প্রকল্পের তথ্য প্রদানের ছক ২০২২-২০২৩ অর্থ বছর

<table>
<tr><th>ক্রমিক নং</th><th>বিষয়বস্তু</th><th>বিবরণ</th></tr>
<tr><td>১</td><td>প্রকল্পের নাম</td><td>‘গবেষণা ও উন্নয়নের মাধ্যমে তথ্য প্রযুক্তিতে বাংলা ভাষা সমৃদ্ধকরণ (১ম সংশোধিত)’ প্রকল্প</td></tr>
<tr><td>২</td><td>মেয়াদ</td><td>০১/০৭/২০১৬/-৩০/০৬/২০২৪ (নো-কস্ট এক্সটেনশন ৩০/০৬/২০২৪ পর্যন্ত)</td></tr>
<tr><td>৩</td><td>প্রাক্কলিত ব্যয়</td><td><table><tr><th>অর্থের উৎস</th><th>পরিমান</th></tr><tr><td>বৈদেশিক সাহায্য</td><td>নাই (লক্ষ টাকায়)</td></tr><tr><td>জিওবি</td><td>(১৫৮৯৬.৬৯) (লক্ষ টাকায়)</td></tr><tr><td>মোট</td><td>(১৫৮৯৬.৬৯) (লক্ষ টাকায়)</td></tr></table></td></tr>
<tr><td>৪</td><td>জনবল</td><td>কর্মকর্তা: ০২ জন
পরামর্শক: ০৯ জন
কর্মচারী: ০৭ জন</td></tr>
<tr><td>৫</td><td>পটভূমি</td><td>বাংলা ব্যবহারের দিক থেকে পৃথিবীতে প্রভাবশালী ভাষাগুলোর অন্যতম। বাংলা ভাষাভাষীর রয়েছে রক্তস্নাত ভাষা-আন্দোলনের ইতিহাস। দেশ ও ভাষার মর্যাদা রক্ষায় এই জাতির রয়েছে গৌরবময় ঐতিহ্য, রয়েছে ভাষার প্রতি দরদ, ভাষাকে সমুন্নত রাখার চেতনা। কিন্তু এ কথা সত্য যে, বাংলা ভাষাকে প্রযুক্তি বান্ধব করার ক্ষেত্রে এদেশে প্রয়োজনীয় ভিত্তি তৈরি হয়নি। বিশেষ করে কম্পিউটিংয়ে বাংলা ভাষাকে অভিযোজিত করার ক্ষেত্র-- খুব বেশি অগ্রসর হয়নি। বাংলা ভাষাকে প্রযুক্তিবান্ধব হিসেবে গড়ে তোলার বাংলা ভাষা সমৃদ্ধকরণ প্রকল্প কাজ করে যাচ্ছে।

দেশকে ‘ডিজিটাল বাংলাদেশ’ হিসেবে গড়ে তোলার একটি প্রধান শর্ত বাংলা ভাষাকে প্রযুক্তিবান্ধব করা। তথ্য ও যোগাযোগ প্রযুক্তিতে বাংলা ভাষার ব্যবহার ও প্রয়োগ হলে দেশের প্রশাসনিক ব্যবস্থা, শিক্ষা ব্যবস্থা ও যোগাযোগ কাঠামোতে নতুন পরিবর্তন সূচিত হবে। আন্তর্জাতিক পর্যায়ে বাংলা ভাষাকে যথাযথ মর্যাদা দান ও উৎকর্ষে পৌঁছানো সম্ভব হবে। এই প্রকল্পের আওতায় গৃহীত পদক্ষেপসমূহ বাস্তবায়িত হলে ডিজিটাল মাধ্যমে বাংলা ভাষার ব্যবহার বিস্তৃত হবে এবং 'ফ্রন্টিয়ার টেকনোলজি'তে বাংলা ভাষাকে ব্যবহার উপযোগি করার 'প্রাইমারি রিসোর্স' প্রস্তুত হবে।</td></tr>
<tr><td>৬</td><td>লক্ষ্য ও উদ্দেশ্য</td><td>লক্ষ্য:
বাংলা ভাষার জন্য বিভিন্ন প্রযুক্তির মাধ্যমে (ওয়েব, মোবাইল, কম্পিউটার) ব্যবহারযোগ্য ৪০টি সফটয়্যার/টুলস/ রিসোর্স উন্নয়ন করা, যাতে ডিজিটাল মাধ্যমে বাংলা ভাষা ব্যবহার করতে কোনো প্রতিবন্ধকতা না থাকে। একই সঙ্গে ভ্যালুয়েবল রিসোর্স তৈরির মাধ্যমে বিশ্বে বিভিন্ন পর্যায়ে ও প্রতিষ্ঠানে (যেমন জাতিসংঘ) বাংলা ভাষার অবস্থানকে আরো উন্নত ও মর্যাদাপূর্ণ করা।
উদ্দেশ্য:
গবেষণা ও উদ্ভাবনের মাধ্যমে গ্লোবাল প্ল্যাটফর্মে নেতৃস্থানীয় ভাষা হিসেবে বাংলা ভাষাকে প্রতিষ্ঠা করা।
আইসিটি সহায়ক বাংলা ভাষার বিভিন্ন ফিচার প্রমিতকরণ।
বাংলা কম্পিউটিং এর জন্য টুলস উন্নয়ন, টেকনোলজিস ও বিষয়বস্তু উন্নয়ন।
আইসিটির ক্ষেত্রে বাংলা সমৃদ্ধকরণ উন্নয়ন ও আধুনিকায়নের নিমিত্ত পরীক্ষা, জরিপ এবং গবেষণা পরিচালনা করা।</td></tr>
</table>

৭	প্রকল্পের উল্লেখ্যযোগ্য কম্পোনেন্ট	বাংলা ভাষার জন্য ৪০টি সফটয়্যার/টুলস/রিসোর্স উন্নয়ন করা হবে। এর ফলে আন্তর্জাতিক পর্যায়ে বাংলা ব্যবহারের সুযোগ তৈরি হবে। সম্পূর্ণ বাংলা করপাস এবং বাংলা স্টাইল-শিট সম্পন্ন হলে বিশ্বমানের বাংলা কম্পিউটিং-এর ভিত্তি তৈরি করা যাবে। এই ৪০টির মধ্যে প্রধানতম ১৬টি উপাংশের সংক্ষিপ্ত পরিচয় বাংলা ও ইংরেজি শিরোনামসহ (ইংরেজি মূল, বাংলা সহজে অনুধাবনের জন্য ঈষৎ পরিমার্জিত) নিম্নে উল্লেখ করা হলো : **ক) বাংলা টেক্সট করপাস (**Bangla Syntactic Treebank Corpus with Processing Pipeline and Distribution Platform) বাংলা ভাষার জন্য একটি প্রতিনিধিত্বমূলক করপাস তৈরি হচ্ছে। এটি মূলত অ্যানোটেটেড সিনট্যাকটিক ট্রি-ব্যাংক করপাস হবে। এর পরিমাণ কমপক্ষে ১০০মিলিয়ন। যার অন্তত শতকরা ১০ ভাগ গোল্ড স্ট্যান্ডার্ড রক্ষা করবে। করপাসটি প্রতিনিয়ত নতুন র-ডেটা যুক্ত হওয়ার সুযোগ থাকবে, ফলে এটি দ্রুত বিশাল ভান্ডারে পরিণত হবে। এই করপাস ব্যবহার করে 'মেশিন লানিং' এর মাধ্যমে কৃত্রিম বুদ্ধিমত্তাসম্পন্ন বাস্তব জীবনে ব্যবহারযোগ্য অ্যাপ্লিকেশন তৈরি করা যাবে এবং এর ফলে বাংলা ভাষাকে বৈজ্ঞানিকভাবে বিশ্লেষণ করা যাবে। এই কম্পোনেন্টে করপাস সংগ্রহ, প্রক্রিয়াজতকরণ এবং বিতরণ ব্যবস্থা তৈরির জন্য কয়েকটি টুলস তৈরি হবে যা একটি পাইপলাইনের মাধ্যমে যুক্তি থাকবে। করপাসে প্লাটফরমে যেসকল অ্যানালাইসিস ফিচার থাকছে: ওয়ার্ড/ফ্রেজ ফ্রিকোয়েন্সি, এনগ্রাম, কনকর্ড্যান্স, কোলেকেশন প্রভৃতি। **খ) সঠিক: বাংলা বানান ও ব্যাকরণ সংশোধক (**Development of Bangla Spell & Grammar Checker) বানান পরীক্ষক ও ব্যাকরণ সংশোধক হলো বাংলা ভাষার শব্দ,বাক্য স্বয়ংক্রিয়ভাবে সম্পাদনা করার সফটওয়্যার। এই সফটওয়্যার কেবল ভুল বানান চিহ্নিত করবে তা নয়, বরং স্বয়ক্রিয়ভাবে সংশোধনের পরামর্শ দেবে। বিশেষ করে, একই রকম উচ্চারণ কিন্তু বানান ভিন্ন, একই রকম বানান কিন্তু অর্থ ভিন্ন এমন কনটেক্সট নির্ভর বানান ভুল বিষয়ে সংশোধনী দেবে। ব্যাকরণ সংশোধক ভুল বাংলা বাক্য জানাতে সাহায্য করবে। সরল ও জটিল বাক্যের প্রচলিত সাধারণ ভুলগুলো চিহ্নিত করে ব্যবহারকারীর কাছে বিকল্পসহ সঠিক বাক্য উপস্থাপন করবে। এই বানান ও ব্যাকরণ পরীক্ষক সফটওয়্যারটি বাংলা একাডেমির প্রমিত বানানবিধি ও প্রমিত বানান অভিধানকে অনুসরণ করবে। **গ) বর্ণ: বাংলা ওসিআর ও হাতের লেখা শনাক্তকরণ (**Further Improvement of Bangla OCR Developed by ICTD & Integrating Hand Writing Recognition System) ছাপানো অক্ষর ও হাতের লেখা শনাক্ত করার সফটওয়্যার হলো ওসিআর। এর মাধ্যমে যেকোনো অপরিবর্তনযোগ্য ফরম্যাটে (পিডিএফ, জেপেগ) থাকা বর্ণমালাকে কম্পিউটার শনাক্ত করতে পারে এবং দ্রুত তা সার্চেবল টেক্সটে রূপান্তর করে দিতে পারে। এই কম্পোনেন্টের মাধ্যমে কম্পিউটার কম্পোজ, টাইপরাইটারে এবং লেটারপ্রেসে ছাপা বাংলা ডকুমেন্টকে স্ক্যান করে সার্চেবল ক্যারেকটারে রূপান্তর করা যাবে। এই কম্পোনেন্ট বাংলা হাতের লেখা শনাক্ত করে তা টেক্সটে রূপান্তর করে দিতে পারবে। এছাড়াও স্টাইলাস দিয়ে প্যালেটে বা ডিভাইস স্ক্রিনে বর্ণ লেখার পর তা বাংলা ইউনিকোড টেক্সটে রূপান্তর করা যাবে। **ঘ) কথা: বাংলা টেক্সট টু স্পিচ এবং স্পিচ টু টেক্সট সফটওয়্যার (**Development of Bangla Speech to Text & Text to Speech Software) এই কম্পোনেন্টের মাধ্যমে বাংলা কথাকে লেখায় এবং লেখাকে কথায় রূপান্তরিত করা যাবে। রেকর্ড করা বা চলমান বাংলা কথাকে লেখায় রূপান্তর করবে স্পিচ টু টেক্সট (এসটিটি) সফটওয়্যার। এর মাধ্যমে বাংলা ভাষার ভাষণ ও বক্তব্য দ্রুত লিখিত বা কম্পোজ অবস্থায় পাওয়া যাবে। বিভিন্ন সাক্ষাৎকার, বিবৃতি দ্রুত যন্ত্রের মাধ্যমে অনুলিখন করা যাবে। পক্ষান্তরে টেক্সট টু স্পিচ (টিটিএস) অ্যাপ্লিকেশন হলো ডিজিটাল টেক্সটকে উচ্চারিত শব্দে রূপান্তর করার পদ্ধতি। এই অ্যাপ্লিকেশন যাদের চোখের দৃষ্টির কারণে পড়তে অসুবিধা হয় তাদের উপকারে আসবে। এর মাধ্যমে সরকারি জরুরি বিজ্ঞপ্তি, নির্দেশনা, পত্রিকার শিরোনাম/ তাজা খবর শোনা যাবে। ওয়েবসাইটে প্রকাশিত লেখা সহজে শোনা যাবে। **ঙ) জনমত: বাংলা টেক্সটের সেন্টিমেন্ট বিশ্লেষণ (**Development of Sentiment Analysis Software in Bangla) সেন্টিমেন্ট অ্যানালাইসিস টুলস সাধারণত কোনো ডকুমেন্ট বা প্যারাগ্রাফের টেক্সট বিশ্লেষণ করে

		বলতে পারে প্যারাগ্রাফটির বক্তব্য ইতিবাচক, নেতিবাচক না নিরপেক্ষ। এর মাধ্যমে ওয়েবসাইটের মন্তব্য ও ফিডব্যাক বিশ্লেষণ করা যায়। এর মাধ্যমে দ্রুত বাজার-জরিপ, জনমত জরিপ করা, নির্বাচন উত্তর জনমত যাচাই দ্রুত করা যাবে। এই কম্পোনেন্টের মাধ্যমে একটি টেক্সটের মাধ্যমে প্রকাশিত অনুভূতি যেমন হাসি, আনন্দ, রাগ প্রভৃতি জানা যাবে। সেন্টিওয়ার্ডনেট। এর কয়েকটি উল্লেখযোগ্য ফাংশনাল প্যারামিটার হলো, ৩ ক্লাস সেন্টিমেন্ট (পজেটিভ নেগেটিভ নিউট্রাল), ৫ ক্লাস সেন্টিমেন্ট (স্ট্রংলি পজেটিভ, উইকলি পজেটিভ, নিউট্রল, স্ট্রংলি নেগেটিভ, উইকলি নেগেটিভ), ৬ ক্লাস ইমোশন (আনন্দ, ভয়, রাগ, দুঃখ, বিরক্তি, বিস্ময়)। **চ) অনুবাদক: বাংলা ম্যাশিন ট্রান্সলেটর** (Development of the Bangla Machine Translator) যান্ত্রিক অনুবাদের মাধ্যমে সাধারণত তথ্যমূলক বাক্যগুলো স্বয়ংক্রিয়ভাবে সহজে অনুবাদ করা যায়। এই ধরনের অনুবাদকের মাধ্যমে তথ্যমূলক বাংলা, দৈনন্দিন বাংলা, প্রাতিষ্ঠানিক রচনা/ডকুমেন্টস/ নথি, সংবাদ বিজ্ঞপ্তি, সর্বশেষ সংবাদসহ কুশলাদি ও সংক্ষিপ্ত সংলাপ দ্রুত নির্ভুলভাবে অনুবাদ করা সম্ভব হবে। এই অনুবাদকের মাধ্যমে বাংলা থেকে ইংরেজি এবং ইংরেজি থেকে বাংলা ছাড়াও বাংলা থেকে স্প্যানিশ, ফরাসি, জার্মান, রুশ, মান্দারিন, জাপানিজ, কোরিয়ান, আরবি, হিন্দি ভাষায় এবং উল্লিখিত ভাষাগুলোকে থেকে বাংলায় অনুবাদ করা যাবে। এই যান্ত্রিক অনুবাদক তৈরির জন্য প্যারালাল করপাস তৈরি করা হবে যা পরবর্তী সময়ে রিসোর্স হিসেবে ব্যবহার করা যাবে। **ছ) বাংলা ভার্চুয়াল ব্যক্তিগত সহকারী (Development of Bangla Virtual Private Assistant)** বাংলা ভার্চুয়াল প্রাইভেট অ্যাসিস্ট্যান্টের মাধ্যমে মোবাইল বা অনুরূপ ডিভাইসে বাংলা ভাষায় নির্দেশনা দিয়ে সার্ভিস পাওয়া যাবে যেমনটি গুগল প্রাইভেট অ্যাসিসট্যান্ট, সিরি, করটানা, অ্যালেক্সা করে থাকে। এই কম্পোনেন্টের প্রাথমিকভাবে ব্যবহারের ক্ষেত্রগুলো হলো: ক. নিয়মিত যোগাযোগের কাজগুলো করা যেমন, নিকটস্থ দোকানে পণ্য বা খাবার অর্ডার করা, কোনো জরুরি ফোন নম্বর বের করে কল করা ইত্যাদি। অর্থাত্ এর দ্বারা বাংলা ভাষায় ডিভাইস নিয়ন্ত্রণ করা যাবে। খ) বাংলা এআই সহকারীর সঙ্গে কথোপকথন বা চ্যাটবট। গ. নেটিভ রিপোজিটরি থেকে প্রশ্ন অনুযায়ী তথ্য বের করে প্রশ্নকারীকে উত্তর জানানো। প্রাথমিকভাবে সরকারি জনগুরুত্ববাহী ১০টি সেবা পাওয়া যাবে এই অ্যাপ্লিকেশনের মাধ্যমে। ঘ. বাহ্যিক ইনফরমেশন রিপোজটরি যেমন গুগল বা সমরূপ থেকে তথ্য এনে ব্যবহারকারীকে জানানো। এই কম্পোনেন্টেটি মাল্টিমোডাল মাল্টিটাস্কিং অ্যাপ্লিকেশন। ব্যবহারকারীর কাছে পুরো অ্যাপ্লিকেশনটি একটি সলুশন হিসেবে উপস্থাপিত হবে। প্রয়োজনে বিটুবি মডেলে এই অ্যাসিসট্যান্ট ব্যবহার করা যাবে। **জ) ইশারা: বাক ও শ্রবণ প্রতিবন্ধী ব্যক্তির জন্য সাইন ল্যাঙ্গুয়েজ কনভার্টার** (Development of Software for Disable People) এই কম্পোনেন্টের মাধ্যমে বাক ও শ্রবণ প্রতিবন্ধী ব্যক্তিদের জন্য সাইন টু স্পিচ সফটওয়্যার উন্নয়ন করা হচ্ছে। এর মাধ্যমে একজন ব্যবহারকারী কোনো স্মার্টফোন, ট্যাবলেট বা ক্যামেরাযুক্ত পিসির সামনে দুই হাত, মুখ ও শরীরের উর্ধ্বাংশের সমন্বয়ে ইঙ্গিত ভাষা প্রকাশ করবেন। সফটওয়্যার এই ইঙ্গিত ভাষাকে বাংলা ইউনিকোড টেক্সটে রূপান্তর করবে। কয়েকটি বিশেষ ডোমেইনে এই সিস্টেম কাজ করবে। প্রয়োজনে এই টেক্সট সাথে সাথে উচ্চারিত কথায় রূপান্তরিত হবে। ফলে বাক ও শ্রবণ প্রতিবন্ধীদের প্রাথমিক যোগাযোগের সীমাবদ্ধতা দূর করা সম্ভব হবে। **ঝ) আলো: দৃষ্টি প্রতিবন্ধী ব্যক্তির জন্য স্ক্রিন রিডার** (Development of Screen Reader Software) স্ক্রিন রিডার সফটওয়্যার এর মাধ্যমে দৃষ্টি প্রতিবন্ধী ব্যক্তি বা স্বল্প দৃষ্টিসম্পন্ন ব্যক্তিরা কম্পিউটার বা মোবাইল ব্যবহার করতে পারবেন। কম্পিউটারের পর্দায় ভেসে আসা আইকনের নাম ও ডকুমেন্টের বাংলা লেখা পড়ে শোনাবে এই সফটওয়্যার। কম্পিউটারের ইন্টারফেসে থাকা বাটন বা আইকন বাংলা ভাষায় চিনিয়ে দেবে। বাংলায় কমান্ড দেয়া যাবে, যার মাধ্যমে দৃষ্টিহীন ব্যক্তি সহজে কম্পিউটার বা মোবাইল ব্যবহার করতে পারবে। এই সফটওয়্যারের মাধ্যমে বাংলা ভাষায় কম্পিউটারের কমান্ড শোনা যাবে, ফলে প্রতিবন্ধীদের জন্য বিদেশি ভাষার প্রতিবন্ধকতা দূরীভূত হবে। **ঞ) ইউ-বোর্ড: জাতীয় কিবোর্ড** (Improvement of the National Keyboard (Bangla))

		কম্পিউটারে নির্বিঘ্নে বাংলা কম্পোজ করার জন্য বাংলাদেশ কম্পিউটার কাউন্সিল কর্তৃক এর 'জাতীয় বাংলা কিবোর্ড'-কে আরো উন্নত করা হচ্ছে। এর ফলে বিভিন্ন জনপ্রিয় অপারেটিং সিস্টেম যেমন, উইন্ডোজ, ম্যাকওএস, লিনাক্স, অ্যান্ড্রয়েড ও আইওএস এ একই মান ও প্রক্রিয়া অনুসরণ করে বাংলা লেখা সম্ভব হবে। **ট) বাংলা ফন্ট কনভার্টার (**Development of the Bangla Font Interoperability Engine) বিভিন্ন অপারেটিং সিস্টেম, ওয়েব ও মোবাইল প্লাটফর্ম-এ বাংলা লেখা ফন্ট স্থানান্তরের সময় ভেঙে যায়। এ ফন্টভাঙা সমস্যা থেকে মুক্তি পাওয়ার জন্য একটি 'ফন্ট এনকোডিং কনভার্টার' তৈরি করা হচ্ছে, যা 'ফন্ট ইন্টারঅপারেবল ইঞ্জিন' হিসেবে কাজ করবে। **ঠ) ধ্বনি: বাংলা থেকে আইপিএ কনভার্টার (**Bangla to IPA Automatic Converter) আইপিএ হলো আন্তর্জাতিক ধ্বনিমূলক বর্ণমালা। আন্তর্জাতিক ধ্বনিমূলক বর্ণমালা মানুষের দ্বারা উচ্চারিত প্রায় সব ধ্বনির লিখিত রূপকে প্রকাশ করা যায়। পৃথিবীর প্রায় সব ভাষার জন্য 'ইন্টারন্যাশনাল ফোনেটিক অ্যাসেসিয়েশন' এই বর্ণমালা প্রণয়ন ও প্রমিতকরণ করে থাকে। ভাষাবিজ্ঞানীসহ বিদেশি ভাষার ছাত্র-শিক্ষক, স্পিচ-প্যাথলজিস্ট, গায়ক, অনুবাদক এই বর্ণমালা ব্যবহার করে থাকে। এই কম্পোনেন্টের মাধ্যমে বাংলা ইউনিকোড টেক্সটকে আইপিএতে রূপান্তর করা যাবে। সাধারণত এই কনভার্টার ব্রড ও ন্যারো ট্রান্সক্রিপশন রীতি অনুসরণ করে তৈরি হচ্ছে। এই কনভার্টার তৈরির ফলে দ্রুত বাংলা ভাষার উচ্চারিত রূপকে আন্তর্জাতিক মান অনুসারে লেখা যাবে। **ড) নৃগোষ্ঠী ভাষার ডিজিটাইজেশন (**Digital resources and Keyboard for Ethnic Minority group's language) বাংলাদেশের ক্ষুদ্র-নৃগোষ্ঠীর ভাষাগুলোর অধিকাংশ খুব স্বল্প পরিসরে তথ্য প্রযুক্তির জগতে ব্যবহৃত হয়। অধিকাংশ ক্ষেত্রেই ভাষার মানসম্পন্ন ডকুমেন্টশন, রিসোর্স ও ম্যাটেরিয়াল নেই। এদের মধ্যে কয়েকটি বিপন্ন অবস্থায় রয়েছে। অনেক ভাষার পর্যাপ্ত ডিজিটাল ডেটা নেই, ফন্ট ও এনকোডিং নেই। অনেক ভাষার লিপিও নেই। ডিজিটাল রিসোর্স তৈরির মাধ্যমে এই ভাষাগুলোকে প্রযুক্তি জগতে ব্যবহারের উপযোগী করা হবে । এই লক্ষ্যে, যেসব ভাষার নিজস্ব বর্ণমালা রয়েছে, সেসব ভাষার জন্য কিবোর্ড সফটয়্যার উন্নয়ন করা হচ্ছে। যেসব ভাষার লিপি নেই বা লিপি নিয়ে বিতর্ক রয়েছে তাদের জন্য একটি 'মাস্টার টেমপ্লেট' তৈরি করা হবে, যাতে তাঁরা নিজেদের মতো কিবোর্ডের কাজ করতে পারে। একই সঙ্গে অন্যান্য বিপন্ন ভাষাগুলোকে ডিজিটাল আর্কাইভে সংরক্ষণের ব্যবস্থা করা হচ্ছে। **ঢ) প্রমিতকরণ ও স্টাইল গাইড (**Development of Bangla Style Guide) প্রযুক্তির সঙ্গে ভাষার সম্মিলনের প্রথম শর্ত হলো ভাষার রীতি এবং ভাষা ব্যবহারে উপাদানসমূহকে প্রমিতকরণ করা। এই কম্পোনেন্টের মাধ্যমে প্রমিতকরণের সকল লজিস্টিকস সাপোর্ট প্রদান করা হবে। দেশের শীর্ষস্থানীয় ভাষা বিশেষজ্ঞ, তথ্যপ্রযুক্তিবিদ, লেখক, লিপি বিশেষজ্ঞ, শিক্ষাবিদ, প্রশাসক, সম্পাদক, আইনজ্ঞ প্রমুখের সমন্বয়ে গঠিত কমিটি এই প্রমিতকরণ করা হবে। প্রমিতকরণের প্রাথমিক বিষয়গুলো হলো: বাংলা ক্যারেকটার , বাংলা সর্টিং অর্ডার , ইঙ্গিতভাষা/সাইন ল্যাংগুয়েজ , বিরাম চিহ্ন প্রয়োগ-রীতি প্রমিতকরণ, ওয়েবে ও মুদ্রণ জগতে টাইপোগ্রাফিকাল স্টাইল শিট, প্রকাশনায় অন্যভাষার সঙ্গে বাংলা ভাষার ব্যবহারের রীতি (ইংরেজি, আরবি, সংস্কৃত, চাকমা প্রভৃতি) নির্ধারণ, বাংলা গ্লোসারি, টীকা ব্যবহারের রীতি (ফুটনোট ও এন্ডনোট), গ্রন্থপঞ্জি ও নির্ঘণ্ট লেখার নিয়ম নির্ধারণসহ লিপি ও ভাষাসংক্রান্ত প্রযুক্তিজগতে প্রয়োগযোগ্য সকল মান প্রমিতকরণ। ওয়ার্কশপ ও পাবলিক কনসালটেশনের মাধ্যমে মানগুলো চূড়ান্ত করা হবে। প্রমিত বিষয়গুলো ভলিউম আকারে মুদ্রিত করা হবে এবং ই-বুক ও ওয়েব-সংস্করণেও প্রকাশ করা হবে। ণ) **বাংলা CLDR উন্নয়ন এবং ইউনিকোড কনসোরটিয়মে জমা দেয়া (**Development of Bangla CLDR resource and submit to Unicode) ইউনিকোড কমন লোকাল ডাটা রিপোজিটরি (Unicode Common Locale Data Repository বা CLDR) হলো বিশ্বের প্রধান ভাষাসমূহের সহায়ক সফটওয়্যার হিসেবে মূল বিল্ডিং ব্লক যোগানদাতা। এটি স্থানীয় ইউনিকোড বিষয়ে বৃহত্তম ও প্রমিত তথ্য ভান্ডার। আন্তর্জাতিক কোম্পানিসমূহ তাদের সফটওয়্যার আন্তর্জাতিকায়ন ও স্থানীয়করণে এই তথ্য ভান্ডার ব্যবহার করে থাকে এবং ডিএলডিআর প্রদত্ত মান অনুসরণ করেন । বাংলা ভাষার ক্ষেত্রেও এই উপাংশের মাধ্যমে সিএলডিআর ভান্ডার উন্নয়ন ও প্রমিতকরণ করে তা ইউনিকোড কনসোর্টিয়মে জমা দেয়া হবে এবং অনুমোদনের প্রয়োজনীয় পদক্ষেপ নেয়া হবে।

		গ) ইন্টিশন: সার্ভিস ডেলিভারি প্লাটফর্ম (Integrated platform for Bangla Text, Speech and image related service distribution)প্রকল্পের সবগুলো সার্ভিস এক প্লাটফর্ম থেকে প্রদান করার জন্য www.bangla.gov.bd ওয়েবসাইটি চালু করা হয়েছে। এই ওয়েবসাইটি প্রকল্পের সবগুলোর সার্ভিসের মিলনবিন্দু হিসেবে কাজ করবে। এখন পর্যন্ত প্রকল্পের ১টি কম্পোনেন্টে (ধ্বনি) সর্বসাধারণের ব্যবহারের জন্য উন্মুক্ত করা হয়েছে। ৩টি কম্পোনেন্টের (সঠিক, বর্ণ, জনমত) পরীক্ষামূলক সংস্করণ অভ্যন্তরীণভাবে ব্যবহার করা হচ্ছে। ডেভেলপমেন্ট কার্যক্রম শেষে প্রকল্পের কম্পোনেন্টগুলো ক্রমান্বয়ে এই প্লাটফর্মে রিলিজ করা হবে। একজন ব্যবহারকারী ব্রাউজারের মাধ্যমে এই প্লাটফর্মে প্রবেশ করে উল্লিখিত সার্ভিসগুলো ব্যবহার করতে পারবেন। এছাড়াও এপিআই এর মাধ্যমে বিভিন্ন সরকারী সংস্থাকে চাহিদা অনুসারে জিটুজি সার্ভিস প্রদান করা হচ্ছে। ইতিমধ্যেই এটুআই ও ই-গভঃ সার্ট প্রকল্প থেকে উদ্দিষ্ট সার্ভিস গ্রহণ করছে।
৮	বাস্তবায়ন অগ্রগতি	২০২২-২০২৩ **অর্থ বছরের বাস্তবায়ন অগ্রগতি** **১। কৃত্রিম বুদ্ধিমত্তা প্রযুক্তিভিত্তিক বাংলা বানান সংশোধক:** বাংলা বানান সংশোধনে কৃত্রিম বুদ্ধিমত্তা প্রযুক্তিভিত্তিক বাংলা বানান সংশোধক ‘সঠিক’ গত ২১শে ফেব্রুয়ারি ২০২৩ তারিখে পরীক্ষামূলক ভার্সন উন্মুক্ত করা হয়েছে। এই সফটওয়্যার ব্যবহারে বাংলা একাডেমির প্রমিত বানান বিধি ও প্রমিত বানান অভিধানকে অনুসরণ করা হয়েছে। 'সঠিক' বানান সংশোধক হলো বাংলা ভাষার শব্দ, বাক্য কৃত্রিম বুদ্ধিমত্তার সাহায্যে সম্পাদনা করার সফটওয়্যার। এই সফটওয়্যার কেবল ভুল বানান শনাক্ত করবে তা নয়, বরং স্বয়ংক্রিয়ভাবে সংশোধনের পরামর্শ দেবে। সফটওয়্যারটি বিভিন্ন ধরনের এরর যেমন নন-ওয়ার্ড এরর, রিয়েল ওয়ার্ড এরর শনাক্ত করতে পারে। এ ছাড়া প্রায়ই যেসব বানান ভুল হয়, সেসব বানানসহ অসতর্কতাবশত লেখা 'টাইপো' দ্রুত শনাক্ত করতে পারে। তবে একটি শব্দের বানান শুদ্ধ হলেও ওই পরিস্থিতিতে শব্দটি ভুল হলে অ্যাপ্লিকেশন একে ভুল হিসেবে শনাক্ত করবে। কৃত্রিম বুদ্ধিমত্তা প্রযুক্তির মাধ্যমে তৈরি হওয়ায় এর রয়েছে কনটেক্সচুয়াল এরর চেকিংসহ বেশ কয়েকটি আকর্ষণীয় ফিচার; যা একাডেমি প্রকাশনা, মুদ্রণ জগৎসহ অনলাইনে শুদ্ধ বানানে লেখার অভিজ্ঞতা বদলে দেবে। **২। ৫টি ক্ষুদ্র নৃগোষ্ঠি ভাষার কিবোর্ড:** বাংলাদেশে ক্ষুদ্র নৃগোষ্ঠীদের মাতৃভাষার অনেকগুলো বিপন্ন। এই ভাষাগুলো রক্ষা করতে হবে এবং বিকশিত করার সুযোগ দিতে হবে। সরকার সম্প্রতি এসব বিপন্ন ভাষা ডিজিটাইজড করার উদ্যোগ নিয়েছে। এরই লক্ষ্যে ৫টি ক্ষুদ্র নৃগোষ্ঠি ভাষা, যথা- চাকমা, মারমা, ম্রো, সানতালি এবং তঞ্চঘ্যা এর কিবোর্ড প্রস্তুত করা হয়েছে। এরফলে এসব ভাষাভাষিগণ তাদের নিজস্ব ভাষায় কম্পিউটার বা মোবাইলে লেখালেখি করতে হবে। প্রস্তুতকৃত কিবোর্ডগুলো উদ্বোধনের অপেক্ষায় রয়েছে। এছাড়া কয়েকটি ক্ষুদ্র নৃগোষ্ঠী ভাষার কিবোর্ড প্রস্তুত কার্ক্রক্রম চলমান রয়েছে। **৩। বাংলা স্টাইল গাইড প্রণয়ন:** বাংলাদেশ কম্পিউটার কাউন্সিল কর্তৃক বাস্তবায়নাধীন ‘গবেষণা ও উন্নয়নের মাধ্যমে তথ্য প্রযুক্তিতে বাংলা ভাষা সমৃদ্ধকরণ’ প্রকল্পের আওতায় তথ্য প্রযুক্তিতে বাংলা বর্ণ/ক্যারেকটার

<table>
<tr><td></td><td></td><td>প্রমিতকরণ, বাংলা সর্টিং অর্ডার প্রমিতকরণ, ইঙ্গিতভাষা/সাইন ল্যাংগুয়েজ প্রমিতকরণ, বিরাম চিহ্ন প্রয়োগ-রীতি প্রমিতকরণ, ওয়েবে ও মুদ্রণ জগতে টাইপোগ্রাফিকাল স্টাইল শিট প্রমিতকরণ, প্রকাশনায় অন্যভাষার সঙ্গে বাংলা ভাষার ব্যবহারের রীতি (ইংরেজি, আরবি, সংস্কৃত, চাকমা প্রভৃতি) নির্ধারণ, বাংলা গ্লোসারি, টীকা ব্যবহারের রীতি (ফুটনোট ও এন্ডনোট), গ্রন্থপঞ্জি ও নির্ঘণ্ট লেখার নিয়ম নির্ধারণসহ লিপি ও ভাষাসংক্রান্ত প্রযুক্তিজগতে প্রয়োগযোগ্য সকল মান প্রমিতকরণের কাজ চলমান রয়েছে। এরি প্রেক্ষিতে গত ১৭ জানুয়ারি ২০২৩ তারিখ ঘটিকায় বঙ্গবন্ধু আন্তর্জাতিক সম্মেলন কেন্দ্রে কর্মদলসমূহের সদস্য ও অংশীজনবৃন্দ অংশগ্রহণে একটি কর্মশালা অনুষ্ঠিত হবে। এরই আলোকে ইতোমধ্যে চারটি বিষয়ের খসড়া চূড়ান্ত করা হয়েছে এবং গাইডলাইন আকারে ক্লোজগ্রুপে পুস্তক প্রকাশ করা হয়েছে। শীঘ্রই এসব মান উন্মুক্ত করা হবে।</td></tr>
<tr><td></td><td></td><td>আর্থিক অগ্রগতি: ১০০%।

বাস্তব অগ্রগতি: ১০০%।</td></tr>
<tr><td>৯</td><td>প্রশিক্ষণ</td><td>প্রযোজ্য নয়</td></tr>
<tr><td>১০</td><td>সেমিনার/কর্মশালা/আয়োজিত ইভেন্ট ও প্রতিযোগিতা</td><td>সেমিনার/কর্মশালার বিবরণ সহ অংশগ্রহণকারীর মোট সংখ্যা

আয়োজিত ইভেন্ট
<table>
<tr><th>SL</th><th>Date</th><th>Component Name</th><th>Total Participants</th></tr>
<tr><td>1</td><td>21/01/2023</td><td>Development of the Bangla Machine Translator (MT) (SD 14)</td><td>125</td></tr>
<tr><td>2</td><td>21/02/2023</td><td>Development of Bangla CLDR Resource and submit to Unicode (SD-13)</td><td>125</td></tr>
<tr><td>3</td><td>24/02/2023</td><td>Improvement of the National Keyboard (Bangla)(SD-10)</td><td>125</td></tr>
<tr><td>4</td><td>28/02/2023</td><td>Integrated Platform for Bangla Text, Speech and image related service distribution (SD-19)</td><td>125</td></tr>
<tr><td>5</td><td>05/03/2023</td><td>Improvement of the National Keyboard (Bangla)(SD-10)</td><td>125</td></tr>
<tr><td>6</td><td>09/03/202</td><td>Development of Bangla</td><td>125</td></tr>
</table></td></tr>
</table>

	3	Style Guide (SD-11)	
7	03/04/2023	Development of Bangla CLDR Resource and submit to Unicode (SD-13)	125
8	08/05/2023	Development of Sentiment analysis software in Bangla (SD-18)	125
9	24/06/2023	Development of Software for Disable People (SD-17)	125

আয়োজিত ইভেন্ট:

- **'বাংলার জন্য কৃত্রিম বুদ্ধিমত্তা ২.০' (AI for Bangla 2.0) প্রতিযোগিতা ও ফন্টের জন্য বর্ণ ডিজাইন প্রতিযোগিতা:**

কৃত্রিম বুদ্ধিমত্তার মাধ্যমে বাংলা লেখা পড়ে শোনাবে কম্পিউটার। এমনই এক উদ্ভাবনী প্রযুক্তি তৈরি করে বাংলায় কৃত্রিম বুদ্ধিমত্তা প্রযুক্তি বিষয়ক 'এআই ফর বাংলা ২.০' প্রতিযোগিতার চ্যাম্পিয়ন হয়েছে শাহজালাল বিজ্ঞান প্রযুক্তি বিশ্ববিদ্যালয়ের 'সাস্ট ১৯৫২'। 'সাস্ট ১৯৫২' এর বাংলা লেখা থেকে কথায় রূপান্তর ভাষা প্রযুক্তি বিষয়ক 'বাংলা নিউরাল টেক্সট টু স্পিস' ছাড়াও আরও সাতটি উদ্ভাবনকে চূড়ান্তভাবে পুরুস্কৃত করা হয়েছে।

তথ্য প্রযুক্তিতে বাংলা ভাষাকে সমৃদ্ধ করা এবং গবেষণায় উৎসাহ প্রদানের লক্ষ্যে আইসিটি বিভাগের বাংলাদেশ কম্পিউটার কাউন্সিলের 'গবেষণা ও উন্নয়নের মাধ্যমে তথ্যপ্রযুক্তিতে বাংলা ভাষা সমৃদ্ধকরণ' প্রকল্পের উদ্যোগে এ প্রতিযোগিতার চূড়ান্ত ফলাফল ঘোষণা করা হয়।

রাজধানীর হোটেল ইন্টারকন্টিনেন্টালে আয়োজিত এক অনুষ্ঠানের মাধ্যমে উক্ত প্রতিযোগিতার বিজয়ীদের হাতে পুরুস্কার ও সম্মাননা প্রদান করা হয়। এতে প্রধান অতিথি হিসেবে উপস্থিত ছিলেন আইসিটি ডিভিশনের সচিব জনাব মোঃ সামসুল আরেফিন। বাংলাদেশ কম্পিউটার কাউন্সিলের নির্বাহী পরিচালক রণজিত কুমারের সভাপতিত্বে অনুষ্ঠিত এ অনুষ্ঠানে বিশেষ অতিথি হিসেবে উপস্থিত ছিলেন প্রতিথযশা শিক্ষাবিদ ও প্রযুক্তিবিদ ড. মুহম্মদ জাফর ইকবাল।

এছাড়া একই প্রকল্পের আওতায় বাংলা ডিজাইন ফন্ট প্রতিযোগিতার ফলাফলও ঘোষণা করা হয়। প্রথম পুরস্কার হিসেবে এক লক্ষ টাকা প্রাইজমানি জিতেছেন জনাব কাজী মোঃ মহসিন। এছাড়া, শহীদ শরিফ রাসেল। মোঃ আনোয়ার হোসেন, সাজেদুর রহমনা সবুজ, মোঃ আলআমিন, ময়দুল হাসান রাসেল, মোঃ মনজুর হোসেন, ওসমান হাতা মিরন, মোল্লা শরীফ এবং মোঃ জাহিদুল ইসলামকে সর্বমোট দুই লক্ষ পয়ত্রিশ হাজার টাকার প্রাইজমানি ও সম্মাননা প্রদান করা হয়।

- **অমর একুশে বইমেলায় মাসব্যাপি অ্যাক্টিভেশন কার্যক্রম:**

প্রতিবছরের ন্যায় এবারও প্রকল্পের উদ্যোগে অমর একুশে বইমেলায় স্টল স্থাপনের মাধ্যমে মাসব্যাপি অ্যাক্টিভেশন কার্যক্রম পরিচালনা করা হয়। স্টলে প্রকল্পে আওতায় উন্নয়কৃত সফটওয়্যার সমুহের পরীক্ষামূলক ভার্সন সর্বসাধারণের ব্যবহারের জন্য উন্মুক্ত রাখা হয় এবং ব্যবহারকারীদের মতামত গ্রহণ করা হয় যা পরবর্তীতে সফটওয়্যারকে অধিকতর ব্যবহারকারীবান্ধব করার ক্ষেত্রে ভুমিকা রাখবে।

- **বাংলা স্টাইল গাইড বিষ্যক কর্মশালাঃ**

তথ্যপ্রযুক্তির সঙ্গে মানুষের মুখের ভাষার সম্মিলনের প্রথম শর্ত হলো ভাষার রীতি এবং ভাষা ব্যবহারের উপাদানসমূহকে প্রমিতকরণ করা। তথ্য ও যোগাযোগপ্রযুক্তি বিভাগের বাংলাদেশ কম্পিউটার কাউন্সিল (বিসিসি) তথ্যপ্রযুক্তিতে বাংলা ভাষা ব্যবহার প্রমিতকরণের কর্মসম্পাদনে ভূমিকা রাখছে। এরই ধারাবাহিকতা গবেষণা ও উন্নয়নের মাধ্যমে তথ্যপ্রযুক্তিতে বাংলা ভাষা সমৃদ্ধকরণ প্রকল্প, ১৬টি কম্পোনেন্টের মাধ্যমে বাংলা ভাষাসংশ্লিষ্ট প্রায় ৪০টি সার্ভিস, টুলস ও রিসোর্স তৈরির কাজ করছে। এর মধ্যে বাংলা স্পেল চেকার, ওসিআর, স্পিচ টু টেক্সট, টেক্সট টু স্পিচ, ম্যাশিন ট্রান্সলেশনের মতো সফটওয়্যার রয়েছে (www.bangla.gov.bd দ্রষ্টব্য)। সফটওয়্যার ডেভেলপ করার পাশাপাশি কম্পোজ ও 'ফন্ট ভেঙে যাওয়া'সহ কম্পিউটারে বাংলা

		লিখনের নানা প্রচলিত সমস্যা দূরীকরণের জন্য কিবোর্ড, কনভার্টার, ফন্ট তৈরি কার্যক্রমে চলমান রয়েছে। এছাড়াও বিদ্যমান প্রযুক্তিগুলো পর্যালোচনা সমস্যা নিরসনের জন্য প্রয়োজনীয় গবেষণা ও গাইডলাইন তৈরির জন্য 'বাংলা স্টাইল গাইড' কম্পোনেন্টের কার্যক্রম শুরু হয়েছে। রাজধানীর বঙ্গবন্ধু ইন্টারন্যাশনাল কনফারেন্স সেন্টারে এ বিষয়ক একটি কর্মশালার আয়োজন করা হয়। যেখানে দেশের বিজ্ঞ শীক্ষাবিদ, গবেষক, লেখক, সাংবাদিকসহ অনেকে অংশগ্রহণ করেন।
১১	**প্রকল্পের চলমান কিছু কার্যক্রমের স্থিরচিত্র**	 **Figure 1: 'বাংলার জন্য কৃত্রিম বুদ্ধিমত্তা ২.০' (AI for Bangla 2.0) প্রতিযোগিতা ও ফন্টের জন্য বর্ণ ডিজাইন প্রতিযোগিতা প্রতিযোগিতার পুরস্কার বিতরণী অনুষ্ঠানে সম্মানিত অতিথিদের সাথে বিজয়ীরা।** **Figure 2: রাজধানীর বঙ্গবন্ধু ইন্টারন্যাশনাল কনফারেন্স সেন্টারে স্টাইল গাইড বিষয়ক কর্মশালায় উপস্থিত দেশের বিজ্ঞ শিক্ষাবিদ, গবেষক, লেখক, সাংবাদিকসহ গুণীজন।**

Figure 3: অমর একুশে বইমেলায় মাসব্যাপি অ্যাক্টিভেশন কার্যক্রমে বাংলার স্টলে মাননীয় আইসিটি প্রতিমন্ত্রী জনাব জুনায়েদ আহমেদ পলক

Figure 4: রাজধানীর আগারগাঁওয়ে আইসিটি ভবনে 'বাংলাদেশের নৃগোষ্ঠী ভাষার ফন্ট ও কিবোর্ড নির্মাণ' শিরোনামে আলোচনা সভা অনুষ্ঠিত হয়

Figure 5: Further improvement of Bangla OCR developed by ICTD & integrating hand writing recognition system কম্পোনেন্টের আর্কিটেকচার সফটওয়ারের বৈশিষ্ট্য ও রূপরেখা পর্যালোচনার জন্য ঢাকা বিশ্ববিদ্যালয়ের সিনেট ভবনের সম্মেলন কক্ষে বিশ্ববিদ্যালয়ের ভাষাবিজ্ঞান বিভাগের

Figure 6: Bangla to IPA Automatic Converter (SD-22) কম্পোনেন্ট বিষয়ে সেমিনার কক্ষ, আধুনিক ভাষা ইনস্টিটিউট, ঢাকা বিশ্ববিদ্যালয়ে এ সম্ভাব্য ব্যবহারকারী, উন্নয়ন সহযোগী ও গবেষকদের নিয়ে একটি সেমিনার অনুষ্ঠিত হয়।