькими текстовыми сообщениями. Предложенное решение задачи автоматической обработки коротких текстовых сообщений позволяет определить, на каком этапе взаимодействия находятся корреспонденты, оценить, стоит ли наблюдателю обратить внимание на потенциальную угрозу безопасности.

ЛИТЕРАТУРА

1. *McDonald, D.* Natural language generation systems / D. McDonald L., Bolc. – Springer-Verlag, 1988. – 388 p.
2. *Остин, Дж.* Слово как действие / Дж. Остин // Новое в зарубежной лингвистике. – Вып. 17: Теория речевых актов. – М., 1986. – С. 122–129.

3. *Серль Дж. Р.* Классификация иллокутивных актов // Новое в зарубежной лингвистике. – Вып. 17: Теория речевых актов. – М., 1986. – С. 170–194.
4. *Сазонова, П. А.* Автоматическое порождение фраз естественного языка по OWL-модели, семантике и прагматике / П. А. Сазонова // Сб. материалов междунар. науч.-практ. конф. «Анализ изображений, сетей и текстов» (АИСТ). – Екатеринбург : Урал. Фед. ун-т, 2014 г.
5. *Segalovich, I.* A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine / I. Segalovich // Proceedings of MLMTA-2003. – Las Vegas, 2003. – P. 273–280.
6. *Korobov, M.* Morphological Analyzer and Generator for Russian and Ukrainian Languages / M. Korobov // Analysis of Images, Social Networks and Texts. – Lecture Notes in Computer Science. – Vol. 9308. – Cham: Springer, 2015. – P. 320–332.
7. *Straka, M.* Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe / M. Straka, J. Strakova // Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. – Vancouver, 2017. – P. 88–99.
8. *Stanza: A Python Natural Language Processing Toolkit for Many Human Languages* / P. Qi, Y. Zhang, J. Bolton [et al.] // Proceedings of ACL 2020: System Demonstrations. – 2020. – P. 101–108.
9. *Scikit-learn: Machine Learning in Python* / F. Pedregosa, G. Varoquaux, A. Gramfort [et al.] // Journal of Machine Learning Research. – 2011. – Vol. 12. – P. 2825–2830.
10. *Bag of Tricks for Efficient Text Classification* / A. Joulin A., E. Grave E., P. Bojanowski [et al.] // Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics. – Valencia, 2017. – P. 427–431.
11. *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* / J. Devlin J., M.-W. Chang M.-W., K. Lee K. [et al.] // Proceedings of NAACL-HLT 2019. – Minneapolis, 2019. – P. 4171–4186.
12. *DeepPavlov: Open-Source Library for Dialogue Systems* / M. Burtsev, A. Seliverstov, R. Airapetyan, [et al.] // Proceedings of ACL 2018: System Demonstrations. – Melbourne, 2018. – P. 122–127.
13. *Kuraton, Y.* Adaptation of Deep Bidirectional Multilingual Transformers for Russian Language / Y. Kuraton, M. Arkhipov // arXiv preprint arXiv: 1905.07213. – 2019.
14. *Рубашко, Н. К.* Разработка лексико-грамматического классификатора флективных языков / Н. К. Рубашко // Вест. Мин. гос. лингвист. ун-та. Сер. 1. Филология. – № 11. – Минск : МГЛУ 2003. – С. 84–89.
15. *Антонов, С. Г.* Модели текста естественного языка для решения задач автоматической обработки. Оценка возможности преобразования лингвистических данных для разных языков // С. Г. Антонов, М. В. Федюкин / Информатизация и связь. – № 4 – М., 2022. – С. 85–97.

Поступила в редакцию 10.06.2026